# The Political Economy Causes of the Soviet Great Famine, 1932–33*

Andrei Markevich†, Natalya Naumenko‡and Nancy Qian§

(Incomplete)

We document several new facts about the Soviet Great Famine, 1932–33. First, there was no aggregate food shortage. Second, mortality rates were highly unequal across regions. Third, mortality rates were unrelated to regional per capita food production, but positively associated with Ukrainian population share, even after controlling for weather, contemporaneous food production, demographic and pre-famine political differences. Finally, political loyalty to the the regime was positively associated with famine mortality, collectivization and procurement intensity, but only in regions where ethnic Ukrainians resided. These findings go against traditional explanations of famine based on aggregate food production shortfalls, and are consistent with Sen's political-economic theory of famine. Our estimates imply that ethnic bias against Ukrainians in Soviet policy explains 77% of famine deaths in Russia, Ukraine and Belarus, and 92% in the Ukraine.

**JEL:** N4, P2 **Keywords:** Strategic Killing, 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 CERGE Applied Economic Workshop, 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, Northwestern Economic History Lunch and Northwestern Development Lunch. 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

In 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] Mortality rates were very unequal across ethnic groups. 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] In Kazakhstan, where most ethnic Kazakhs lived, 1.3 to 1.5 million out of a population of 6.9 million died (Cameron, 2018). Until the late 1980s, the Soviet government denied that any famine occurred. Even now, the causes of the famine are highly controversial.

The traditional explanation of famine is based on the occurrence of aggregate food shortage. This could be due to bad weather, or overpopulation (Malthus, 1798). In the case of the Soviet Famine, historians argue that bad weather and pre-famine economic policies reduced productivity and caused poor harvests (e.g., Davies and Wheatcroft, 2004; Kondrashin, 2008; Kotkin, 2017). Some historians emphasize that mortality rates were high for all ethnicities (Kondrashin, 2008). In contrast, Sen (1981) famously provides an alternative political economy explanation. He argues that famines are not caused by aggregate food shortages. Instead,

[1] Conquest (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).

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

[3] The most cited total famine death toll for the U.S.S.R. is seven million (footnote 1). There are no systematic data on ethnic-specific mortality rates. For the Ukraine, where 80% of the population were ethnic Ukrainians, Meslé et al. (2013) estimates excess deaths to be 2.6 million, while Rudnytskyi et al. (2015) estimates it to be 3.9 million. If one conservatively assumes that famine mortality rates were equally distributed across ethnic groups within the Ukraine and that no ethnic Ukrainians died from famine outside the Ukraine, then ethnic Ukrainian deaths constitute 30% ($0.8 \times 2.6/7 = .3$) to 45% ($0.8 \times 3.9/7 = .45$) of the total famine deaths. Note that the results in this paper refutes both assumptions. Thus, these are lower bound estimates.

they are due to unequal distribution where the political rulers control the food and oppressed populations starve. In the Soviet context, many have gone as far as arguing that Stalin and the central government waged an intentional "war" or "terror" "by starvation" to crush dissent from the peasantry which was particular pronounced among certain ethnic minorities such as the Ukrainians (e.g., Conquest, 1986; Graziosi, 2015; Ellman, 2007; Mace, 2004; Snyder, 2010). The Ukrainian Parliament declared the *Holodomor* (death by starvation) a genocide in 2003 and the European Parliament declared it a "crime against humanity" in 2008.

The main goal of our study is to make progress in understanding the causes of the Soviet Great Famine and the degree to which Soviet policies contributed to famine mortality. Despite the attention and controversy surrounding these questions, there is little systematic empirical analysis. Most arguments are based on descriptive or narrative evidence. The main difficulty is the lack of representative sub-national data that would allow the econometrician to evaluate competing hypotheses.

The primary contribution of this paper is to address this limitation by constructing what is to the best of our knowledge the largest and most comprehensive data on the Soviet Union in the interwar period. We use data from archival and published sources, as well as geo-spatial weather and agricultural suitability data to form province and district-level panels that cover the three most populous republics of the U.S.S.R.: Russia, Ukraine and Belarus. Our data include the main variables that have been discussed in the famine literature. The large sample size and panel structure allow us to rigorously analyze the data and distinguish between the prevailing hypotheses by controlling for confounding factors and exploring a rich set of heterogeneous treatment effects.

We proceed in several steps. First, we conduct a food accounting exercise. We use data on national grain production and total population to document that reported per capita production was approximately six times the level necessary to avoid deaths from starvation. This implies that famine was due to inequality in food distribution across the population. The

Soviet centrally planned economy procured food from rural areas to give to urban areas and to export. Using data on total procurement and rural retention, we also document that per capita rural retention was approximately five times more than the level necessary to avoid famine mortality. This implies that inequality not only existed between urban and rural areas, but also across rural areas. While we note that there may be measurement error in the official grain statistics, there are no accounts of measurement error large enough to explain the entire surplus. Our results of sufficient aggregate food supplies hold if we use Davies and Wheatcroft's (2004) revised estimates. They are true for each of the three republics, for which we have mortality data, as well as for the entire U.S.S.R.

Second, we document spatial patterns between rural famine mortality rates and food production.[4] We find that mortality is positively associated with per capita food production during the famine if we do not control for Ukrainian population share. When we control for Ukrainian population share, the latter is strongly positively associated with famine mortality and food production is uncorrelated with famine mortality. All regressions control 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.

These results go against the theory that the famine occurred despite the best intentions of the Soviet government – i.e., the government was unable to ship adequate amounts of food relief from regions that had good harvests to regions that had poor harvests. If that were the case, mortality rates should be higher in less productive areas (and highest in urban areas). Thus, the results imply that government procurement and distribution of food is important for explaining the famine. These estimates also support the belief that understanding the

[4]We do this partly because there is no ethnic-specific mortality or food production data.

causes of Ukrainian famine mortality is central to understanding the causes of the famine. They indicate that for two regions with the same level of per capita production during the famine, the region with more Ukrainians suffered higher mortality. To address the possibility that grain production may be misreported, we use both reported grain production as well as production predicted by weather and agricultural suitability. To address the possibility that mortality is misreported, we also use natality data following the logic that birth rates were lower in regions that suffered famine more intensely. The results are consistent.

Third, motivated by the previous patterns in the data, we examine Ukrainian mortality more closely. We find that the positive association between famine mortality and Ukrainian population share is robust to alternative measures of ethnic composition, to omitting the Republic of Ukraine (the Ukrainian S.S.R.) from the sample, and to controlling for demographic structure as well as pre-famine political and economic characteristics. These results imply that Ukrainian famine mortality is not an artifact of pre-existing conditions or historical features of the Ukrainian S.S.R.. which may have made the areas where Ukrainians lived a target for Soviet policy. The fact that the pattern is present even if we exclude the Ukraine implies that ethnicity, rather than the republic of residence, is what matters.

Fourth we attempt to provide more direct evidence on the role of Soviet policies on famine mortality. We begin by estimating heterogeneous treatment effects with proxies for loyalty to the Soviet regime or zealousness in implementing its policies (i.e., the triple interaction effects of Ukrainian population share, the famine year dummy variable and proxy variables). We find that the positive association between Ukrainian population share and famine mortality is 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 in regions that sent more delegates to the 1930 Party Congress to vote for forcing collectivization (a measure of commitment to Stalin and/or collectivization of the local Party elite). These results imply that Soviet policies or the zealousness with which

they were implemented systematically targeted regions with a higher share of ethnic Ukrainians. Interestingly, the estimates for voting delegates show that the number of both ethnic and non-ethnic Ukrainian delegates was associated with higher mortality in the regions with a higher Ukrainian population share , but 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 the documentary evidence that ethnically Ukrainian bureaucrats in the regions populated by ethnic Ukrainians were more likely to protest Soviet policy during the famine.

To delve further into the mechanisms, we repeat the heterogeneous effects estimates with collectivization as the dependent variable. Agricultural collectivization was the main rural Soviet economic policy implemented at the time. It was meant to boost production by achieving economies of scale, which was useful for mechanization, and increase the government's control over food distribution. Consistent with the view that collectivization was an important contributor to famine mortality, the heterogeneous treatment estimates for collectivization mirror those for mortality. We also find similar patterns when we examine planned and reported procurement rates as the dependent variable. The former re-enforces the interpretation that high Ukrainian famine mortality was an outcome of deliberate Soviet policy because it demonstrates that the regime had procured more from Ukrainian-populated areas. In contrast, we find that loyalty to the regime and political zealousness were negatively correlated with mechanization for regions with a higher share of ethnic Ukrainians. Thus, Ukrainians experienced the negative, but not the positive side of collectivization.

Finally, we 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 show that the baseline result of excess Ukrainian mortality during the famine is true when we control for 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 also across smaller ones within the large units. With province-year fixed effects, we are also able to control for region-level features of the famine (e.g. local politics) which historians often refer to explain variation in mortality rates. We also demonstrate that the discrete downward jump in mortality rates when crossing the border from the Ukraine to Russia disappears if we control for Ukrainian population share. This is consistent with the view that food was procured systematically from ethnic Ukrainians.

A back-of-the-envelope calculation shows that ethnic bias in Soviet policies against Ukrainians explains 92% of total famine mortality in the Ukraine and 77% in Russia, Ukraine and Belarus.

We supplement the main analysis with an additional exercise. We show that mortality was not associated with Ukrainian population share during the 1892 famine, 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).

The findings support Sen's (1981) thesis that large 20th century famines were caused by political economy factors and unequal food distribution rather than aggregate food shortages. Moreover, our results support scholars of the Soviet famine which attribute high famine mortality to the Soviet regime's strategic decisions, and show that understanding high Ukrainian mortality and ethnic bias is inextricable from understanding the causes of this famine. Note that our study takes as given the factors which contributed to grain production in 1932 or those that made the population particularly vulnerable to grain shortage. To the extent that such factors are related to state policy (e.g., the reduction of livestock during the first year of collectivization), our interpretation understates total policy effect.

To the best of our knowledge, our study is the first systematic and rigorous empirical

analysis of the causes of the Soviet Great Famine. We contribute to the large literature on the causes of famines in development economics and economic history. Earlier works have examined contexts such as China (e.g., Meng et al., 2015), India (e.g., Sen, 1981), and Ireland (e.g., Ó Gráda, 1999).[5] We add to these by documenting a new mechanism, ethnic bias, in a new and important context, the Soviet Union. We add to studies from historians on the causes of the Soviet Famine discussed earlier. These studies have failed to be conclusive on the role of state policy because there is no direct documentary evidence of Stalin having "ordered" a famine. The empirical analysis in our study thus addresses an important limitation by allowing us to infer policy and reject alternative explanations with the data.

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 in all human history. We complement macro calibrations of Soviet industrialization policies by Allen (2003) and Cheremukhin et al. (2017).[6] 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 S.S.R..

Finally, we add to the political economy literature on ethnic conflict (Chassang and Padró i Miquel, 2009). The Stalinist regime, where all groups suffer persecution, but some much more than other is consistent with Padró i Miquel, 2007, which argues that rulers of one group strategically oppress members of other groups to prevent leaders from other groups from credibly committing to not oppress members of the first group if ever in power.[7] Our findings are in line with the theoretical predictions for mass killings and genocide from Esteban et al. (2015) that strategic killings are more likely with high levels of natural resources and low labor productivity. The fact that these events unfolded ten years after the founding of the U.S.S.R. in 1922 and were a part of the policy of centralization is consistent with the

[5] See Ó Gráda (2009) and Alfani and Ó Gráda, eds (2017) for an overview.

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

[7] Also, see Blattman and Miguel (2010) for an overview of the Civil War literature and ethnic conflict in that context.

view that ethnic homogenization is complementary to nation building (Alesina and Reich, 2015).[8] Our findings in the context of the U.S.S.R., which had high state capacity, complement Heldring (2020), which provides empirical evidence that state capacity is an important factor in mass killings in the Rwandan context. Our results add to the understanding of the source of current Ukrainian-Russian tensions, which recent studies have found to be related to economic (Korovkin and Makarin, 2019) and political outcomes (Rozenas and Zhukov, 2019).

This paper is organized as follows. Section 2 summarizes the historical background. Section 3 presents the food accounting exercise. Section 4 presents the main province-level analysis. Section 5.6 presents the district-level analysis. 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.[9] 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-

[8] Ou 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.

[9] Davies 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 tremendous 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 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.

## 2.2 Food Production and Distribution

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 secure its industrialization spurts. For that, 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 and which the government would be able to control directly. 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 out of these two 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). That often resulted in payments per labor day of a collective-farmer 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 focus on 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. We show that a drop in the number of horses because of collectivization was not statistically larger in regions with higher share of ethnic Ukrainians. Unfortunately, we do not have the data to study other channels.

## 2.3 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 (Graziosi, 2015). 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 1923, to ensure loyalty of ethnic minorities to the new regime, the Soviet government first launched a policy of indigenization (korenizatsiya), which aimed to promote national languages and culture in regions where minorities represented local majority. But fears that such policy also contributed to the rise of nationalism periodically caused delays in its implementation in the 1920s. A start of collectivization strengthened centralization of the Soviet regime that was in contrast to the indigenization policy (Graziosi, 2015; Martin, 2001). As many historians note, there is no explicit evidence that Stalin "ordered" the famine.[10] 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. [11]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, Stalin received multiple evidence indicating the reluctance of individual Party leaders in Ukraine at all levels to cooperate with the starvation of so many peasants.[12] 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 (Kotkin, 2017). In the end, concerns about national opposition to the regime became particular strong that led to a de facto termination of the indigenization policy in the autumn of 1932 (Graziosi, 2015; Martin, 2001). On December, 14, 1932, the Soviet government issued a classified decree in which the government insisted on complete fulfillment of grain procurement tasks in Ukraine,

[10]E.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.

