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

Andrei Markevich†, Natalya Naumenko‡and Nancy Qian§

(Incomplete)

This paper documents several new facts about the Soviet Great Famine, 1932–33. There was no aggregate food shortage. Regional mortality rates were unrelated to per capita food production, but positively associated with ethnic Ukrainian population share. Political loyalty to and peasant resistance against the regime were positively associated with famine mortality and state food procurement in regions populated by ethnic Ukrainians. The findings show that, all else equal, ethnic Ukrainians suffered disproportionally high famine mortality and imply ethnic bias in famine-era policies. A back-of-the-envelope calculation indicates that ethnic bias against Ukrainians explains 77% of famine deaths in Russia, Ukraine and Belarus, and 92% in Ukraine.

JEL: N4, P2 **Keywords:** Strategic Mass Killing, Ethnic Conflict

*We are extremely grateful to Ekaterina Zhuravskaya for her generous support; and thank Maxim Ananyev, Paul Catañeda Dower, Michael Ellman, Ruben Enikolopov, Scott Gehlbach, Mikhail Golosov, Mark Harrison, Oleg Khlevniuk, Alexey Makarin, Victor Malein, Joel Mokyr, Gerard Padro-i-Miquel, Steven Nafziger, Marco Tabellini, Chris Udry for their insights; the participants of the Trinity College Applied Economics Workshop, CERGE-EI Applied Economics Workshop, EIEF workshop, HECER (Helsinki Center of Economic Research) Labor and Public Economics seminar, 2020 Summer Workshop in the Economic History and Historical Political Economy of Russia, Northwestern Economic History Lunch, Northwestern Development Lunch, George Mason Economic History Lunch, the Ridge Conference, the ASSA/AEA On the Role of and Culture and Community 2020 session. For excellent research assistance, we thank Pavel Bacherikov, Nikita Karpov, Ivan Korolev and Carlo Medici. 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] A striking feature of the Soviet Great Famine is the concentration of mortality for Ukrainians: overall Soviet famine mortality rates were 3.7 to 7%, while in the Soviet Republic of Ukraine (Ukraine henceforth), mortality rates were 9 to 13.4%. If one examines mortality by ethnicity, ethnic Ukrainians were only 21% of the pre-famine Soviet population, but constituted at least 30% to 45% of total famine deaths.[3]

The causes of the famine have been a topic of much controversy. Some emphasize that mortality rates were similarly high for all ethnicities during the Soviet Great Famine.[4] Others acknowledge that some groups suffered higher mortality, but argue that this was due to bad weather and the unintended effects of Soviet policies that reduced production: e.g., ethnic Ukrainians were more concentrated in agriculturally productive areas, which suffered larger declines in harvest.[5] This explanation ultimately assumes that the immediate cause of the famine was the drop of food production. In contrast, Sen (1981) argues that famines in the 20th century are caused not by aggregate food shortages, but by unequal distribution: the political rulers control the food and the oppressed populations starve. This view is shared by studies which argue that the Soviet famine was a "terror" intentionally waged by the government on the peasantry, in particular, on the Ukrainian peasantry.[6]

The main goal of our study is to make progress in understanding the causes of the Soviet

[1] See the Background Section.

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

[3] See Background Section for the details of the calculation. Another group that suffered very high famine mortality was the Kazakhs. We do not study this group because there are no famine-era mortality data from Kazakhstan, where most Kazakhs lived.

[4] See Kondrashin (2008).

[5] See, for example, Davies and Wheatcroft (2004) and Kotkin (2017).

[6] See, for example, Conquest (1986), Graziosi (2015), Ellman (2007) and Mace (2004).

Great Famine and the degree to which Soviet policies contributed to famine mortality. The controversial debate thus far has reached an impasse because of the lack of direct documentary evidence that Stalin "ordered" a famine. Our paper fills the gap in the literature by using data to evaluate competing hypotheses and to connect policy to famine mortality.

The main difficulty is the lack of representative sub-national data. Existing arguments are based on narrative or descriptive evidence. To address this, we collect data on the Soviet Union in the interwar period (1922–40). We use data from archival and published sources, as well as geo-spatial weather and agricultural suitability data to construct province- and district-level panels that cover the three republics of the U.S.S.R. — Russia, Ukraine and Belarus. To the best of our knowledge, these are the largest and most comprehensive dataset for this period of the Soviet Union. The large sample size and panel structure allow us to distinguish between competing hypotheses by controlling for confounding factors, and to provide rigorous empirical evidence on the causes of the famine by exploring a rich set of heterogeneous treatment effects.