[11]For example, he was known to widely recruit members of ethnic minorities into his secret police (Gregory, 2009); with rare exceptions first party secretaries in national republics also did not belong to the titular nation in the region Martin (2001).

[12]E.g. in a letter to his deputy Lazar Kaganovich from August, 11 1932, Stalin mentioned that party district committees in about fifty districtes in Ukraine had spoken out against state procurement quotas; he expressed his concerns that the Soviet government 'could lose Ukraine' Davies et al., eds, 2003.

the North Caucasus and the Western region and required to arrest those communists on the ground and local officials who failed these tasks. Simultaneously, the government acknowledged that the previous policy of indigenization often contributed to promotions of former Ukrainian nationalists in the communist party and local state apparatus and blamed them for the sabotage of state procurement policy. [13] We will later explore the extent to which such ethnic allegiances played a role.

# 3 Food Accounting

The goal of this exercise is to estimate per capita food production and per capita food requirement for the Soviet Union and examine whether production was sufficient to avoid the famine. The most important source of food, which was the target of government procurement, was grain. The data we use come from archival and published sources.[14] In the paper, we will discuss the data as they become relevant.

We start with official data on population, production and procurement (rows (1) to (3) and (5) in Table 1). Row (5) presents reported procurement as a share of production and shows that it increased over time from 14.9% in 1927 to 30.7% in 1939, with the peak being the famine years, when they were 32.9% and 27.2% in 1931 and 1932. Note that food produced in a given year is used to feed the population the following year. Thus, we focus on production in 1932 to study mortality in 1933.

Table 1 row (7) shows that per capita grain production in 1931 and 1932 were 433kg and 428kg, lower than the previous four and subsequent three years. However, production in

[13] E.g. the decree stated: "... frivolous, not arising from the cultural interests of the population, not Bolshevik 'Ukrainization' of almost half of the regions of the North Caucasus, in the complete absence of control over the Ukrainization of schools and the press by the regional authorities, gave legal form to the enemies of the Soviet government to organize resistance to the activities and tasks of the Soviet government by the kulaks, officers, re-emigrant-Cossacks, members of the Kuban Rada, etc." (Danilov et al., eds, 1999-2006Vol. 3, Document 226).

[14] See Appendix A provides the source of each variable.

1929, when there was no famine, was only slightly higher at 465kg. We convert grain from kilograms to calories per day using calories per one kilogram of Russian grain estimated by Lositskij (1920) in row (8).

In rows (12) and (13), we present two levels for caloric requirements. The first is the "business as usual" measure that maximizes labor productivity and healthy child development. These assume that all rural prime age males do heavy labor and all urban prime age males do light work. We use official Soviet estimates for caloric requirements from Lositskij, ed (1928), which are higher than the estimates for other countries or international standards. They are 3,750 and 2,750 calories per day for the two types of labor. We adjust the requirements by the demographic composition (e.g., age, gender) using Soviet official data on relative requirements Lositskij (1926) and official data on demographic structure.

The second caloric requirement is for "staying alive". For this, we use the 900 calories required for prime age males provided by Dasgupta and Ray (1986). We adjust it in the same way as the first threshold to account for demographic composition.

Row (12) shows that for business as usual, the U.S.S.R. required 2,439 to 2,437 calories per capita during 1931 and 1932. Per capita production in row (7) for these years, 3,716 and 3,675 calories, are 152% and 151% higher than these requirements. Row (11) shows that to avoid mortality, the U.S.S.R. required 621 and 622 calories on average. Production was 599% and 591% higher than these requirements.

It is important to briefly discuss the reliability of the historical data. The raw data used to generate the aggregate 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. The government used various cross-check procedures to avoid manipulation at lower levels of government. One important exception, however, is the aggregate grain production in the early 1930s. Grain production was viewed as one of the key and public indicators for Soviet

economic health, which reflected the success of the new Bolshevik regime. As such, it was controversial and there exists much debate over the accuracy of reported aggregate grain production (e.g., Wheatcroft and Davies, 1994; Davies and Wheatcroft, 2004). To address this, we use Davis and Wheatcroft's (2004) adjusted estimates as a lower bound for production in rows (4) and (9). They are lower than official estimates, but do not overturn the point of sufficient aggregate production for avoiding famine. [15]

In the centrally planned food distribution system, food is procured from rural areas to urban areas and for export, and it is known that famine mortality rates were lower in urban areas.[16] To investigate whether aggregate procurement is sufficient for explaining famine mortality without additional inequality in food distribution across the rural population, we calculate average rural grain retention (row 10). We use data on the reported amount of grain procured by the central government.[17] We convert retention into calories in row (11).

These calculations show that average rural retention was 128% and 141% of the higher threshold and 503% and 553% of the lower threshold for avoiding mortality. Thus, aggregate procurement of food to supply urban areas and exports cannot explain the famine alone. For the famine to have occurred, there must have been unequal food distribution across the rural population.

Table 2 repeats the exercise for the three republics for which we have mortality data: Russia, Ukraine and Belarus. Rural per capital food retention (rows 9, 20 and 31) during the famine is always higher than food requirement. It is interesting to note that the surplus is the largest in the Ukraine, which suffered the highest famine mortality among the three republics.

There are several caveats to keep in mind. The first the concern of measurement error in

[15]Tauger (2001) argues that the true 1932 grain harvest was meager 50 million tons that is the most conservative estimate in the literature. This transforms into 2,630 calories per day that is still above our estimate of the "business as usual" caloric requirement.

[16]Grain exports during 1931-32 were approximately five million metric tons (7% of total production) (Nove, 1969).

[17]These data are counts of actual procurement and not estimates. In general, historians view procurement figures as of much better quality because procured grain was phisically observable by state officials. See the Data Appendix for precise references.

the data. The second is that some of the production may be wasted (e.g., due to poor storage). Lositskij (1920) estimates waste for wheat and rye to be approximately 5% in Russia. We do not know of estimates for the early 1930s. While these factors may be relevant, we have not heard any reliable estimates of mis-reporting that would overturn the main point that the famine would not have occurred if food was equally distributed across the population, and in particular, the rural population.

# 4 Spatial Inequality

Given that aggregate production and rural retention were too high to cause mortality, this section investigates the notion that famine intensity was unequal across the population.

## 4.1 Mortality

We observe mortality rates at the province level for 1923 to 1940. We adjust these and all other province variables discussed later in the paper to the 1932 provincial units. Our sample includes nineteen provinces from the three largest republics of the Soviet Union: Belarus, Ukraine, and most of Russia. Altogether, the sample includes 84% of the 1926 Soviet population and 88% of the 1928 Soviet grain production.[18] The average province has 6.5 million people in 1926.

Note that the data misreporting issues we discussed in the previous section all focus on aggregate (i.e., national) grain production and we know of no claims that misreporting of any variable was correlated with ethnic composition at disaggregated levels of government. Nevertheless, we will carefully keep the possibility of mis-reporting in mind and address it

[18]The 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 our panel, Belarus and Ukraine stand for one region each. In our district-level analysis, we explore variation in mortality within Ukraine and within provinces of the Russian Federation.

when relevant in the paper.

Figure 3a plots mortality rates, the number of deaths divided by total population, in each province and year, from 1923 to 1940. It shows that mortality rates for the Soviet Union are reasonably constant over time at approximately twenty per 1,000, but spike in 1933 to nearly forty per 1,000. The figure also presents mortality rates over time for each republic. We see that mortality rates are usually higher in Russia than the Ukraine or Belarus, but the spike in famine mortality is the highest for the Ukraine, where it increased from approximately eighteen per 1,000 to nearly 60 per 1,000 in 1933. In contrast, Russia experienced an increase from approximately 22 per 1,000 to 30 per 1,000, while mortality in Belarus remained relatively constant at approximately fifteen per 1,000.

Figure 1a maps excess mortality in 1933, the year with the highest famine mortality, for the provinces in our sample.[19] The map illustrates significant geographic variation. The Ukraine and the southern provinces of Russia suffer much higher mortality rates than other regions.

Another way to examine unequal mortality during the famine is to examine the variation in famine mortality across space and see if it increases during the famine. Figure 4a plots the standard deviation in mortality rates across provinces normalized by mean mortality for 1923—1940. It shows that there is always variation in famine mortality across provinces, but it is constant over time, with the exception of the famine, when it increases dramatically. We can only examine the whole sample because of the limited number of provinces per republic.

These figures are consistent with the aggregate food accounting exercise by showing that there was significant inequality in famine severity.

[19] We calculate 1933 excess mortality as mortality in 1933 minus mortality in 1928.

## 4.2 Natality

To address the concern that the mortality data may be misreported we repeat the estimates with natality data, the logic being that live births should be decreasing in the famine severity.[20] We present figures analogous to those shown for mortality. Figure 3b shows that average natality rates begin to decline around 1928 and reach the lowest levels in 1933 and 1934. The decline is the largest for the Ukraine. Interestingly, note that national birth rates remained low in 1934, when mortality rates had already recovered. This is consistent with the fact that those who were starving in 1933 were unable to become pregnant and give birth in 1934.

Figure 4b plots the standard deviation normalized by the mean over time. It shows that the variation increases dramatically during the famine. The natality patterns correspond to the mortality patterns and show spatial inequality in famine intensity.

# 5 Explaining Unequal Famine Intensity

## 5.1 Food Production and Mortality

Our earlier results reject the notion that there was a deficit in aggregate food production. Nevertheless, weather and natural conditions could be the key culprits if conditions were very unequal in 1932 such that some regions produced surplus while others suffered harvest failures, and the government was unable to distribute food from productive areas to unproductive areas. To test this possibility, we investigate the spatial correlation between per capita food production and famine mortality rates. If the famine was due to unequal natural conditions, we should observe a negative association between mortality and production.

We estimate the following equation, where we assume that food produced in a given year is largely used to feed the population in the next calendar year.

[20] Starvation is negatively associated with the probability of pregnancy (and marriage), and is positively associated with probability of miscarrying and stillbirths (Dyson and Ó Gráda, eds, 2002).

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

Mortality rate in province $i$ in year $t+1$ is a function of: per capita grain production, $Grain_{it}$; its interaction with a dummy variable that equals one in the famine year, $Famine_t$; province fixed effects, $\eta_i$; and year fixed effects $\delta_t$. 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.[21] 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 3 column (1) uses reported grain data. It shows that in non-famine years, grain productivity is uncorrelated with mortality rates ($\hat{\beta} \approx 0$), but in famine years, the association is *positive* ($\widehat{\beta + \gamma} \approx 0.124$). This contradicts traditional explanations that attribute starvation to low food production. The positive association between grain production and mortality in the famine years is similar to the Chinese Great Famine (1959-61), which Meng et al. (2015) document and hypothesize is the outcome of the centrally planned procurement system. However, we will later show that this correlation becomes zero when we control for Ukrainian population share, which implies that the underlying mechanisms for the Soviet Famine is quite different from its Chinese counterpart.

In column (2), we address the concern of measurement error in the reported grain data by predicting grain with time-varying weather and time-invariant agro-climatic conditions. We

[21] We control for 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.

use monthly temperature and precipitation data from Matsuura and Willmott (2014) together with province-level grain production for years prior to the establishment of the communist regime, 1901 to 1915, to predict weather-driven production during our sample of interest.[22] The coefficients have similar signs, but are less precise.

Based on these estimates, we conclude that there is no evidence to support the theory that the spatial patterns of the famine were due to weather conditions, or the inability of the government to transport grain from surplus production regions to deficit production regions.

## 5.2 Ethnic Ukrainian Share

Motivated by the debate over potential ethnic bias against Ukrainians discussed in the Introduction and the historical background section, as well as the reversal in the ranking of mortality rates from non-famine to famine years shown in Figure 3b, this section examines the relationship between Ukrainian population share and famine mortality. First, we ask whether mortality in regions with higher shares of ethnic Ukrainians was higher during the famine by estimating the following regression.

$$mortality_{it+1} = \alpha + \beta Ukrainian_i \times Famine_t + \Gamma X_{it} + \eta_i + \delta_t + \varepsilon_{it}, \tag{2}$$

where $Ukrainian_i$ is the share of ethnic Ukrainians amongst the rural population of province $i$ in 1926. Since this is a time-invariant measure, the uninteracted term is absorbed by the province fixed effects.

Our main measure of ethnic composition is from the 1926 Soviet census, which is commonly viewed as one of the best Soviet censuses and reliable (Andreev et al., 1998). Russians were the ethnic majority and constituted 53% of the 1926 Soviet population; Ukrainians were by far the largest ethnic minority and constituted 21% of the Soviet population. According

[22] See the Appendix Section B.

to the 1926 population census, in Ukraine, ethnic Ukrainians comprised 23.2 out of 29.2 million citizens, and additional 5.6 million ethnic Ukrainians lived outside of Ukraine. The second-largest ethnic minority, Belorussians, constituted only 3% of the population.[23]

Grain production and its interaction with famine are now part of the vector of controls, $X_{it}$. Table 3 columns (3) and (4) controls for reported and predicted grain productivity, respectively. The estimates for the interaction of Ukrainian population share and the famine dummy are similar. We will focus on column (4), our baseline, for brevity. The interaction coefficient is 0.051 and statistically significant at the 1% level. Moreover, the interaction of grain production and the famine dummy is reduced in magnitude and statistically zero. Note that the standardized coefficients of the interaction of grain and famine in columns (3) and (4) are much smaller than those for the interaction of Ukrainians and famine. Thus, even if they were precisely estimated, they would be economically less important than ethnic Ukrainian share.[24]

Several facts emerge from these estimates. First, mortality rates were systematically increasing with Ukrainian population share, even when comparing two provinces with the same level of food production and urbanization rates. Second, we note that the estimates for urban population share show that during non-famine years, mortality is lower in more urbanized provinces, which is consistent with the preferential treatment of factory workers in the food distribution system. Interestingly, there is no difference during famine years. Thus, urbanization does not play a large role in explaining the spatial inequality of famine mortality. Our later estimates will always control for urbanization and its interaction the famine dummy, but we will not discuss it again. Finally, the negligible estimate for grain production and its interaction with the famine dummy after controlling for Ukrainian population share implies that the frictions which explained the Chinese famine, namely, the information frictions in the

[23] Appendix Table A.1 lists ethnic groups most often mentioned in the famine literature.

[24] An alternative strategy to account for natural conditions is to directly control for weather variables in the regression. The results are very similar and available upon request.

central procurement system, cannot explain the Soviet famine. Instead, the more important factor is ethnic Ukrainian population share.

Taken literally, column (4) implies that 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, note that one standard deviation in 1933 mortality rates in our sample is 0.013 and one 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 0.825 standard deviation increase in mortality. This is a large effect.

Since the first signs of famine were documented after the 1931 harvest, in column (5), we define the famine dummy variable to equal one in 1931 and 1932. The interaction coefficient is smaller in size, but still large, positive and statistically significant at the 5% level. We examine the timing more explicitly in the next section.

Finally, we present a baseline that replaces the province fixed effects with an uninteracted Ukrainian variable. This allows us to observe the relationship between Ukrainian population share and mortality during non-famine years, and also addresses the concern that province fixed effects over control by absorbing relevant cross-section variation. Column (6) shows that the latter is not a concern as the interaction coefficient is identical to the baseline in column (4). The uninteracted Ukrainian coefficient is -0.007 and statistically significant at the 1% level. This is interesting as it shows that in non famine years, Ukrainian population share is negatively associated with mortality. It is only during the famine that mortality is higher in Ukrainian regions (the sum of the interaction coefficient and uninteracted coefficient is positive, $0.051 - 0.007 = 0.044$, is positive and statistically significant).

### 5.2.1 Dynamic Estimates

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 5 shows a striking pattern. Prior to the famine, from 1924 to 1931, there was little difference in mortality rates across regions.[25] However, regions with a higher share of Ukrainians began to experience higher mortality in 1932 (although the point estimate is statistically indistinguishable from zero) and this difference peaked in 1933. This pattern is consistent with historical evidence that there was a small famine after the 1931 harvest, which was greatly exacerbated after the 1932 harvest. Afterwards, from 1934 to 1940, regions with higher shares of Ukrainians had mortality rates similar to other regions.[26]

Henceforth, we will use the 1932 definition of famine and focus on its effect on mortality in 1933.

### 5.2.2 Alternative Measures of Ukrainian Population Share

The baseline uses Ukrainian share in the rural population because the famine was driven by agricultural policies targeted at the rural population. Table 4 Panel A Columns (2) to (5) show that our results 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

[25] Note that this is slightly different from Table 3 column (6), which shows that Ukrainian population share is negatively associated with mortality in non-famine years because the specification estimated here uses the baseline controls, which includes province fixed effects.

[26] The post-famine patterns 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).

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 (3) 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 (6), 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 0.048 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.[27]

Column (7) controls for the the triple interaction of latitude, longitude and the famine dummy to address the possibility that factors which can affect famine intensity such as social capital (e.g., Durante and Buggle, forthcoming) and Ukrainian population share may be correlated across space. Our estimate is similar with this control.

In column (8), we weight the regression by province-year population. The results are very similar. In column (9), we exclude most influential, according to Cook's Distance, observations. While this reduces the mean share of ethnic Ukrainians in the sample from 10% to 2%, the point estimate of the Ukrainians x Famine coefficient remains almost exactly the same – 0.048 and is still significant at the 1% level. In column (10), we exclude the Ukrainian S.S.R., where 78% of all Ukrainians in our sample reside. The Ukrainian interaction coefficient increases from 0.051 in the full sample (column 1) to 0.086 and is statistically significant at

[27] Our 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.

the 1% level. This means that the baseline results are not driven by a comparison of differences between the Ukrainian S.S.R. and other Soviet regions. Instead, they reflect systematic higher mortality rates for regions with a higher share of *ethnic* Ukrainians.

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.[28] Column (11) estimates our baseline specification for this earlier famine. We find that 1892 famine mortality is not associated with Ukrainian population share.[29]

Table 4 Panel B example natality as the dependent variable. The estimates are all negative and statistically significant at the 1% levels. Figure 6 plots the coefficients from the dynamic estimates. They trace the temporal patterns of the raw natality data shown earlier in Figure 5, which is consistent with Ukrainian population share being an important determinant in birth rates during the famine.

Table 4 presents Huber-White robust standard errors to address heteroskedasticity in parenthesis. However one may be concerned about spatial correlation in the standard errors. To address this, we also present Conley standard errors in square brackets. They are very similar. Henceforth, we return to the Huber-White robust standard errors for the province-level estimates.

[28] Volha Charnysh kindly shared 1885–1896 mortality and natality data with us, Charnysh and McElroy (2020).

[29] In addition to the estimates of the Tsarist famine shown earlier, we can also investigate whether our main results on excess Ukrainian mortality is driven by differences in cultural practices or historical institutions by directly controlling 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.3 shows that our results are robust if we add interactions of these variables with the famine dummy into the baseline specification.

## 5.3 Controlling for Policy and Political Factors

Table 5 additionally controls for the interactions of the famine year dummy with proxies for Soviet policy and loyalty to the regime, which influences the implementation of official policies. We examine measures that are widely believed to have contributed to the intensification of Soviet agricultural policies during famine. Column (1) restates the baseline for comparison. Column (2) controls for per capita grain production in 1928, the beginning of the first Five Year Plan and therefore a common reference for Soviet planning (e.g. see Wheatcroft, 2001). Graziosi (2015) summarizes that Soviet food requisitions "concentrated in the grain-producing areas" (p. 241).

Column (3) 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 and only until the end of Bolshevik rule. Approximately 60% of the population turned out to vote.[30]

Column (4) controls for the number of urban and rural Communist Party Members (averaged over 1922, 1927 and 1931) per one thousand individuals in each province. Party members were key 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, which we will discuss in the next section.

Column (5) 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 a showcase of support for collectivization, all delegates voted in the affirmative and the number of voting delegates can be interpreted as a proxy for loyalty to the regime or commitment to agricultural collectivization.

[30]We 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.

Panel A examines mortality as the dependent variable. Panel B examines natality as the dependent variable. In both cases, the main interaction coefficient for Ukrainians is very robust and always similar to the baseline. These results mean that higher famine mortality and lower famine natality 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 some of the policy interaction coefficient are statistically significant. We postpone the discussion to the next section on heterogeneous results, where we show that the average effects mask meaningful heterogeneity.

## 5.4 Heterogeneous Effects on Mortality and Natality

This section estimates the heterogeneous effects of loyalty to the regime or the zealousness in enforcing Soviet policies on famine mortality in areas with high Ukrainian population share. Table 6 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 of primary interest 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.

Panel A examines mortality. 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 have decreased 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. This is consistent with the Stalinist practice of rewarding loyal populations and penalizing disloyal ethnicities (Polyan, 2001). These results also imply that Stalin perceived ethnic Ukrainians as disloyal.

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 the urban population and mitigating the consequences of famine when they spilled over to urban areas.[31] Rural communists increase mortality only in areas with ethnic 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

[31] For 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).

leadership increased famine mortality, but only in regions with Ukrainians.

Next, we investigate role of the ethnicity of the Party delegate. This is motivated by historical accounts of ethnic Ukrainian Party members opposing the harsh policies during the famine (Kotkin, 2017). We are able to identify the ethnicity of Party Secretaries who were sent to the 1930 Party Congress to vote for collectivization.[32] This allows us to distinguish the effects of having delegates who were Ukrainians themselves versus having delegates who were of other ethnicities.[33] In column (5), we add the triple interactions of each of the two new controls with Ukrainian population share and the famine dummy variable. Both triple interactions are positive and statistically significant, while the double interactions are statistically zero. Moreover, the triple interaction for non-Ukrainian delegates is larger in magnitude than for ethnically 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 in the regions populated by ethnic Ukrainians that led to Ukrainian mortality and ethnic Ukrainian leaders relenting more to the reality of famine in Ukrainian-populated areas.[34]

Panel B examines natality. The coefficients vary slightly in precision, but the signs mirror those for mortality. Note that in column (10), we find that the difference is the effect of ethnic Ukrainian and non-ethnic Ukrainian delegates in Ukrainian-populated areas on birth rate is statistically different at the 5% level, the p-value for the difference between the two triple interaction coefficients is 0.04. It is not presented in the table.

We can examine the dynamic triple interaction effects by repeating the estimates in Table

[32]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. See the Data Appendix.

[33]We do not distinguish between other ethnicities, dominated by Russians, because of limited variation.

[34]These results are also consistent with the view that Stalin often sent "outsiders" to govern to counteract local loyalties and nationalism. See the Background Section.

6, but replacing the famine dummy variable with year dummy variables. Figure 7 shows the timing of the effects of these policy proxies. We find that the effect manifests during the famine. This goes against concerns that the estimates are driven by spurious correlations. For brevity, we do not present the figures for natality.[35]

## 5.5 Soviet Agricultural Policies

Another way to connect the estimates of high Ukrainian famine mortality to Soviet policies is to explore the relationship between Ukrainian population share, the policy and political variables examined in the previous section on the implementation of Soviet economic policies during the famine. We examine collectivization, the main economic policy for rural areas, which was supposed to boost production and procurement, as well as increase mechanization.

We estimate the same specifications as before, except that, for brevity, we go directly to the specification that distinguishes ethnic Ukrainian and non-Ukrainian delegates. Table 7 Panel A examines the share of households in collective farms as the dependent variables. Panel B examines mechanization: the amount of tractor horsepower per capita. 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. For mechanization, the signs are opposite of those for mortality and collectivization. This implies that the same factors that caused Ukrainian-populated-regions to suffer higher mortality and more intense collectivization, also caused these regions to receive less mechanization.[36]

[35] The temporal patterns are consistent with those for mortality, but less precisely estimated. They are available upon request.

[36] 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 7 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.

In Table 8, we examine the other outcomes that were supposed to be affected by collectivization: procurement and production. Panel A examines reported production. The triple interaction effects are mostly statistically insignificant, which suggests that differential grain production does not play an obvious role in explaining excess Ukrainian mortality. In Panel B, the triple interaction effects for reported procurement as a share of reported production are more precisely estimated. The signs are similar to the signs for the triple interaction estimates for collectivization and mortality. This is consistent with the notion that excess procurement played an important role in causing famine mortality in Ukrainian regions. Note that as we discussed earlier, many scholars have cast doubt on official aggregate grain production data. To the best of our knowledge, there has been no dispute about the quality of the procurement data. Nevertheless, both estimates should be interpreted cautiously as suggestive.

## 5.6 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.

Figure 8a shows the mean and normalized standard deviation in mortality rates over time; figures 8b and 8c present mean and normalized standard deviation in district-level mortality for the republics of Russia and Ukraine.[37] They show that mean mortality and the variation across districts increase in 1933, for the full sample and for each republic. These results show that the spatial patterns which exist at the province level for the full sample also exist

[37]District-level analysis does not include the republic of Belarus because we were not yet able to collect 1928 mortality data for Belarus.

across districts within republics. Thus, there is inequality in famine intensity across smaller administrative units.

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 S.S.R. side of the border between Russia and Ukraine.[38] Figure 9a 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 9b, 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 S.S.R..

Second, the disaggregated data allow us to examine whether similar patterns exist across districts within provinces and across provinces. Soviet policies were centrally planned and implemented top-down. If collectivization or procurement targets were partly based on Ukrainian population share and implemented systematically, we expect similar associations across large administrative units as across smaller ones within the large units.

Table 9 column (1) first replicates the baseline specification from the province-level estimate.[39] In addition, we include province-year fixed effects to isolate the within province variation. This allows us to account for famine province-specific factors that many historians have argued as influential on the geography of mortality (e.g., local politics). The results are very robust. The exhibition of similar patterns at different levels of bureaucracy is consistent with the presence of a systematic and centrally planned policy. Columns (2) to (7) show that the results are robust to alternative definitions of ethnic Ukrainians and additional controls.

[38]The 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).

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

## 5.7 Back-of-the-Envelope Calculation

We conduct a simple back-of-the-envelope calculation to understand what famine mortality would have been had there been no ethnic bias – i.e., the interaction coefficient of Ukrainian population share and the famine dummy variable in equation (2) is zero. Conceptually, this is akin to asking what mortality would have been had Ukrainians died at the same rate as other ethnicities, which, in our sample, were mostly Russians. Using the estimates from equation (2), Table 3 column (4), we predict that the number of deaths in non-famine years is on average 2.71 million, and in 1933 is 4.97 million.[40] The difference, 2.26 million ($4.97 - 2.71 = 2.26$ million) is excess deaths due to the famine. If we assign the Ukrainian interaction coefficient to be zero, predicted deaths in 1933 would have been 3.22 million. Thus, famine deaths without ethnic bias would have been the difference between this number and deaths in non-famine years, 0.51 million ($3.22 - 2.71 = 0.51$ million). It follows that ethnic bias contributes 77% ($1 - .51/2.26 = .77$) to famine deaths in our sample. See Table 10 for the calculation.

One way to assess the plausibility of our estimates is to note that non-Ukrainian mortality rates in our sample are low. For example, if we take the estimates provided by historians in terms of total famine deaths of seven million, and subtract the deaths in Kazakhstan (1 to 1.5 million) and the Ukraine (2.6 to 3.9 million), we are left with 1.6 to 3.4 million. This results in famine mortality rate of 14 to 30 per 1,000 for the 114 million residents of Russia. A similar calculation for Ukraine, which had population of 32 million, yields famine mortality rate of 81 to 122 per 1,000. Since the back-of-the-envelope exercise conceptually asks what mortality would have been if ethnic Ukrainians the bulk of whom lived in the Ukrainian S.S.R. died at the same rate as other ethnicities in our sample (mostly Russians), it is then not surprising that ethnic bias contributes to such a large proportion of mortality.