We proceed in several steps. First, we conduct a food accounting exercise to show that enough food was produced to have avoided the famine had distribution been equal. We document that reported per capita production was approximately *six* times the level necessary to avoid deaths from starvation; and that the amount of food left in the rural areas after government procurement (food that is transferred to urban populations or exported) is subtracted was approximately *five* times more than the level necessary to avoid rural famine mortality. Thus, inequality in food distribution existed not only between urban and rural areas, but also across rural areas, and famine mortality cannot be solely due to the transfer of grain from rural to urban areas or for exports.[7]

Second, we document spatial patterns between famine mortality rates and food produc-

[7]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 result of sufficient aggregate food supplies holds if we use Davies and Wheatcroft's (2004) correction for potential over reporting of grain production in the official statistics. It also holds separately for each of the three republics for which we have mortality data.

tion. Consistent with the central role of food inequality, we find that both the mean and the variance of mortality rates across regions increase during the famine.[8] We then regress mortality rates on lagged grain production and its interaction with the famine year indicator, the interaction of pre-famine Ukrainian population share and the famine year indicator, while controlling for urban population share and its interaction with the famine year indicator, and for province and year fixed effects. The results show that mortality is unrelated to per capita food production during the famine year, which goes against market failures and the lack of transportation infrastructure as the key driving force of famine.[9] Mortality is positively correlated with Ukrainian population share during the famine year, but uncorrelated during non-famine years. Since we control for lag grain production and its interaction with the famine year dummy, the results mean that for two regions with the same harvest, the region with a larger share of ethnic Ukrainians suffered higher mortality during the famine.

The main caveats for interpreting these results are measurement error and omitted variables. To address the measurement error, we show that our result is robust to many alternative measures of Ukrainian population share, mortality, and to using natality rates as an alternative proxy for famine severity. We also use predicted grain production (based on weather and geographic conditions) to mitigate concerns about data quality in the official production figures. To address potential omitted variables, we show that the baseline estimates are robust to controlling for pre-famine geographic, climatic, demographic, institutional and political characteristics, such as the scale of the *dekulakization* campaign against wealthy peasants. We also show that the results are robust if we exclude Ukraine from the sample, which means that ethnic rather than administrative boundaries are relevant for the famine. We also docu-

[8] There are no ethnic-specific mortality or food production data.

[9] If the famine was due to the failure in credit or insurance markets, then we should observe higher famine mortality in regions that produced less food per capita (i.e., the victims were unable to produce food, and unable to borrow and purchase food from surplus producers). Similarly, if the famine was due to the lack of transportation infrastructure (i.e., the government was unable to provide adequate relief) or the destruction of the grain market by the communists, we should see higher famine mortality in regions that produced less food per capita.

ment that there is no relationship between Ukrainian population share and famine mortality in the 1892 famine, the last large famine during the Czarist regime, which goes against the possibility that endemic features of ethnic Ukrainians (e.g., genetics, culture) make them more vulnerable to famines.

Next, we provide direct evidence of the role that Soviet policies played in higher famine mortality in Ukrainian areas. We estimate the heterogeneous treatment effects of several variables that proxy for the zealousness with which economic policies were enforced (e.g., grain productivity in 1928) and political loyalty to the central regime (e.g., Bolshevik vote share in the last free election of 1917, the number of rural communists who procured food from peasants, the number of local party secretaries sent to the 1930 Party Congress).[10] We find that the proxies for loyalty and zealousness were more positively associated with famine mortality in regions with larger Ukrainian population. This implies that higher Ukrainian famine mortality was a consequence of Soviet policies being systematically more aggressively enforced towards ethnic Ukrainians. We find suggestive evidence that the presence of ethnic Ukrainian Party bureaucrats slightly moderated the zealousness that increased famine mortality.

We also show that peasant resistance to the Soviet regime in 1931–32 was positively associated with mortality in 1933. The correlation is stronger for the areas with higher share of ethnic Ukrainians. We provide suggestive evidence that, conditional on the degree of implementation of Soviet agricultural policies, Ukrainian regions resisted more. These findings are in line with the interpretation of the famine as a governmental response to peasant hostility to collectivization.

To further connect policy to mortality, we examine the main rural Soviet economic policies as the dependent variables. Consistent with the view that Soviet policy was an important contributor to Ukrainian famine mortality, the estimates for agricultural collectivization and food procurement rates have the same signs as those for mortality. Interestingly, when we ex-

[10] See the paper for a detailed discussion.

amine mechanization as the outcome, our estimates have the opposite signs. Thus, Ukrainians experienced the negative, but not the positive, aspects of collectivization.

Finally, we supplement the main province-level analysis with a district-level analysis. The district-level panel contains two years, 1928 and 1933, and allows us to show that the cross-province patterns between famine mortality and Ukrainian population share also exist *within* provinces. Since Soviet policies were centrally planned and implemented top-down, this pattern is consistent with the famine being a result of a systematic policy. The granularity of the data also allows us to show that the discrete downward jump in mortality rates when crossing the border from Ukraine to Russia disappears if we control for Ukrainian population share. This emphasizes that ethnic and not administrative boundaries mattered for the famine.

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

Our results 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. They also support the view that high Ukrainian and, thus, overall famine mortality is a result of policy. Since our estimates take as given the factors which contributed to grain production or those that made the population particularly vulnerable to grain shortage in 1932, our results understate the role of state policy in the famine. In the conclusion, we provide a speculative discussion of the possible origins of ethnic bias in famine-era policies.