We can repeat the exercise for the republics of Ukraine, Russia, and Belarus separately.

[40] This is very similar to 4.81 million 1933 deaths reported in our sample.

We find that in Ukraine during non-famine years, predicted deaths are 0.53 million. Predicted deaths in 1933 are 2.03 million. The difference, 1.50 million ($2.03 - 0.53$ million), is excess deaths due to the famine. If we assign the Ukrainian interaction coefficient to be zero, predicted deaths in 1933 would have been 0.64 million. Thus, famine deaths without ethnic bias would have been the difference between this number and mortality in non-famine years, 0.12 million ($0.64 - 0.53 = 0.12$, note a small discrepancy due to rounding). Thus, ethnic bias contributes 92% ($1 - .12/1.50 = .92$) of famine deaths in the Ukrainian republic. Since approximately 80% of the population of Ukraine were ethnically Ukrainian, the estimate that ethnic bias explains 92% of famine mortality implies that mortality rates were higher for ethnic Ukrainians than other ethnicities in the Ukraine.

Repeating this exercise for Russia and Belarus, we find that ethnic bias against Ukrainians explains 49% of famine deaths in Russia and 12% of famine deaths in Belarus. Since only 8% of the population of Russia and only 0.7% of the population of Belarus were ethnic Ukrainian, these estimates again imply disproportionately high mortality rates among ethnic Ukrainians in the republics of Russia and Belarus.

# 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 indicate that state policy, and in particular, ethnic bias against Ukrainians played an important role in the famine.

These sorrowful findings prompt several questions of future inquiry. First, more research is needed to reveal the motives behind ethnic bias in Soviet policy. This is particularly interesting, since communist ideology espouses one national identity and no ethnic component. Understanding the drivers of ethnic-biased policy in this context could shed light on the process of nation building and the tradeoffs for the central government, adding to works such as Alesina and Reich (2015), and Alesina et al. (2018). 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. Several recent studies suggest that the consequences for European political economy are long-lasting (Korovkin and Makarin, 2019; Rozenas and Zhukov, 2019, e.g.,).

# References


**Alesina, Alberto and Bryony Reich**, "Nation Building," 2015. NBER Working Paper 18839.

**_ , _ , and Alessandro Riboni**, "Nation-Building, Nationalism, and Wars," 2018. Working Paper.

**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 Istroriya Rossii: 1927-1959*, Moscow: Informatika, 1998.

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

**Cameron, Sarah**, *The Hungry Steppe: Famine, Violence, and the Making of Soviet Kazakhstan*, Ithaca, N.Y.: Cornell University Press, 2018.

**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 Sovetskoj Derevni. Kollektivizatsiya i raskulachivanie. Dokumenti i materialy v 5 tomakh, 1927-1939*, Rosspen, 1999-2006.

**Dasgupta, P. and D. Ray**, "Inequality as a Determinant of Malnutrition and Unemployment: Theory," *The Economic Journal*, 1986, *96*, 1011–1034.

**Davies, R. W., Oleg V. Khlevniuk, E. A. Rees, Liudmila P. Kosheleva, and Larisa A. Rogovaya, eds**, *The Stalin-Kaganovich Correspondence, 1931-36*, New Haven: Yale University Press, 2003.

**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.

**Durante, Ruben and Johannes Buggle**, "Climatic Risk, Cooperation, and the Co-Evolution of Culture and Institutions," *Economic Journal*, forthcoming.

**Dyson, Tim and Cormac Ó Gráda, eds**, *Famine Demography. Perspectives from the Past and Present*, Oxford, UK: Oxford University Press, 2002.

**Ellman, Michael**, "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 uses of hunger Stalin's solution of the peasant and national questions in Soviet Ukraine, 1932 to 1933," in Lubomyr Luciuk Declan Curran and Andrew G. Newby, eds., *Famines in European History. The Last Great European Famines Reconsidered*, Routledge, 2015, pp. 223–260.

**Gregory, Paul**, *Terror by Quota State Security from Lenin to Stalin (an Archival Study)*, Yale University Press, 2009.

**Heldring, Leander**, "The Origins of Violence in Rwanda," *The Review of Economic Studies*, 06 2020. rdaa028.

**Kondrashin, Viktor**, *Golod 1932-1933 godov: Tragediya Sovetskoj 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.

**Lositskij, A.**, "Pishchevoe znachenie sovremennikh produktov pitaniya [Nutritive value of modern food products]," *Vestnik statistiki*, 1920, *1920* (1-4), 102–121.

_ , *Sostoyanie pitaniya gorodskogo naseleniya SSSR, 1919-1924 [The Level of Nutrition of Urban Population of the USSR, 1919-1924]*, TsSU, 1926.

_ **, ed.**, *Sostoyanie pitaniya selskogo naseleniya SSSR v 1920-1924 [The Level of Nutrition of Rural Population in the USSR in 1920-1924]*, TsSU, 1928.

**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.

**Martin, Terry**, *The Affirmative Action Empire: Nations and Nationalism in the Soviet Union, 1923–1939*, Cornell University Press, 2001.

**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 prinuditelnikh migratsij v SSSR*, Moscow: O.G.I, 2001.

**Protasov, L.G., A.V. Repnikov V.V. Zhuravlev, and V.V. Shelokhaev**, *Vserossijskoe uchreditelnoe sobranie. Entsiklopedia. (All-Russian Constitutional Assembly. Encyclopedia)*, Moscow: Rosspen, 2014.

**Rozenas, Arturas and Yuri M. Zhukov**, "Mass Repression and Political Loyalty: Evidence from Stalin's 'Terror by Hunger'," *American Political Science Review*, 2019, *113* (2), 569–583.

**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.

**Tauger, Mark**, "Natural Disaster and Human Actions in the Soviet Famine of 1931-1933," *The Carl Beck Papers in Russian and East European Studies*, 2001.

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

Table 1: Per Capita Food Production and Requirements for the U.S.S.R.

| | Year | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Variable | 1927 | 1928 | 1929 | 1930 | 1931 | 1932 | 1933 | 1937 | 1939 |
| (1) Total population (mil.) | 147.0 | 150.5 | 154.2 | 157.5 | 160.5 | 163.3 | 165.8 | 162.0 | 165.5 |
| (2) Rural population (mil.) | 120.7 | 124.3 | 126.6 | 128.0 | 128.3 | 127.0 | 127.1 | 110.1 | 110.6 |
| (3) Production (mt) | 74.1 | 73.3 | 71.7 | 83.5 | 69.5 | 69.9 | 89.8 | 120.3 | 100.9 |
| (4) DW (mt) | . | . | . | 75.0 | 61.0 | 57.5 | 73.5 | . | . |
| (5) Procurement rate | 14.9% | 14.7% | 22.4% | 26.6% | 32.9% | 27.2% | 26.3% | 26.5% | 30.7% |
| (6) Production pc (kg) | 504 | 487 | 465 | 530 | 433 | 428 | 542 | 742 | 609 |
| **(7) Production pc (cal.)** | **4,329** | **4,183** | **3,994** | **4,555** | **3,716** | **3,675** | **4,651** | **6,373** | **5,231** |
| (8) DW (cal.) | . | . | . | 4,089 | 3,263 | 3,024 | 3,807 | . | . |
| (9) Rural retention pc (kg) | 523 | 503 | 440 | 479 | 364 | 400 | 521 | 803 | 632 |
| **(10) Rural retention pc (cal.)** | **4,486** | **4,318** | **3,775** | **4,113** | **3,122** | **3,436** | **4,469** | **6,890** | **5,425** |
| (11) *Cal. needs pc -- heavy labor* | *2,450* | *2,453* | *2,450* | *2,446* | *2,439* | *2,427* | *2,421* | *2,375* | *2,369* |
| (12) *Cal. needs pc -- avoid mortality* | *619* | *619* | *620* | *620* | *621* | *622* | *622* | *626* | *627* |

*Notes:* Data for population, production and procurement are official statistics. DW is the amount reported by Davis and Wheatcroft (1997). Conversion from grain to calories are based on estimates from Chayanov (1916). Population caloric requirements reported at the bottom of the table adjust for demographic composition (e.g., age, gender, rural/urban). Caloric needs for heavy labor use official Soviet estimates for adult males doing heavy labor (rural) -- 3,750 per day -- and doing mid-level labor (urban) -- 2,750 per day, for relative caloric needs of other groups are based on Lositskij A. (1926). Caloric needs for avoiding mortality are based on the 900 caloreis per day for prime age adult males (Dasgupta and Ray, 1986) and the relative calroic needs of other groups from Lositskij A. (1926).

Table 2: Per Capita Food Production and Requirements for Russia and Ukraine

| Variable | Year | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 1927 | 1928 | 1929 | 1930 | 1931 | 1932 | 1933 | 1937 | 1939 |
| | A. Russian S.S.R. | | | | | | | | |
| (1) Total pop (mil.) | 100.9 | 103.3 | 106.0 | 108.3 | 111.0 | 106.1 | 101.3 | 104.0 | 106.9 |
| (2) Rural pop (mil.) | 83.4 | 85.7 | 87.3 | 88.3 | 88.4 | 87.1 | 86.8 | 69.6 | 70.8 |
| (3) Prod (mt) | 50.8 | 54.4 | 48.0 | 55.3 | 46.2 | 48.1 | 59.6 | 87.4 | 65.9 |
| (4) Proc rate | . | 15.9% | 21.2% | 25.0% | 32.1% | 28.5% | 25.6% | . | . |
| (5) Prod pc (kg) | 503 | 526 | 453 | 511 | 416 | 453 | 588 | 841 | 617 |
| **(6) Prod pc (cal.)** | **4,321** | **4,519** | **3,891** | **4,386** | **3,576** | **3,891** | **5,051** | **7,218** | **5,295** |
| (7) Rural retention pc (kg) | . | 534 | 434 | 470 | 355 | 394 | 511 | . | . |
| **(8) Rural retention pc (cal.)** | **.** | **4,582** | **3,724** | **4,033** | **3,045** | **3,386** | **4,388** | **.** | **.** |
| (9) *Cal. needs pc -- heavy labor* | *2,448* | *2,449* | *2,446* | *2,442* | *2,432* | *2,445* | *2,463* | *2,365* | *2,361* |
| (10) *Cal. needs pc -- avoid mortality* | *618* | *618* | *618* | *618* | *619* | *618* | *616* | *626* | *626* |
| | B. Ukrainian S.S.R. | | | | | | | | |
| (11) Total pop (mil.) | 29.0 | 29.6 | 30.3 | 30.8 | 31.3 | 31.7 | 31.9 | 28.4 | 29.6 |
| (12) Rural pop (mil.) | 23.6 | 24.6 | 24.9 | 25.1 | 25.0 | 24.8 | 25.0 | 18.8 | 18.7 |
| (13) Prod (mt) | 18.6 | 13.9 | 18.7 | 22.7 | 18.3 | 14.5 | 22.0 | 22.5 | 23.8 |
| (14) Proc rate | 22.9% | 13.6% | 28.3% | 33.8% | 39.5% | 29.2% | 27.8% | . | . |
| (15) Prod pc (kg) | 641 | 469 | 618 | 739 | 587 | 457 | 689 | 792 | 805 |
| **(16) Prod pc (cal.)** | **5,506** | **4,023** | **5,307** | **6,342** | **5,039** | **3,927** | **5,919** | **6,797** | **6,912** |
| (16) Rural retention pc (kg) | 607 | 488 | 538 | 600 | 444 | 415 | 637 | . | . |
| **(17) Rural retention pc (cal.)** | **5,213** | **4,189** | **4,620** | **5,149** | **3,815** | **3,559** | **5,467** | **.** | **.** |
| (18) *Cal. needs pc -- heavy labor* | *2,455* | *2,462* | *2,459* | *2,455* | *2,446* | *2,437* | *2,437* | *2,374* | *2,357* |
| (19) *Cal. needs pc -- avoid mortality* | *622* | *621* | *621* | *622* | *623* | *623* | *623* | *629* | *631* |

*Notes:* Data for population, production and procurement are official statistics. Conversion from grain to calories are based on estimates from Chayanov (1916). Population caloric requirements reported at the bottom of the table adjust for demographic composition (e.g., age, gender, rural/urban). Caloric needs for heavy labor use official Soviet estimates for adult males doing heavy labor (rural) -- 3,750 per day-- and doing mid-level labor (urban) -- 2,750 per day, for relative caloric needs of other groups are based on Lositskij A. (1926). Caloric needs for avoiding mortality are based on the 900 caloreis per day for prime age adult males (Dasgupta and Ray, 1986) and the relative calroic needs of other groups from Lositskij A. (1926).

Table 3: Famine Mortality in Ethnic Ukrainian Areas – Baseline Estimate

| | Dependent variable: Mortality in year t+1 | | | | | |
|---|---|---|---|---|---|---|
| | Reported Grain (1) | Predicted Grain (2) | Reported Grain (3) | Predicted Grain (Baseline) (4) | Famine = 1931, 32 (5) | Omit Province FE (6) |
| Ukrainians | | | | | | -0.007*** |
| | | | | | | (0.001) |
| Ukrainians x Famine | | | 0.047*** | 0.051*** | 0.032** | 0.051*** |
| | | | (0.004) | (0.008) | (0.013) | (0.007) |
| *Standardized Coef.* | | | *0.766* | *0.825* | *0.521* | *0.831* |
| Grain | 0.004 | 0.001 | 0.007** | 0.003 | 0.001 | -0.001 |
| | (0.004) | (0.002) | (0.004) | (0.002) | (0.002) | (0.002) |
| Grain x Famine | 0.120** | 0.079 | 0.072 | 0.003 | -0.004 | -0.002 |
| | (0.054) | (0.056) | (0.048) | (0.028) | (0.013) | (0.027) |
| *Standardized Coef.* | *0.484* | *0.417* | *0.290* | *0.017* | *-0.023* | *-0.009* |
| Urbanization | -0.011** | -0.012** | -0.009** | -0.010** | -0.011*** | -0.013*** |
| | (0.004) | (0.004) | (0.004) | (0.004) | (0.004) | (0.002) |
| Urbanization x Famine | 0.014 | 0.010 | 0.011 | 0.005 | 0.010 | 0.007 |
| | (0.019) | (0.014) | (0.015) | (0.010) | (0.006) | (0.015) |
| Observations | 337 | 337 | 337 | 337 | 337 | 337 |
| R-squared | 0.670 | 0.656 | 0.802 | 0.783 | 0.735 | 0.431 |

*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. Famine is an indicator that equals one in 1932 (columns 1--4 and 6) or in 1931 and 1932 (column 5), and zero otherwise. Estimates in columns 1 and 3 control for Reported grain and Reported grain x Famine. Estimates in columns 2 and 4--6 control for Predicted grain and Predicted grain x Famine. Reported and predicted grain are measured in 10s of kilograms per person per day. Estimates in columns 1--5 control for profince fixed effects. All estimates control for year fixed effects. Standard errors robust to heteroskedasticity are in parentheses. Standardized coefficients express the effect of increasing Grain 1932 or Ukrainians by one standard deviation on 1933 mortality in terms of standard deviations. The Data Appendix presents the source of every variable.

# Table 4: Famine Mortality and Natality in Ethnic Ukrainian Areas – Robustness

| | Predicted Grain (Baseline) (1) | Total Ukrainians (2) | Urban Ukrainians (3) | Mother Tongue Ukrainian 1926 (4) | Mother Tongue Ukrainian 1897 (5) | Control for Demographic Structure x Famine (6) | Control for Latitude x Longitude x Famine (7) | Weighted by Population (8) | Omit outliers by Cook's D (9) | Exclude Ukraine S.S.R. (10) | 1892 Famine (11) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | A. Dependent variable: Mortality in year t+1 | | | | | | | | | | |
| Ukrainians x Famine | 0.051*** | 0.055*** | 0.090*** | 0.056*** | 0.058*** | 0.048*** | 0.060*** | 0.045*** | 0.048*** | 0.086*** | 0.000 |
| Robust SE | (0.008) | (0.008) | (0.013) | (0.009) | (0.009) | (0.008) | (0.008) | (0.004) | (0.009) | (0.008) | (0.003) |
| Conley SE | [0.008] | [0.009] | [0.014] | [0.010] | [0.009] | [0.009] | [0.006] | [0.006] | [0.009] | [0.010] | [0.003] |
| *Standardized Coef.* | *0.825* | *0.820* | *0.785* | *0.747* | *0.820* | *0.786* | *0.971* | *0.740* | *0.177* | *0.839* | *-0.004* |
| Observations | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 319 | 319 | 1,297 |
| R-squared | 0.783 | 0.782 | 0.778 | 0.771 | 0.788 | 0.790 | 0.813 | 0.879 | 0.834 | 0.765 | 0.864 |
| | B. Dependent variable: Natality in year t+1 | | | | | | | | | | |
| Ukrainians x Famine | -0.014*** | -0.015*** | -0.025*** | -0.015*** | -0.016*** | -0.015*** | -0.013*** | -0.012*** | -0.035*** | -0.030*** | 0.004*** |
| Robust SE | (0.003) | (0.004) | (0.006) | (0.004) | (0.004) | (0.005) | (0.004) | (0.002) | (0.012) | (0.005) | (0.001) |
| Conley SE | [0.004] | [0.004] | [0.006] | [0.004] | [0.004] | [0.006] | [0.003] | [0.002] | [0.008] | [0.005] | [0.002] |
| *Standardized Coef.* | *-0.431* | *-0.428* | *-0.409* | *-0.379* | *-0.424* | *-0.456* | *-0.405* | *-0.351* | *-0.279* | *-0.518* | *0.102* |
| Observations | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 314 | 319 | 1,297 |
| R-squared | 0.833 | 0.833 | 0.832 | 0.832 | 0.833 | 0.833 | 0.835 | 0.879 | 0.888 | 0.828 | 0.934 |
| Ukrainians | | | | | | | | | | | |
| Mean | 0.104 | 0.095 | 0.055 | 0.074 | 0.085 | 0.104 | 0.104 | 0.104 | 0.021 | 0.061 | 0.172 |
| Std. Dev. | 0.216 | 0.197 | 0.116 | 0.178 | 0.188 | 0.216 | 0.216 | 0.216 | 0.049 | 0.113 | 0.299 |

*Notes:* Observations are at the province and year level. Mortality is the number of deaths divided by population. Natality is the number of live births divided by population. Unless otherwise specified in the column headings, Ukrainians is the share of ethnic Ukrainians in the rural population according to the 1926 Census; in column 2, Ukrainians is the share of ethnic Ukrainians in the total population according to the 1926 Census; in column 3, Ukrainians is the share of ethnic Ukrainians in the urban population according to the 1926 Census; in column 4, Ukrainians is the share of people whose mother tongue is Ukrainian according to the 1926 Census; in columns 5 and 10, Ukrainians is the share of people whose mother tongue is Ukrainian according to the 1897 Census. In columns 1--10, Famine is an indicator that equals one in 1932 and zero otherwise; in column 11, Famine is an indicator that equals one in 1891 and zero otherwise. Estimates in column 6 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). Estimates in column 7 control for Latitude x Famine, Longitude x Famine, and Latitude x Longitude x Famine. Estimates in column 8 are weighted by population. Estimates in column 9 omit observations wich have Cook's D higher than 4/337. Estimates in column 10 omit Ukraine. All estimates control for Urbanization, Urbanization x Famine, Predicted grain, Predicted grain x Famine, and Province- and Year fixed effects. Predicted grain is measured in 10s of kilograms per person per day. Standard errors robust to heteroskedasticity are in parentheses. Conley standard errors are in square brackets; spatial autocorrelation is assumed to exist among observations that are within 900 kilometers of each other. The Data Appendix presents the source of every variable.

Table 5: Famine Mortality and Natality in Ethnic Ukrainian Areas – Robustness to Political Controls

| | Dependent variable: | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A. Mortality in year t+1 | | | | | B. Natality in year t+1 | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Ukrainians x Famine | 0.051*** | 0.055*** | 0.050*** | 0.052*** | 0.050*** | -0.014*** | -0.016*** | -0.014*** | -0.014*** | -0.014*** |
| | (0.008) | (0.008) | (0.008) | (0.007) | (0.008) | (0.003) | (0.003) | (0.003) | (0.003) | (0.003) |
| Grain 1928 x Famine | | 0.092* | | | | | -0.035** | | | |
| | | (0.048) | | | | | (0.015) | | | |
| Bolshevik votes 1917 x Famine | | | -0.013* | | | | | -0.002 | | |
| | | | (0.007) | | | | | (0.004) | | |
| Urban communists x Famine | | | | -0.001 | | | | | 0.001* | |
| | | | | (0.001) | | | | | (0.000) | |
| Rural communists x Famine | | | | 0.003 | | | | | -0.001 | |
| | | | | (0.002) | | | | | (0.001) | |
| Voting delegates 1930 x Famine | | | | | 0.033 | | | | | -0.012 |
| | | | | | (0.021) | | | | | (0.009) |
| Observations | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 337 |
| R-squared | 0.783 | 0.805 | 0.787 | 0.800 | 0.791 | 0.833 | 0.834 | 0.833 | 0.835 | 0.833 |

*Notes:* Observations are at the province and year level. Mortality is the number of deaths divided by population. Natality is the number of live births 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 10s of 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, Predicted grain, Predicted grain x Famine, Province- and Year fixed effects. Standard errors robust to heteroskedasticity are in parentheses. The Data Appendix presents the source of every variable.

# Table 6: Heterogeneous Effects of Political Factors on Famine Mortality in Ethnic Ukrainian Areas

| | Dependent Variable: | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | A. Mortality in year t+1 | | | | | B. Natality in year t+1 | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Ukrainians x Famine | -0.172*** | 0.002 | 0.064 | -0.040* | -0.035 | 0.046** | 0.000 | -0.090** | 0.011 | -0.035 |
| | (0.040) | (0.042) | (0.071) | (0.023) | (0.033) | (0.022) | (0.014) | (0.038) | (0.015) | (0.027) |
| Grain 1928 x Famine | 0.003 | | | | | -0.011 | | | | |
| | (0.022) | | | | | (0.021) | | | | |
| Ukrainians x Grain 1928 x Famine | 2.973*** | | | | | -0.800*** | | | | |
| | (0.619) | | | | | (0.261) | | | | |
| Bolshevik votes 1917 x Famine | | -0.023** | | | | | 0.003 | | | |
| | | (0.012) | | | | | (0.005) | | | |
| Ukrainians x Bolshevik votes 1917 x Famine | | 1.039* | | | | | -0.428*** | | | |
| | | (0.551) | | | | | (0.129) | | | |
| Urban communists x Famine | | | 0.000 | | | | | 0.001*** | | |
| | | | (0.001) | | | | | (0.000) | | |
| Ukrainians x Urban communists x Famine | | | -0.131*** | | | | | -0.001 | | |
| | | | (0.046) | | | | | (0.019) | | |
| Rural communists x Famine | | | 0.000 | | | | | 0.000 | | |
| | | | (0.001) | | | | | (0.001) | | |
| Ukrainians x Rural communists x Famine | | | 0.107*** | | | | | -0.022** | | |
| | | | (0.025) | | | | | (0.009) | | |
| Voting delegates 1930 x Famine | | | | -0.005 | | | | | -0.002 | |
| | | | | (0.011) | | | | | (0.011) | |
| Ukrainians x Voting delegates 1930 x Famine | | | | 2.396*** | | | | | -0.637*** | |
| | | | | (0.396) | | | | | (0.147) | |
| Voting Ukrainian delegates 1930 x Famine | | | | | 0.331 | | | | | -0.105 |
| | | | | | (0.202) | | | | | (0.102) |
| Ukrainians x Voting Ukrainian delegates 1930 x Famine | | | | | 1.798*** | | | | | -0.085 |
| | | | | | (0.516) | | | | | (0.327) |
| Voting non-Ukrainian delegates 1930 x Famine | | | | | -0.024 | | | | | 0.003 |
| | | | | | (0.020) | | | | | (0.012) |
| Ukrainians x Voting non-Ukrainian delegates 1930 x Famine | | | | | 2.227*** | | | | | -0.440** |
| | | | | | (0.328) | | | | | (0.183) |
| Observations | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 337 | 337 |
| R-squared | 0.850 | 0.819 | 0.846 | 0.849 | 0.854 | 0.836 | 0.836 | 0.837 | 0.835 | 0.836 |

*Notes:* Observations are at the province and year level. Mortality is the number of deaths divided by population. Natality is the number of live births 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 10s of 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 to heteroskedasticity are in brackets. The Data Appendix presents the source of every variable.

Table 7: Heterogeneous Effects of Political Factors on Collectivization and Mechanization for Ethnic Ukrainian Areas

| | Dependent variable: | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | A. Collectivization | | | | B. Tractors' horse power | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Ukrainians x Famine | -0.792 | 0.070 | 1.156 | 0.572 | 0.132** | -0.024 | -0.077 | 0.045 |
| | (0.598) | (0.416) | (1.019) | (0.712) | (0.061) | (0.055) | (0.087) | (0.038) |
| Grain 1928 x Famine | 0.619 | | | | -0.030 | | | |
| | (0.505) | | | | (0.032) | | | |
| Ukrainians x Grain 1928 x Famine | 13.293* | | | | -2.654** | | | |
| | (6.787) | | | | (1.096) | | | |
| Bolshevik votes 1917 x Famine | | -0.487*** | | | | 0.020* | | |
| | | (0.084) | | | | (0.011) | | |
| Ukrainians x Bolshevik votes 1917 x Famine | | 7.847*** | | | | -0.969 | | |
| | | (2.947) | | | | (0.609) | | |
| Urban communists x Famine | | | -0.001 | | | | 0.001 | |
| | | | (0.005) | | | | (0.001) | |
| Ukrainians x Urban communists x Famine | | | -0.196 | | | | 0.114 | |
| | | | (0.499) | | | | (0.073) | |
| Rural communists x Famine | | | 0.052*** | | | | -0.001 | |
| | | | (0.017) | | | | (0.001) | |
| Ukrainians x Rural communists x Famine | | | 0.306 | | | | -0.088** | |
| | | | (0.204) | | | | (0.043) | |
| Voting Ukrainian delegates 1930 x Famine | | | | 0.275 | | | | -0.199 |
| | | | | (2.732) | | | | (0.225) |
| Ukrainians x Voting Ukrainian delegates 1930 x Famine | | | | 7.529 | | | | -2.125** |
| | | | | (8.183) | | | | (0.935) |
| Voting non-Ukrainian delegates 1930 x Famine | | | | -0.413 | | | | 0.008 |
| | | | | (0.281) | | | | (0.020) |
| Ukrainians x Voting non-Ukrainian delegates 1930 x Famine | | | | 11.864** | | | | -2.216*** |
| | | | | (4.849) | | | | (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 10s of 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 to heteroskedasticity are in parentheses. The Data Appendix presents the source of every variable.