We are closely connected to two literatures. First, we add to studies on the causes of famine. Earlier works have examined contexts such as China (e.g., Li and Yang, 2005; Meng, Qian, and Yared, 2015), India (e.g., Sen, 1981; Burgess and Donaldson, 2017), and Ireland (e.g., Ó Gráda, 1999).[11] The debate about state culpability in the Soviet famine has been

[11]Li and Yang (2005) estimates the dynamic effects of China's Great Leap Forward policies on the Chinese Great Famine, 1959–61. Meng, Qian, and Yared (2015) documents that there was no aggregate food shortage during the Chinese Great Famine, mortality was positively associated with food production, and attribute part of the famine mortality to centralized food procurement policies. Sen (1981) argues that the Bengal famine was

inconclusive because of the lack of historical documents with direct evidence. The empirical analysis in our study addresses this limitation by using data to test competing explanations and connect famine mortality to policy. Our findings complement Naumenko (2021), which documents a positive association between collectivization and famine mortality across Ukrainian districts.

Second, our study adds to a small but rapidly growing literature on the Russian and Soviet political economy during the 19th and 20th centuries. In providing evidence on the political economic determinants of ethnic-specific persecution, we are most similar to Grosfeld, Sakalli, and Zhuravskaya (2020), which finds that anti-Jewish pogroms during 1800–1927 were triggered by a break of borrower-lender relationships in times of political turmoil. Our study complements Gregory, Schröder, and Sonin (2011); Castañeda Dower, Markevich, and Weber (forthcoming), which explain why dictators, such as Stalin, would kill citizens who are not real enemies; and Egorov and Sonin (2011), which considers the tradeoffs for dictators like Stalin, in maximizing the loyalty of followers.[12] Also related are studies which show that historical features have had important persistent effects over time, such as in the case of the abolition of serfdom (Markevich and Zhuravskaya, 2018; Buggle and Nafziger, 2019), forced migration (Bauer, Braun, and Kvasnicka, 2013; Becker, Grosfeld, Grosjean, Voigtlander, and Zhuravskaya, 2020), peasant rebellions (Castañeda Dower, Finkel, Gehlbach, and Nafziger, 2018; Finkel and Gehlbach, 2020), mass repressions (Talibova and Zhukov, 2018), and anti-Semitism (Grosfeld, Rodnyansky, and Zhuravskaya, 2013; Acemoglu, Hassan, and Robinson, 2019).[13]

Related to these are studies of modern ethnic conflict in Serbia-Croatia (DellaVigna,

due to unequal food distribution between surplus producers, the failure of credit and insurance markets, and food hoarding by the British Colonial government. Burgess and Donaldson (2017) finds that access to railroads reduced famine severity in Colonial India. Ó Gráda (1999) provides a comprehensive study of the economic and political causes of Irish Famine of 1847. See Ó Gráda (2009) for a discussion of major famines in history.

[12] For a comprehensive overview of the political economy problems faced by autocrats, see Gehlbach, Sonin, and Svolik (2016) and Egorov and Sonin (2020).

[13] Also, see Finkel, Gehlbach, and Kofanov (2017) for a study of the causes of peasant rebellions in 1917.

Enikolopov, Mironova, Petrova, and Zhuravskaya, 2014) and the cost of ethnic diversity in Russia (Egorov, Enikolopov, Makarin, and Petrova, 2020), which can be viewed as legacies of historical ethnic tensions. Two recent papers that highlight the potential legacy of the famine in Ukraine today are particularly relevant to our study: Korovkin and Makarin (2019) documents the effect of ethnic tensions between Ukrainians and Russians on firms; and Rozenas and Zhukov (2019) documents the impact of famine-induced ethnic tensions on political outcomes. Gorodnichenko and Roland (2017); Roland (2010) document the long-run effects of Communism more generally. In studying the consequences of Soviet economic policies, we complement macro calibrations of Soviet industrialization policies by Cheremukhin, Golosov, Guriev, and Tsyvinski (2017).[14]

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.[15] 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

[14]Note that Cheremukhin, Golosov, Guriev, and Tsyvinski (2017) intentionally excludes the human costs of the famine from their study.

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

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 contributed 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.

Conquest (1986) estimates total famine deaths to be 7 million. Davies and Wheatcroft (2004) estimates 5.5 to 6.5 million deaths. Ellman (2005) cites "'about eight and a half million' victims of famine and repression in 1930–33". Russian famine historian Kondrashin (2008) gives a range between 5 and 7 million victims. Russian historical demographers estimate 7.2 to 10.8 million famine victims (Polyakov and Zhiromskaya, 2000). In 2008, the Russian State Duma postulated that within the territories of the Volga Region, the Central Black Earth Region, Caucasus, Ural, Crimea, Western Siberia, Kazakhstan, Ukraine and Belarus, the estimated famine death toll was 7 million people (State Duma, 2008).