Table 8: Heterogeneous Effects of Political Factors on Grain Production and Procurement in Ethnic Ukrainian Areas

| | Dependent Variable: | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | A. Reported Production | | | | B. Reported Procurement/Reported Production | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Ukrainians x Famine | -0.084 | -0.037 | -0.049 | 0.161 | -0.241 | 0.410 | 1.790** | 0.373 |
| | (0.169) | (0.126) | (0.274) | (0.191) | (0.321) | (0.292) | (0.754) | (0.500) |
| Grain 1928 x Famine | -0.224** | | | | 0.625*** | | | |
| | (0.101) | | | | (0.228) | | | |
| Ukrainians x Grain 1928 x Famine | 1.033 | | | | 10.555** | | | |
| | (2.334) | | | | (4.694) | | | |
| Bolshevik votes 1917 x Famine | | 0.006 | | | | -0.164* | | |
| | | (0.028) | | | | (0.090) | | |
| Ukrainians x Bolshevik votes 1917 x Famine | | 0.356 | | | | 5.526** | | |
| | | (0.820) | | | | (2.320) | | |
| Urban communists x Famine | | | 0.006** | | | | -0.008 | |
| | | | (0.003) | | | | (0.008) | |
| Ukrainians x Urban communists x Famine | | | -0.049 | | | | -0.016 | |
| | | | (0.154) | | | | (0.307) | |
| Rural communists x Famine | | | -0.005 | | | | 0.016 | |
| | | | (0.005) | | | | (0.014) | |
| Ukrainians x Rural communists x Famine | | | 0.052 | | | | 0.330** | |
| | | | (0.075) | | | | (0.142) | |
| Voting Ukrainian delegates 1930 x Famine | | | | 1.145* | | | | -3.189** |
| | | | | (0.663) | | | | (1.388) |
| Ukrainians x Voting Ukrainian delegates 1930 x Famine | | | | -3.051 | | | | 15.296*** |
| | | | | (2.454) | | | | (5.356) |
| Voting non-Ukrainian delegates 1930 x Famine | | | | -0.178** | | | | 0.253* |
| | | | | (0.077) | | | | (0.148) |
| Ukrainians x Voting non-Ukrainian delegates 1930 x Famine | | | | -0.673 | | | | 12.329*** |
| | | | | (1.658) | | | | (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 10s of 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 to heteroskedasticity are in parentheses. The Data Appendix presents the exact source of every variable.

Table 9: Famine Mortality in Ethnic Ukrainian Areas – Using district-level data

| | Dependent variable: Mortality | | | | | | |
|---|---|---|---|---|---|---|---|
| | Baseline with Province-Year FE | Total Ukrainians | Urban Ukrainians | Mother Tongue Ukrainian | Controlling for gender ratio | Controling for Latitude x Longitude x Famine | Omit Ukraine SSR |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Ukrainians x Famine | 0.039*** | 0.044*** | 0.043*** | 0.047*** | 0.040*** | 0.040*** | 0.022*** |
| | (0.006) | (0.007) | (0.010) | (0.007) | (0.006) | (0.006) | (0.007) |
| *Normalized* | *0.506* | *0.545* | *0.423* | *0.592* | *0.506* | *0.509* | *0.190* |
| Urbanization | 0.004 | 0.000 | -0.006 | 0.005 | 0.004 | 0.004 | 0.001 |
| | (0.009) | (0.008) | (0.009) | (0.008) | (0.008) | (0.009) | (0.007) |
| Urbanization x Famine | -0.008* | -0.004 | -0.006 | -0.008* | -0.008* | -0.007* | 0.002 |
| | (0.004) | (0.004) | (0.005) | (0.004) | (0.004) | (0.004) | (0.004) |
| Grain suitability x Famine | 0.022*** | 0.022*** | 0.029*** | 0.021*** | 0.023*** | 0.026*** | 0.013** |
| | (0.005) | (0.005) | (0.006) | (0.005) | (0.005) | (0.005) | (0.005) |
| Observations | 3,513 | 3,515 | 2,052 | 3,505 | 3,513 | 3,513 | 2,734 |
| R-squared | 0.770 | 0.773 | 0.786 | 0.777 | 0.770 | 0.773 | 0.742 |
| Provinces | 18 | 18 | 18 | 18 | 18 | 18 | 17 |
| Mortality 1933 | | | | | | | |
| Mean | 0.038 | 0.038 | 0.041 | 0.038 | 0.038 | 0.038 | 0.030 |
| Std. Dev. | 0.028 | 0.028 | 0.029 | 0.028 | 0.028 | 0.028 | 0.019 |
| Ukrainians | | | | | | | |
| Mean | 0.237 | 0.227 | 0.215 | 0.219 | 0.237 | 0.237 | 0.069 |
| Std. Dev. | 0.363 | 0.348 | 0.290 | 0.358 | 0.363 | 0.363 | 0.165 |

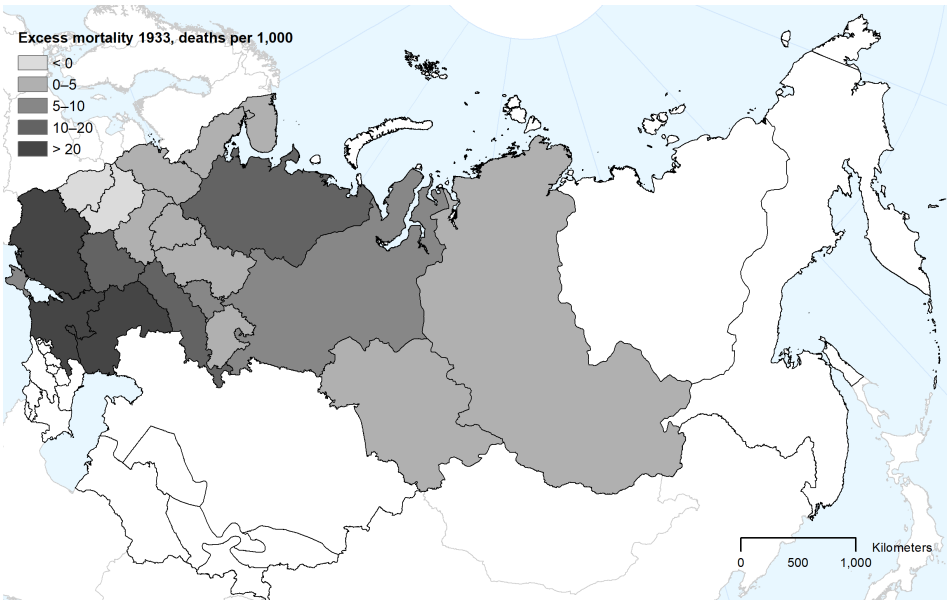
*Notes:* Observations are at the district and year level, the sample has two years: 1928 and 1933. Ukrainians is the share of ethnic Ukrainians in the rural population (columns 1, 5--7), total population (column 2), urban population (column 3), or the share of people whose mother tongue is Ukrainian (column 4) according to the 1926 Census. Famine is an indicator that equals one in 1933 and zero in 1928. Estimates in column 5 control for the Male/female ratio x Famine (according to the 1926 Census). Estimates in column 6 control for the district centroid Latitude x Famine, Longitude x Famine, and Latitude x Longitude x Famine. Standard errors clustered at the district level are in parentheses. The Data Appendix presents the source of every variable.

Table 10: Back-of-the-Envelope Calculation

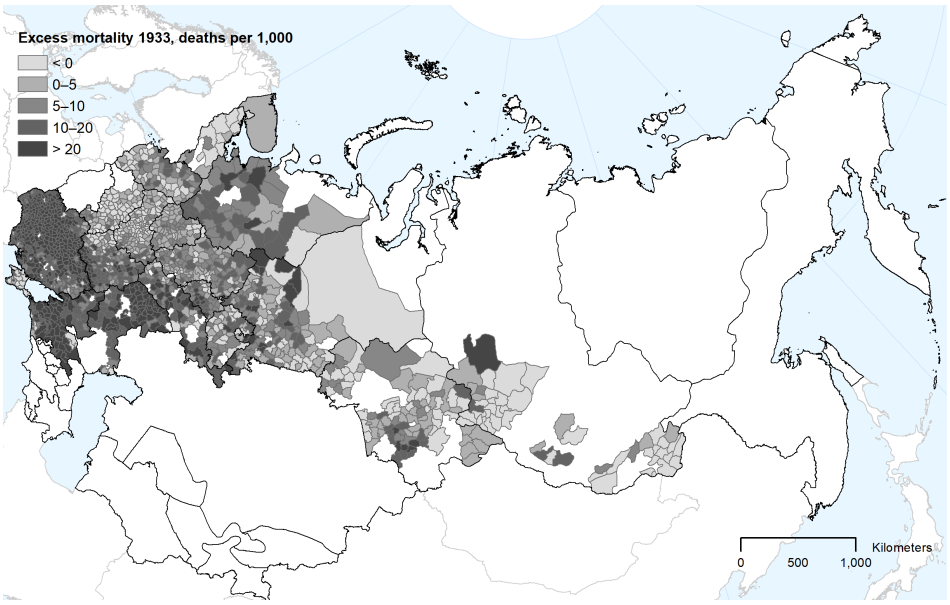
| | Full Sample | Ukraine | Russia | Belarus |
|---|---|---|---|---|
| (1) 1933 deaths if no famine (famine dummy = 0 ), mln | 2.71 | 0.53 | 2.10 | 0.08 |
| (2) Reported 1933 deaths, mln | 4.81 | 1.86 | 2.88 | 0.07 |
| (3) Predicted 1933 deaths (famine dummy = 1, Ukrainian = as reported), mln | 4.97 | 2.03 | 2.84 | 0.10 |
| (4) if no bias (famine dummy = 1, Ukrainian = 0), mln | 3.22 | 0.64 | 2.48 | 0.10 |
| (5) Total famine deaths: (3)-(1), mln | 2.26 | 1.50 | 0.74 | 0.02 |
| (6) if no bias: (4)-(1), mln | 0.51 | 0.12 | 0.38 | 0.02 |
| (7) Famine deaths due to bias: 1 - (6) / (5) | 0.77 | 0.92 | 0.49 | 0.12 |

Figure 1: Excess Mortality 1933

(a) Province Map



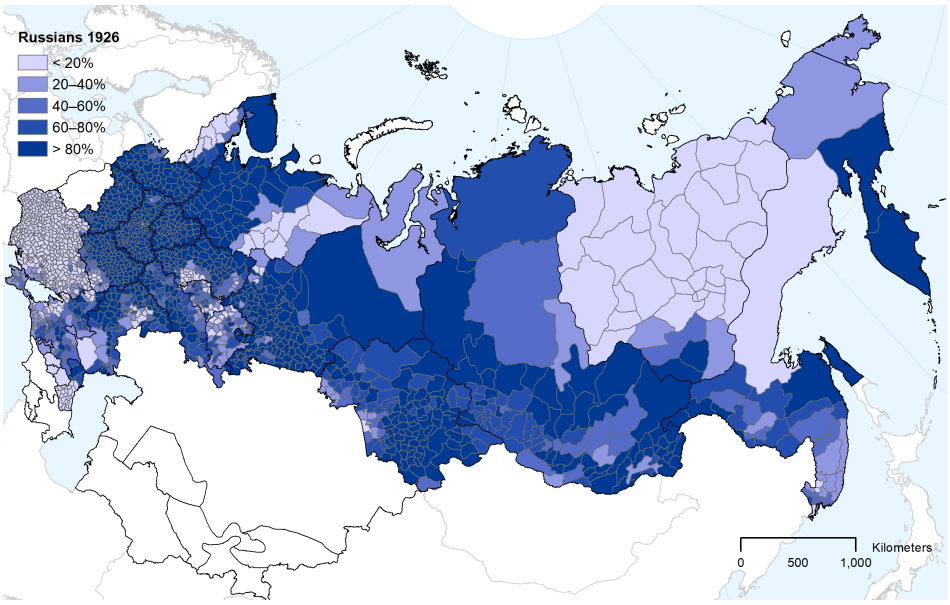

(b) District Map



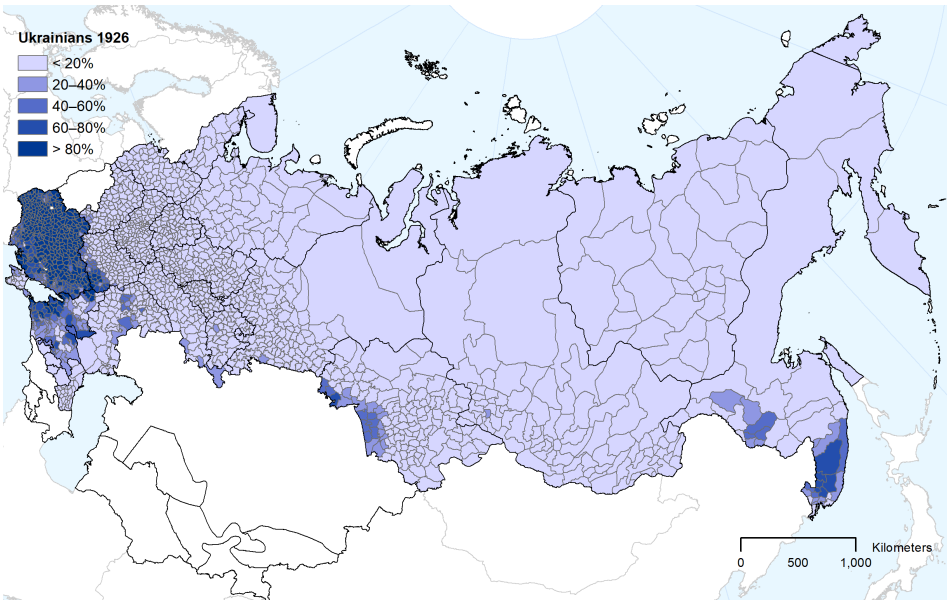

*Notes:* Excess mortality 1933 is mortality in 1933 minus mortality in 1928. *Source:* See the Data Appendix.