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 efforts. 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). De-classified secret police reports reveal much active resistance from peasants, mostly in the form of arson, killing communist officials in the rural areas, demonstrations, or the dissemination of anti-Soviet leaflets. Wealthier, more productive peasants, or those actively resisting collectivization, were persecuted as kulaks. As a part of *dekulakization* campaign, approximately two million peasants were exiled to Siberia and other remote areas, amongst whom approximately 500,000 perished (Viola, 2007).

Collective farms worked in teams. Food was produced and stored by the collective, and later delivered to state procurement facilities. 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 —– that is, a 100% tax on food production surplus. In practice, food was procured even if peasants were left with below subsistence amounts. There are many documents showing that Stalin advocated for over-procurement as a method to discipline the peasants, whom he believed to have intentionally understated their production capacity (Danilov, Manning, and Viola, 1999–2006; Davies and Wheatcroft, 2004).

Collectivization could have contributed to famine through several channels: facilitating higher procurement because it gave the government better control over the harvest, and/or reducing grain production because of poor incentives and the destruction of productive individual farms. In addition, collectivization reduced the traditional buffer savings of food, such as the backyard production of potatoes and the destruction of livestock, 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.[16]

2.3 Ukrainians

As the Introduction discussed, famine mortality rates in Ukraine were much higher than the Soviet average. There are no systematic data on ethnic-specific mortality rates. One way to approximate ethnic Ukrainian mortality rates is to use the most cited total famine death toll for the U.S.S.R., seven million (Conquest, 1986); and mortality rates of 2.6 million (Meslé, Vallin, and Andreev, 2013) to 3.9 million (Rudnytskyi, Levchuk, Wolowyna, Shevchuk, and Kovbasiuk, 2015) for Ukraine. If famine deaths were equally distributed between ethnic Ukrainians (80% of Ukraine) and others, and assuming that no ethnic Ukrainians died outside 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.

The Ukrainian nationalism was a two-edged sword for the Soviet government. On the one hand, the support of an organized group that controlled much of the most agriculturally productive regions was very valuable to the Bolsheviks. A minority party in 1917, the Bolsheviks needed cooperation of other groups for their political survival during the early years of the regime. On the other hand, any form of nationalism threatened the authority of the Soviet government.

Thus, the balance between eliciting the support of ethnic minorities and reducing what the Bolsheviks perceived as their subversive influence was delicate. Moreover, it shifted over time as the Bolsheviks increased their political and administrative power. In 1923, the year following the establishment of the Soviet union, the Soviet government launched a policy of indigenization (*korenizatsiya*) to ensure cooperation of ethnic minorities. The policy aimed to promote national languages and culture in regions where minorities constituted local ma-

[16]Initially, before the famine, the state also aimed to remove all private holdings, including small personal plots for potatoes in the peasants' backyards.

jority. From the very beginning, the fears that such policy also contributed to the rise of nationalism periodically caused delays in implementation. In the late 1920s, the Soviet state began to collectivize agriculture, which increased the tensions among the Bolsheviks over indigenization policy, particularly in agriculturally rich regions (Graziosi, 2015; Martin, 2001).

These tensions increased further during the famine. In 1931, party members began to report food shortages to Stalin asking to reduce procurement as the famine intensified. In 1932, Stalin received multiple reports indicating the reluctance of Party leaders at all levels in Ukraine to facilitate the starvation of so many peasants.[17] In response, Stalin sent special commissions headed by his closest deputies, Vyacheslav Molotov and Lazar Kaganovich, neither of whom were ethnic Ukrainians, to exact the full force of Soviet policies in Ukraine and North Caucasus, the two key grain producing regions, where most ethnic Ukrainians lived (Kotkin, 2017).

Concerns about nationalist opposition to the regime was so strong by the autumn of 1932, that the indigenization policy was *de facto* terminated (Graziosi, 2015; Martin, 2001). On December, 14, 1932, the Politburo of the Communist party and the Soviet government issued a classified decree in which the government insisted on complete fulfillment of grain procurement in Ukraine, North Caucasus and the Western region and required the arrests of communists and local officials who failed in this task. In the same decree, the communist leaders accused Ukrainian nationalists within the Communist Party and local bureaucracy of sabotaging grain procurement. The decree required regional authorities in Ukraine, the North Caucasus, and the Western region to crush any resistance of "counter-revolutionaries" and nationalists and to fulfill procurement quotas (Danilov, Manning, and Viola, 1999–2006, Volume 3, Document 226).

Historians note that Stalin believed that the Ukrainian peasantry was one of the essential

[17]E.g., in a letter to his deputy Lazar Kaganovich from August 11, 1932, Stalin mentioned that party district committees in about fifty districts in Ukraine had spoken out against state procurement quotas; he expressed his concerns that the Soviet government 'could lose Ukraine' (Davies, Khlevniuk, Rees, Kosheleva, and Rogovaya, 2003).

difficulties for collectivization because "the national question [is], in essence, a peasant question" (Stalin, 15 April 1925 as quoted in Graziosi 2015).The view that Ukrainian resistance to collectivization was particularly strong is consistent with the descriptive evidence. Table 1 regresses a per capita measure of peasant resistance to collectivization against the population share of Ukrainians in rural areas, the share of households that are collectivized in 1931–32, the interaction of these two variables, and the urban population share using a cross-section of nineteen provinces.[18] The positive coefficient for collectivization shows that peasant resistance was higher in provinces that experienced more collectivization. The positive interaction coefficient shows that resistance to collectivization is increasing in the share of Ukrainians in rural areas.[19]

The main question that our empirical analysis asks is whether such a view translated into systematic persecution of ethnic Ukrainians during the famine, as has been argued by scholars such as Conquest (1986), Ellman (2007), Graziosi (2015), Mace (2004) and Snyder (2010).

3 Food Accounting

The goal of this exercise is to estimate per capita food production and per capita food requirements for the Soviet Union and to examine whether production was sufficient to avoid the famine. The most important source of food, which was also the main target of government procurement, was grain. The data we use come from archival and published sources.[20] 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 2). Row (6) presents reported procurement as a share of production and shows

[18]The dependent variable is a principal component of the three measures of peasant resistance from declassified Soviet Secret Police reports: terrorist acts, mass demonstrations and leaflets. These measures are discussed later in the paper in Section 5.3. See the Data Appendix for more detailed variable descriptions and sources.

[19]See Section 4.1 for more discussion of the sample.

[20]See Appendix D for the source of each variable.

that it increased over time from 14.9% in 1927 to 30.7% in 1939, with the peak during the famine years, when procurement share was 32.9% and 27.2% in 1931 and 1932 respectively. 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 2 row (7) shows that per capita grain production in 1931 and 1932 were 433kg and 428kg, respectively, lower than the previous four and subsequent three years. However, production in 1929, when there was no famine, was only slightly higher at 465kg. Row (8) converts grain from kilograms to calories per day using calories per one kilogram of Russian grain estimated by Lositskij (1920).

Rows (12) and (13) present two levels for caloric requirements. The first is the "business as usual" measure that maximizes labor productivity and healthy child development. This measure assumes 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 (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 (age and gender) using Soviet official data on relative requirements (Lositskij, 1926) and the 1926 population census data on demographic composition.

The second caloric requirement is the "staying alive" measure. 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 2,427 calories per capita during 1931 and 1932. Per capita grain production in row (8) for these years, 3,716 and 3,675 calories, are 152% and 151% higher than these requirements. Row (13) shows that to avoid mortality, the U.S.S.R. required 621 and 622 calories on average. In 1931 and 1932, grain production was 599% and 591% higher than these requirements.

It is important to discuss the reliability of the historical data. The raw data used to gener-

ate 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. One important exception, however, is the aggregate grain production in the early 1930s. Grain production was viewed as the key, and more importantly, public, indicator 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.[21]

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.[22] To investigate whether aggregate grain 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.[23] We convert retention into calories in row (11). These calculations show that average rural grain retention was 128% and 141% of the "business as usual" threshold and 503% and 553% of the "staying alive" threshold. Thus, aggregate procurement of food to supply urban areas and exports cannot explain the famine. For the famine to have occurred, there must have been unequal food distribution across the rural population.

Appendix Table A.2 repeats the exercise for Ukraine. Rural per capita grain retention

[21]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.

[22]Grain exports during 1931–32 were approximately five million metric tons, 7% of the total production (Nove, 1992).

[23]These data are counts of actual procurement and not estimates. See the Data Appendix for references.

during the famine is always higher than the food required to avoid famine.

There are several caveats to keep in mind. The first is the concern of measurement error in the data that we discussed earlier. 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 or waste that are large enough in magnitude to overturn the main point that the famine would not have occurred if food was equally distributed across the population, or the rural population.

4 Spatial Inequality

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

4.1 Mortality

We observe mortality rates at the province level from 1923 to 1940. We adjust these and all other province variables to the 1932 provincial units. Our sample includes nineteen provinces from the three most populous republics of the Soviet Union: Belarus, Ukraine and Russia. Altogether, the sample includes 84% of the 1926 Soviet population and 88% of the 1928 Soviet grain production.[24] The average province has 6.5 million people in 1926.

Note that the data misreporting issues discussed in the previous section all focus on aggregate (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,

[24]The only Russian regions not covered by our panel are Far East, Yakutia, and the republics of North Caucasus. These regions comprised less than 3% of the 1926 population of Russia. 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 Russia.

we will carefully address the possibility of misreporting when relevant in the paper.

Figure 1a plots mortality rates (the number of deaths divided by total population) in the whole sample and in Ukraine only from 1923 to 1940. It shows that mortality rates for the Soviet Union are reasonably constant over time at approximately twenty per thousand, but spike in 1933 to nearly forty per thousand. Mortality rates are usually smaller in Ukraine than in the rest of our sample, but the 1933 spike is much higher for Ukraine, where it increased from approximately eighteen to nearly sixty per thousand.[25]

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 2a plots the standard deviation in mortality rates across provinces normalized by mean mortality for 1923–40. It shows that there is always variation in mortality across provinces, but it increases dramatically during the famine. Thus, the increase in mortality during the famine was accompanied by an increase in inequality in mortality across regions. This is consistent with the aggregate food accounting exercise in suggesting that food availability was very unequal during the famine.