Figure 2: Rural Ethnic Composition 1926

(a) Russians

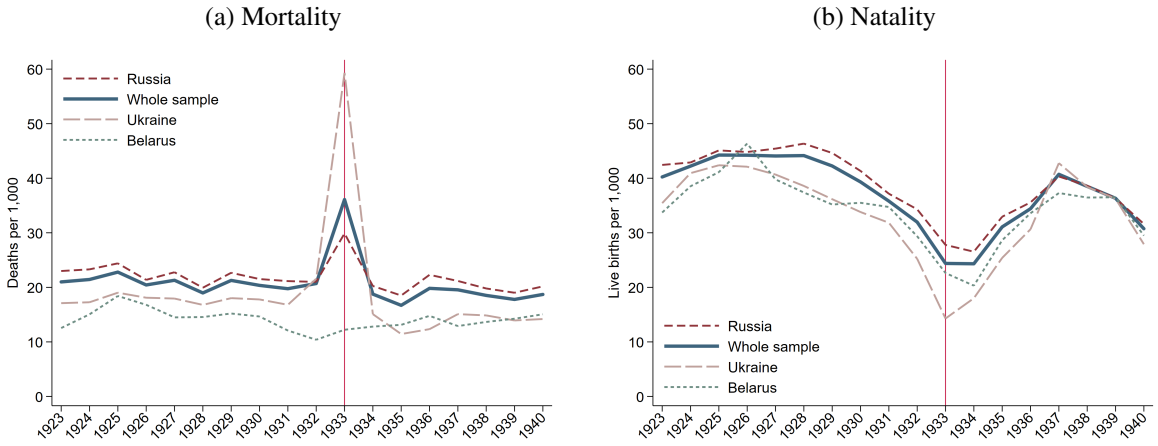



(b) Ukrainians



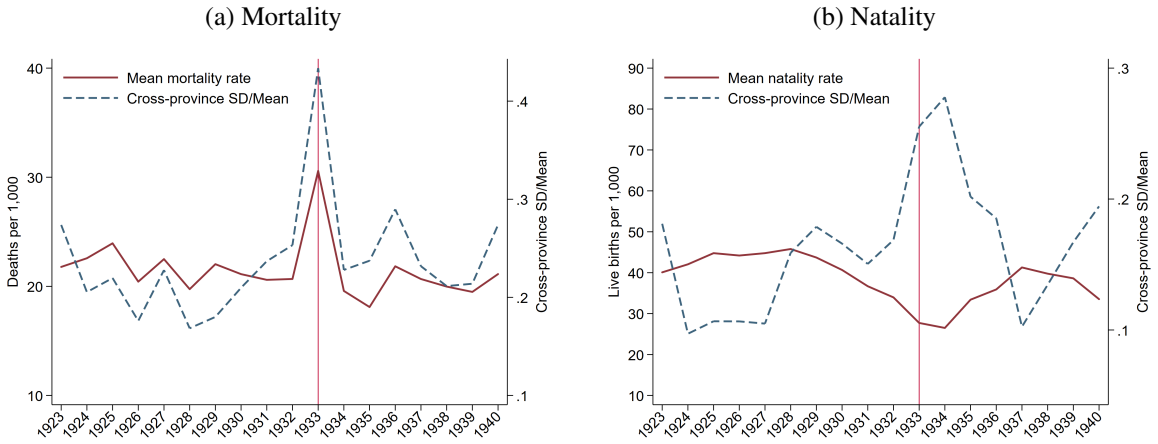

*Notes:* Share of ethnic Russians and Ukrainians in the rural population according to the 1926 Population Census.
*Source:* See the Data Appendix.

Figure 3: Mortality and Natality Rates over Time




*Notes:* Mortality is the number of deaths per 1,000 individuals. Natality is the number of live births per 1,000 individuals. *Source:* See the Data Appendix.

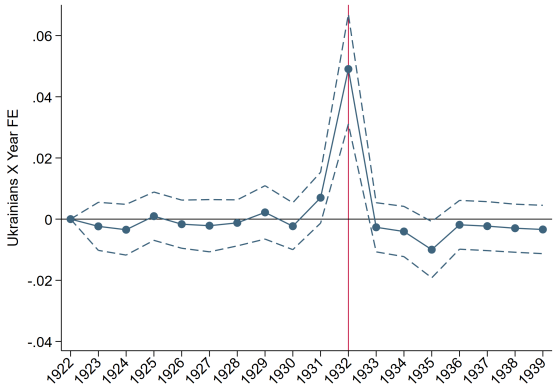
Figure 4: Cross-Province Mean and Standard Deviation of Mortality and Natality Rates




*Notes:* Mean mortality (natality) rate is the average mortality (natality) rate across provinces in each year. Cross-province SD/Mean is the standard deviation in mortality (natality) rates across provinces in year $t$ divided by the mean mortality (natality) rate in year $t$. *Source:* See the Data Appendix.

Figure 5: The Dynamic Effect of Ukrainian Population Share on Famine Mortality

(a) Ukrainians × Year FE



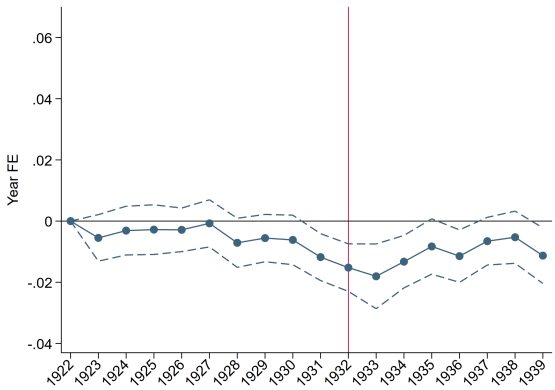

(b) Year FE


*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 interacted with year indicators, predicted grain interacted with year indicators, 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 the Data Appendix.

Figure 6: The Dynamic Effect of Ukrainian Population Share on Famine Natality

(a) Ukrainians × Year FE

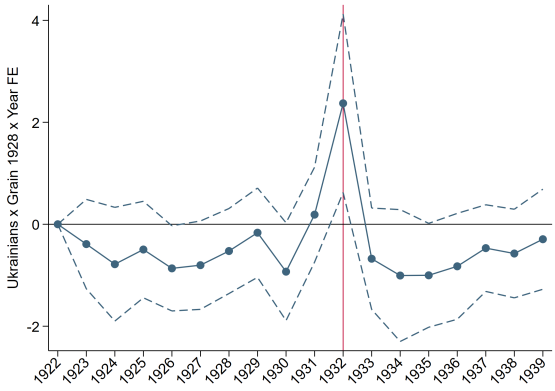

(b) Year FE



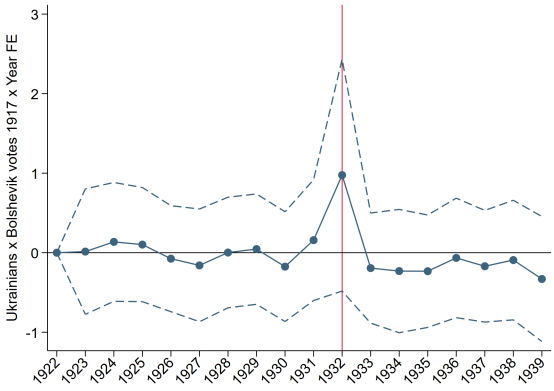

*Notes:* The figures show regression coefficients with their 95% confidence intervals from regressing natality in year $t + 1$ on the rural share of ethnic Ukrainians interacted with year indicators, urbanization interacted with year indicators, predicted grain interacted with year indicators, year indicators (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 the Data Appendix.

Figure 7: Heterogenous Effects of Ukrainian Population Share and Political Factors on Famine Mortality

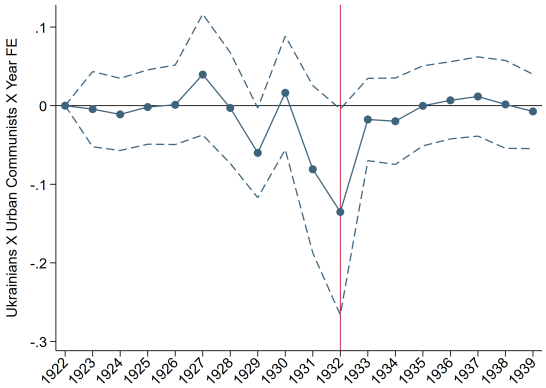
(a) Ukrainians × Grain 1928 × Year FE



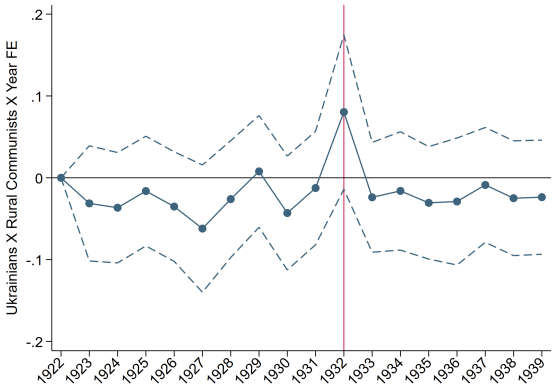

(b) Ukrainians × Bolshevik votes 1917 × Year FE



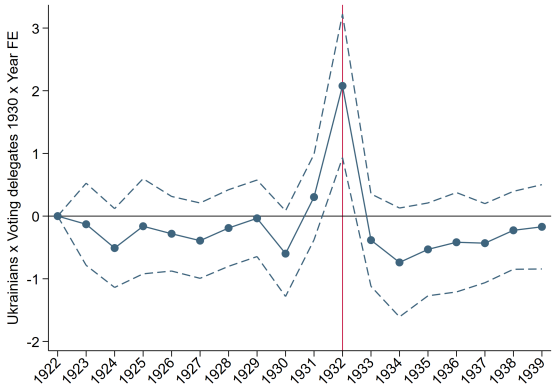

(c) Ukrainians × Urban Communists × Year FE

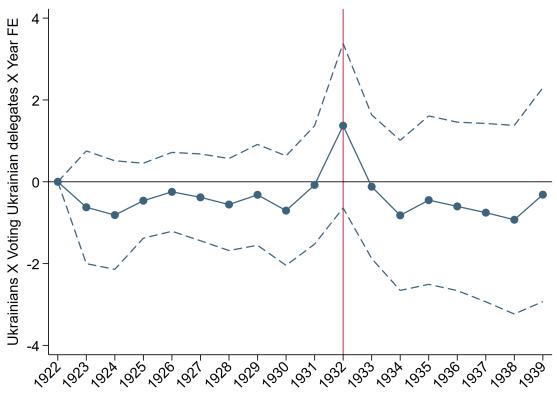



(d) Ukrainians × Rural Communists × Year FE

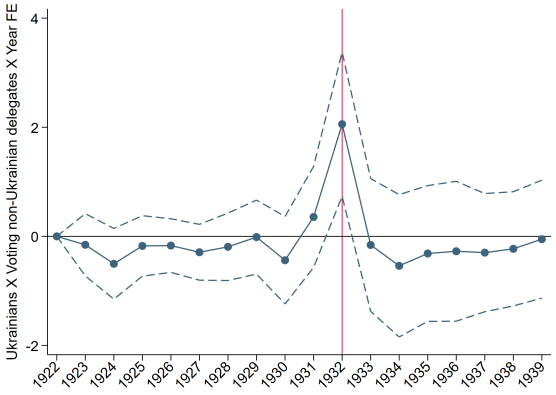



(e) Ukrainians × Voting delegates 1930 × Year FE






(f) Ukrainians × Voting Ukrainian delegates 1930 × Year FE

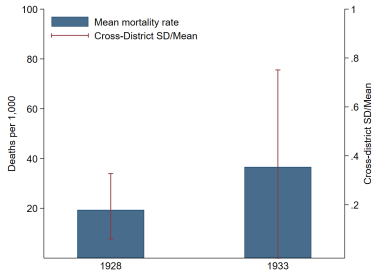



(g) Ukrainians × Voting non-Ukrainian delegates 1930 × Year FE

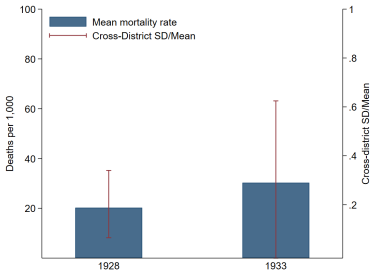
*Notes:* Figures (a), (b) and (e) plot coefficients and their 95% confidence intervals estimated from separate regressions. Figures (c) and (d) are estimated from one regression. Figures (f) and (g) are estimated from one regression. The dependent variable in each regression is mortality in year $t +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

## Figure 8: Cross-District Mean and Standard Deviation of Mortality Rates

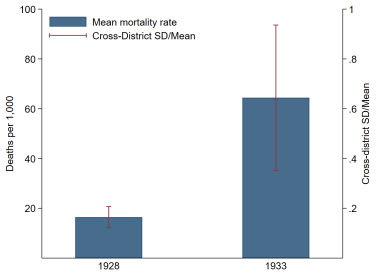
(a) Russia and Ukraine




(b) Russia




(c) Ukraine

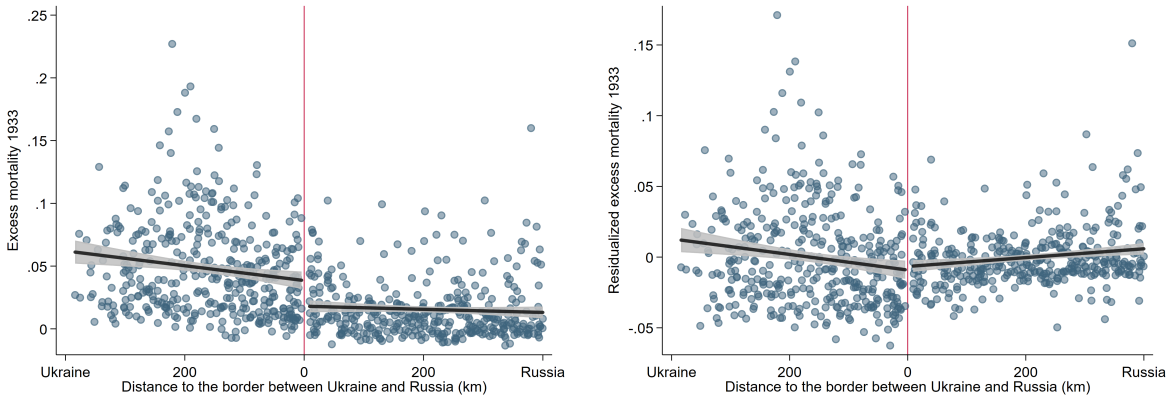



*Notes:* Mean mortality rate is the average mortality rate across districts in each year. Cross-province SD/Mean is the standard deviation in mortality rates across districts in year $t$ divided by the mean mortality rate in year $t$. *Source:* See the Data Appendix.