4.2 Natality

To address the possibility that the mortality data may be misreported, we repeat the estimates with natality data: live births should be decreasing with the famine severity.[26] We present figures analogous to those shown for mortality. Figure 1b 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 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

[25] Appendix Figure A.1a maps famine excess mortality, defined as mortality in famine 1933 year minus mortality in "normal" 1928 year, for the provinces in our sample. Ukraine along with the southern provinces of Russia suffer much higher excess mortality than other regions.

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

starving were unable to become pregnant in 1933 and to give birth in 1934.

Figure 2b plots the standard deviation of natality 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 that spatial inequality increased during the famine.

5 Explaining Unequal Famine Intensity

5.1 Food Production and Mortality

The empirical patterns thus far 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 so unequal in 1932 that some regions produced surplus while others suffered harvest failures, and the government was unable or unwilling 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 production and mortality.

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

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

Mortality rate in province i during year $t+1$ is a function of: per capita grain production in province i during year t, $Grain_{it}$; its interaction with a dummy variable that equals one in the famine year, $Famine_t$; province fixed effects, η_i; and year fixed effects δ_t. The additional

controls, X_{it}, include the urban population share in province i during year t and its interaction with the famine dummy variable. The latter variable accounts for the fact that the urban and rural areas were treated very differently by food policies.[27] Our baseline defines the famine dummy to be equal to one in 1932 because 1933 was the year with the highest mortality rates when the famine became apparent in all regions. We estimate standard errors adjusted for spatial correlation.[28]

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* ($\widehat{\beta + \gamma} \approx 0.12$). This contradicts traditional explanations that attribute starvation to low food production. The positive association between grain production and mortality in the famine year is similar to the Chinese Great Famine (1959–61), which Meng, Qian, and Yared (2015) document and hypothesize is the outcome of the centrally planned procurement system. 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 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.[29] 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 mortality were due to weather conditions, or the inability of the government to transport grain from surplus production regions to deficit production regions. This rejects the "despite-best-intentions" theory of the famine.[30]

[27] We control for a time-varying urbanization measured at the province and year level. The results are similar if we control for urbanization reported by the 1926 Census interacted with the famine dummy. These results are available upon request.

[28] We follow the recommendations by Colella, Lalive, Sakalli, and Thoenig (2019) in choosing to adjust for spatial correlation within 1,500 kilometers (the mean province width in our sample is 1,300 km).

[29] See the Appendix Section B.

[30] In fact, it is difficult for anything other than government policies (e.g., market failures) to generate a positive association between mortality rates and food productivity. For example, consider the classical example of credit

5.2 Ethnic Ukrainian Share and Mortality

Motivated by the debate over potential ethnic bias against Ukrainians, this section examines the relationship between Ukrainian population share and famine mortality. Recall that there is no ethnic-specific data on mortality or food production. Thus, we examine the association between famine mortality and ethnic Ukrainian population share.

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

$Ukrainian_i$ is the share of ethnic Ukrainians in 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. The other variables have the same definition as before, with the exception that per capita grain production in province i during year t, $Grain_{it}$, and its interaction with the famine dummy variable are now included in the vector of controls, X_{it}.

Our main measure of ethnic composition comes from the 1926 Soviet census, which is commonly viewed as one of the highest quality Soviet censuses (Andreev, Darskij, and Kharkova, 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. In Ukraine, ethnic Ukrainians comprised 23.2 out of 29.2 million citizens, and an additional 7.9. million ethnic Ukrainians lived in Russia and Belarus. The second-largest ethnic minority, Belorussians, constituted only 3% of the total population. Moreover, while Ukrainians were an ethnic minority in the national population, they were the largest ethnic groups in "grain-producing" areas (where production exceeded local subsistence levels in non-famine years). Ukrainians comprised 45% and Russian comprised 42% of the

or insurance market failures. The argument is that harvest shocks lower the wages of peasants in stricken areas, so that they cannot buy food from those who produced surpluses. The lack of credit and insurance markets makes them unable to smooth food consumption during the shock and survive the famine. In these cases, mortality should always be higher in places that produce less food.

population of grain-producing areas.[31]

Table 3 columns (3) and (4) control for reported and predicted grain production, 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 is statistically significant at the 1% level. Moreover, the interaction of grain production and the famine dummy is reduced in magnitude to statistically zero. Note that the standardized coefficients of the interaction of grain and famine dummy in columns (3) and (4) are smaller than those for the interaction of Ukrainians and famine dummy. Thus, even if they were precisely estimated, they would be economically less important than the ethnic Ukrainian share.

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.[32] 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 the famine year. Thus, urbanization does not play a large role in explaining the spatial inequality of famine mortality. The baseline estimates will always control for urbanization and its interaction with 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 central procurement system, cannot explain the Soviet famine. Instead, the more important factor is ethnic Ukrainian population share.