Figure 9: Excess Mortality in 1933 and Distance from the Ukrainian-Russian Border

(a) Excess mortality 1933

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




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

# Appendix

# A Data Appendix (Not for Publication)

## A.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 S.S.R., and the North Caucasus ethnic territories: Chechen Autonomous Province, Cherkess Autonomous Province, Dagestan Autonomous S.S.R., 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**

- 1920: 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.
- 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) *"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 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) *"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 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, 17 1926 Population Census.
- 1928–1932: is interpolated between 1927 and 1933.
- 1933: Russian state archive of economy (hereafter, 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.

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.

**Births and Deaths**

- 1923: 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 5.
- 1924: 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 8.
- 1925: Tsentralnoye Statisticheskoye Upravleniye SS.S.R. [Central Statistical Office of the U.S.S.R.] (1928) *"Yestestvennoye dvizheniye naseleniya Soyuza S.S.R. 1923–1925 [Natural movement of the population of the U.S.S.R.]"*, Volume I, Issue 1, Table 1.
- 1926: Yestestvennoye dvizheniye naseleniya Soyuza S.S.R. v 1926 g, Izdaniye TsSU SS.S.R. (1929), Table 1
- 1927–1932: Belarus, Ukraine – RGAE 1562/329/256; Russia – Demoscope.ru.

- 1933–1940: Demoscope.ru.

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.

### Natality and Mortality

Natality is the number of live births divided by population (what demographers call crude birth rate). 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, Russian (*Velikorusskiy*), and Ukrainian (*Malorusskiy*). The 1926 Census reports population by self-proclaimed ethnicity and by mother tongue, we use both. Data are calculated in our province borders using hand-created district-level 1897 and 1926 maps. The 1897 map is from ristat.org .

### 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 S.S.R. 1918–1923. Za pyat let raboty Tsentralnogo Statisticheskogo Upravleniya [A collection of statistical information on the U.S.S.R. 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) *"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], Part III, Tables 3 and 4.

- 1924: 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 III, Tables 6 and 7.

- 1925–1927: Statisticheskoye izdatelstvo TsSU SS.S.R. [Statistical Publishing House of the Central Statistical Office of the U.S.S.R.] (1929) *"Selskoye khozyaystvo SS.S.R. 1925–1928. Sbornik statisticheskikh svedeniy k XVI Vsesoyuznoy partkonferentsii [Agriculture of the U.S.S.R. 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 SS.S.R.. Statisticheskiy spravochnik 1932 [The national economy of the U.S.S.R.. 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 SS.S.R.. Yezhegodnik 1935 [Agriculture of the U.S.S.R.. 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.

**Procurement**

- 1924: Tsentralnoye Konventsionnoye Byuro Khlebozagotoviteley [Central Conventional Bureau of Grain Procurers] (1928) *"Yezhegodnik khlebnoy torgovli N1 [Yearbook of grain trade N 1]"*, Table 6.

- 1925: Tsentralnoye Konventsionnoye Byuro Khlebozagotoviteley [Central Conventional Bureau of Grain Procurers] (1928) *"Yezhegodnik khlebnoy torgovli N1 [Yearbook of grain trade N 1]"*, Table 14.
- 1926: Tsentralnoye Konventsionnoye Byuro Khlebozagotoviteley [Central Conventional Bureau of Grain Procurers] (1928) *"Yezhegodnik khlebnoy torgovli N1 [Yearbook of grain trade N 1]"*, Table 22.
- 1927: Statisticheskoye izdatelstvo TsSU SS.S.R. [Statistical Publishing House of the Central Statistical Office of the U.S.S.R.] (1929) *"Selskoye khozyaystvo SS.S.R. 1925–1928. Sbornik statisticheskikh svedeniy k XVI Vsesoyuznoy partkonferentsii [Agriculture of the U.S.S.R. 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 SS.S.R. [People's Commissariat of Supply of the U.S.S.R.] (1932) *"Yezhegodnik khlebooborota N4 [Yearbook of grain turnover N 4]"*, Tables 3 and 10.
- 1930: Narodnyy Komissariat Snabzheniya SS.S.R. [People's Commissariat of Supply of the U.S.S.R.] (1932) *"Yezhegodnik khlebooborota N4 [Yearbook of grain turnover N 4]"*, Table 29 and Table 36
- 1931: Komitet po zagotovkam S.-Kh produktov pri SNK SS.S.R. [Committee for Procurement of Agricultural Products under the Council of People's Commissars of the U.S.S.R.] (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 SS.S.R. [Committee for Procurement of Agricultural Products under the Council of People's Commissars of the U.S.S.R.] (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 SS.S.R. [Committee for Procurement of Agricultural Products under the Council of People's Commissars of the U.S.S.R.] (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 SS.S.R. [Statistical publishing house of the Central Statistical Office of the U.S.S.R.] (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 SS.S.R. i RSFSR. Ekonomiko-statisticheskiy sektor [State Planning Committee of the U.S.S.R. and the RSFSR. Economic and statistical sector] (1931) *"Kolkhozy v 1929 g. Itogi sploshnogo obsledovaniya kolkhozov [Collective farms in 1929. Results of a comprehensivy survey of collective farms]"*.
- 1930: Gosplan SS.S.R.. Upravleniye Narodnokhozyaystvennogo Ucheta [State Planning Committee of the U.S.S.R.. 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. Kolkhoztsentra SS.S.R. i RSFSR [Publishing House of the Collective Farm Center of the U.S.S.R. and the RSFSR] (1931) *"Kolkhoznoye stroitelstvo v SS.S.R. [Collective farms building in the U.S.S.R.]"*, p. 15 and Davies and Wheatcroft (2004), Table 27.
- 1932: RGAE 1562/82/271.
- 1933: *"Plan. Zhurnal Gosplana i TsUNKhU SS.S.R. [Plan. Journal of the State Planning Committee and TsUNKhU U.S.S.R.]"*, 2-1933.
- 1934–1936: RGAE 1562/82/271.
- 1937: interpolated between 1936 and 1938.
- 1938: Gosplanizdat (1939) *"Selskoye khozyaystvo Soyuza S.S.R. 1939 (Statictichеskiy spravochnik) [Agriculture of the U.S.S.R. 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) *"Vserosssiyskaya perepis chlenov RKP 1922 goda [All-Russian census of the members of the RCP in 1922]"*, Issue 3, Table 6.
- 1927: Statisticheskiy otdel TsK VKP(b) [Statistical Department of the Central Committee of the CPSU(b)] (1927) *"Vsesoyuznaya partiynaya perepis 1927 goda. Chislennyy 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 SS.S.R. (Statisticheskiy spravochnik) [Collective farms of the U.S.S.R. (Statistical handbook)]
- 1929: horse power of tractors belonging to collective farms and to machine-tractor stations from Gosplan SS.S.R. i RSFSR. Ekonomiko-statisticheskiy sektor [State Planning Committee of the U.S.S.R. and the RSFSR. Economic and statistical sector]

(1931) *"Kolkhozy v 1929 g. Itogi sploshnogo obsledovaniya kolkhozov [Collective farms in 1929. Results of a comprehensivy survey of collective farms]"*, Tables 1 and 2.

- 1930: horse power of tractors belonging to collective farms is from Gosplan SS.S.R.. Upravleniye Narodnokhozyaystvennogo Ucheta [State Planning Committee of the U.S.S.R.. 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)]";* horse power of tractors belonging to machine-tractor stations is from Tsentralnoye Upravleniye Narodnokhozyaystvennogo Ucheta Gosplana SS.S.R. [The Central Statistical Administration of Gosplan] (1935) *"Sotsialisticheskoye stroitelstvo SS.S.R. (Statisticheskiy yezhegodnik), 1935 g. [Socialist construction of the U.S.S.R. (Statistical Yearbook), 1935]"*, Part II.6, Table 3.
- 1931–1934: Tsentralnoye Upravleniye Narodnokhozyaystvennogo Ucheta Gosplana SS.S.R. [The Central Statistical Administration of Gosplan] (1935) *"Sotsialisticheskoye stroitelstvo SS.S.R. (Statisticheskiy yezhegodnik), 1935 g. [Socialist construction of the U.S.S.R. (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.

### Religious composition

Religious composition is from the 1897 Population Census, available at ristat.org .

### Shares of repartition commune land and private land

Data on commune and private land ownership are originally from the 1905 land census. We calculate province shares from district (*uezd*)-level figures and district (*uezd*)-level map from Castañeda Dower and Markevich (2020).

## A.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: State archive of the Russian federation (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.

### 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.

## A.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.

### Births and 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]"*.

# B 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.1). This is consistent with the lack of major technological changes in the Soviet agriculture before the 1930s argued by historians (Allen, 2003).

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

| | 1926 census | | 1939 census | |
|---|---|---|---|---|
| | Population (mil.) | % Population | Population (mil.) | % Population |
| | (1) | (2) | (4) | (5) |
| | 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. Regression 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 |

*Source: the 1926 and 1939 Population Censuses.*

Table A.2: The Effect of Weather and Natural Conditions on Grain Production

| | Depentent 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 prvince 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. The Data Appendix presents the source of every variable.

Table A.3: The Effect of Ukrainian Population Share on Famine Mortality – Robustness to Controlling for Historical Institutions

| | Dependent Variable: Mortality in year t+1 | | | | | |
|---|---|---|---|---|---|---|
| | Baseline | | Baseline with info on land 1905 | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Ukrainians x Famine | 0.051*** | 0.049*** | 0.035*** | 0.041*** | 0.043*** | 0.034*** |
| | (0.008) | (0.004) | (0.005) | (0.007) | (0.009) | (0.005) |
| Share of catholics 1897 x Famine | | -0.123*** | | | | |
| | | (0.045) | | | | |
| Share of orthodox christians 1897 x Famine | | -0.017 | | | | |
| | | (0.020) | | | | |
| Share of peasant land in repartition commune 1905 x Famine | | | | 0.010 | | |
| | | | | (0.007) | | |
| Share of peasant households in repartition commune 1905 x Famine | | | | | 0.011 | |
| | | | | | (0.008) | |
| Peasant and private land gini 1905 x Famine | | | | | | -0.012 |
| | | | | | | (0.010) |
| Observations | 337 | 337 | 286 | 286 | 286 | 286 |
| R-squared | 0.783 | 0.885 | 0.898 | 0.899 | 0.899 | 0.900 |

*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 control for the share of Catholics in the population according to the 1897 Census interacted with the famine indicator and for the share of Orthodox Christians in the population according to the 1897 Census interacted with the famine indicator. Estimates in Column 3 replicate the baseline on a restricted sample of 16 provinces for which data from the 1905 Land Census is available (omitted are the Urals, and Western and Eastern Siberia). Estimates in Column 4 control for the share of peasant land in repartition commmune according to the 1905 Land Census interacted with the famine indicator. Estimates in Column 5 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 6 control for land gini coefficient from the 1905 Land Census interacted with the famine indicator. All estimates control for Urbanization, Urbanization x Famine, Predicted grain, Predicted grain x Famine, Province- and Year fixed effects. Standard errors robust to heteroskedasticity are in brackets. The Data Appendix presents the source of every variable.

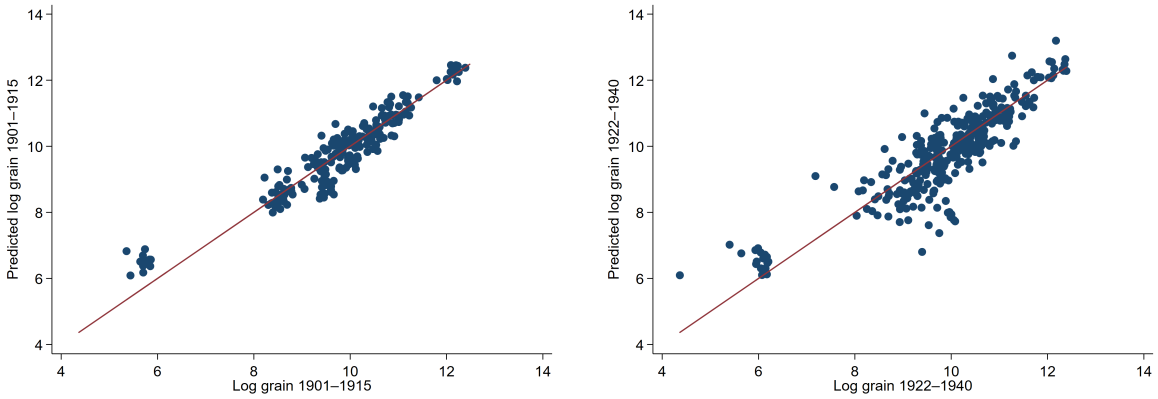
Table A.4: Heterogeneous Effects of Political Factors and Ukrainian Population Share on Rural Horse Power – Controlling for the lagged number of work horses

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

Figure A.1: Reported and Predicted Grain

(a) In-Sample Fit, $R^2 = 0.90$ (b) Out-of-Sample Fit, $R^2 = 0.77$




*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 B for details., The Data Appendix presents the source of every variable.