[31] Appendix Table A.1 Panel A lists ten largest ethnic groups in the Soviet Union, Panel B lists ten largest ethnic groups in our sample, and Panel C lists ten largest ethnic groups in grain-producing provinces of our sample. "Grain-producing" classification is from the official grain procurement statistics.

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

Taken literally, column (4) implies that in a province that was 100% ethnically Ukrainian, famine mortality rate would have been higher than in a province with no ethnic 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.

In columns (5) and (6), we account for pre-famine collectivization rate and for the exile of *kulaks* which occurred just prior to the famine, in 1930–31. Collectivization and the elimination of the wealthiest and most productive peasants could have reduced production in a way that is not fully accounted for in our predicted grain measure. We address this by adding an interaction of the famine period dummy and collectivization (measured as the share of rural households in collective farms in 1932) and *dekulakization* (measured as the number of kulak households exiled from each region in 1930–31 divided by the 1930 population). The Ukrainian population share interaction coefficient is robust to these controls.[33,34]

Column (7) presents a baseline that replaces the province fixed effects with an uninteracted Ukrainian population share variable. This allows us to observe the relationship between Ukrainian population share and mortality during non-famine years and to address the concern that province fixed effects over control by absorbing relevant cross-section variation. Column (7) shows that the latter is not a problem since the interaction coefficient is identical to the baseline in column (4). It also shows that the uninteracted Ukrainian coefficient is -0.007 and statistically significant at the 1% level. This means that in non-famine years, Ukrainian

[33]The collectivization interaction term is positive and statistically significant that is in line with findings in Naumenko (2021). We do not control for collectivization rate in the baseline because collectivization itself was a policy outcome. The kulaks interaction term is positive but is not precisely estimated. Appendix Table A.4 investigates alternative measures of the Soviet dekulakization campaign.

[34]Another important aspect of the pre-famine economy is the large drop in livestock during the years of forced collectivization. Appendix Table A.5 re-estimates the baseline controlling for various measures of the drop in livestock from 1929 to 1932 interacted with the famine indicator. Our estimates are robust; we do not include the drop in livestock in the baseline because, like collectivization and dekulakization, it is a policy outcome.

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 (and all controls) with dummy variables for all years instead of only 1932. Each Ukrainian interaction with year t coefficient 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 3 shows a striking pattern. Prior to the famine, there was little difference in mortality rates across regions.[35] However, regions with a higher share of Ukrainians began to experience higher mortality in 1932 and this difference peaked in 1933. This pattern is consistent with historical accounts of a small famine after the 1931 harvest, which was greatly exacerbated after the 1932 harvest. Afterwards, regions with higher shares of Ukrainians had mortality rates similar to other regions.[36]

5.2.2 Alternative Measures

The baseline uses total mortality as the dependent variable because this variable is available for a larger sample than rural or urban mortality, and rural Ukrainian population share as the explanatory variable because the famine was driven by agricultural policies targeted at the rural population. Table 4 examines the sensitivity of our estimates to alternative ways of measuring Ukrainian population share and mortality. In Panel A, column (1) restates

[35]Note that this is different from Table 3 column (7), which shows that Ukrainian population share is negatively associated with mortality in non-famine years because the specification estimated here uses the baseline controls, including province fixed effects.

[36]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).

the baseline. Columns (2) and (3) replace total mortality with urban and rural mortality, respectively. The results confirm that higher famine mortality in regions with a larger share of ethnic Ukrainians was mostly a rural phenomenon. The estimate for urban mortality in column (2) is small in magnitude and statistically insignificant. The estimate in column (3) for rural mortality is large and statistically significant at the 1% level. Figure 5 presents the year by year estimates for rural mortality. As with total mortality, it shows a sharp increase in the association between mortality and Ukrainian population share during the famine years.

Columns (4) to (7) of Table 4 use different measures of our main independent variable. Our results are nearly identical if we use the total share of Ukrainians in column (4). In column (5), we use urban Ukrainian population share; the coefficient is positive, statistically significant and larger than the baseline. The increase in magnitude is mechanical and is due to the fact that the share of urban Ukrainians is smaller than the rural or total shares. This can be seen in the similarity of the standardized coefficients presented in italics in the table. Columns (6) and (7) use the share of people whose mother tongue is Ukrainian according to the 1926 and 1897 Population Censuses. The estimates are robust to these alternative measures. Henceforth, we will use the 1926 rural Ukrainian population share as the explanatory variable.

Since the first signs of famine were documented after the 1931 harvest, in column (8), we define the famine dummy variable to be equal to one in 1931 and 1932. The interaction coefficient is smaller in size, but still large, positive and statistically significant at the 1% level. The decrease in magnitude is due to the smaller difference in mortality between regions with higher Ukrainian population share and regions with lower Ukrainian population share in 1932, as we show in the previous section.

We will discuss Panel B estimates later.

5.2.3 Demographic and Geographic Controls, Influential Observations

Table 5 examines the sensitivity of our estimates to additional demographic and geographic controls and to the omission of outliers or specific regions. Panel A column (1) restates the baseline. Column (2) controls 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 is significant at the 1% level. Thus, higher famine mortality in areas with more ethnic Ukrainians is not driven by the differences in the demographic composition between Ukrainian-populated regions and other regions.[37]

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, column (3) controls for the triple interaction of latitude, longitude and the famine dummy and all lower-term interactions. Our estimate is similar with this control.

In column (4), we weight the regression by a province-year population so that the estimated coefficient is population weighted. The results are similar. In column (5), we exclude the most influential observations, defined by Cook's Distance (omit observations with a distance greater than $4/n$, where n is the number of observations in the sample).[38] While this reduces the mean share of ethnic Ukrainians in the sample from 10% to 2%, the point estimate of the Ukrainian coefficient remains almost exactly the same as in the full sample. In column (6), we exclude Ukraine, where 75% 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 the 1% level. This means that the baseline results are not

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

[38] See Belsley, Kuh, and Welsch, 1980.

driven by a comparison of differences between Ukraine and other Soviet regions. Instead, they reflect systematically higher mortality rates for regions with a higher share of *ethnic* Ukrainians. In column (7), in addition to excluding Ukraine, we exclude three other regions where food production was particularly concentrated: Lower Volga, North Caucasus, and West Siberia.[39] Our results are robust.

To see whether higher famine mortality in areas with higher shares of ethnic Ukrainians 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 famine of 1892, the last large famine in the late Russian empire, using province-level mortality data from 1885 to 1913.[40] Column (8) estimates our baseline specification for this earlier famine. We find that 1892 famine mortality is not associated with Ukrainian population share. This result is important because it goes against the hypothesis that higher ethnic Ukrainian mortality is due to endemic features of Ukrainians that could lower their resistance to famine.[41]

[39] Appendix Table A.6 also controls for the interactions of various proxies of pre-Soviet regional wealth with the famine dummy variable. These proxies are the nominal regional income per capita in 1897, real regional income per capita in 1897, regional labor productivity in 1897, regional rural labor productivity in 1897 (upper and lower bound estimates) from Markevich (2019), the value of agricultural equipment in 1910, the number of horses in 1916, the number of cows in 1916, livestock in 1916 (from Castañeda Dower and Markevich, 2018), and the first principal component of these measures. Our estimate of the Ukrainian interaction coefficient is robust.

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

[41] Another way to test the relevance of long-run cultural or institutional features of ethnic Ukrainians is to control for these variables in the baseline. One important historical institution in this context is the repartition commune. Living in one required cooperative behavior, and, according to the 1905 land census, these communes were less widespread in Ukrainian-populated regions than in Russian-populated regions. If the values of cooperation were transmitted intergenerationally, this difference could contribute to the difference in mortality between the two ethnicities. In addition, we control for other potentially important variables such as the share of serfs in 1858, three years before the abolition of serfdom, 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 (see Appendix for precise references to the data sources). Appendix Table A.7 shows that our results are robust if we add interactions of these variables with the famine dummy into the baseline specification.

5.2.4 Natality

As before, we address the concerns that the mortality data are mismeasured by using natality as an alternative outcome variable. In tables 4 and 5, Panel B present the same specification as in Panel A with natality as the dependent variable. The estimates are all negative and statistically significant at the 1% level. Figure 4 plots the coefficients from the dynamic estimates. They trace the temporal patterns of the raw natality data shown earlier in Figure 3, which is consistent with Ukrainian population share being an important determinant in birth rates during the famine. Note that the timing of the effect of famine on natality is not as sharp as it is for mortality. This is consistent with the fact that natality is a more sensitive measure, it responds to more moderate levels of food deprivation which do not yet lead to excess mortality. This means that we can alternatively use a wider time window to define the famine year dummy for the natality estimates. The results are qualitatively similar and are available upon request.

5.3 Controlling for Policy and Political Factors

Table 6 additionally controls for the interactions of the famine year dummy with proxies for Soviet policies, loyalty to the regime, which may have influenced the implementation of official policies, and peasant resistance to the Soviet regime in 1931–32. Panels A and B report the estimates with mortality and natality as the dependent variables, respectively. Our main coefficient of interest is robust to these additional controls. We will discuss estimates in Panel A in detail. Column (1) restates the baseline for comparison.

First, we examine measures that are widely believed to have contributed to the intensification of Soviet agricultural policies during the famine. Column (2) controls for per capita grain production in 1928, i.e., the first year 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). In line with this observation, the coefficient on per capita grain production interaction is positive and statistically significant at the 10% level.

Column (3) proxies for the population's commitment and/or loyalty to the Bolsheviks with vote share from the 1917 Constituency Assembly election. The 1917 election was a universal election, the first and only until the end of the Bolshevik rule. Approximately 60% of the population turned out to vote.[42] The coefficient on the 1917 Bolsheviks vote share interaction is negative and statistically significant at the 5% level. This is consistent with the view on the famine as a penalty imposed to the disloyal population.

Column (4) controls for the number of rural Communist Party Members (averaged over 1922, 1927 and 1931) per one thousand individuals in each province. Party members were the key enforcers of state policy in the countryside, and we interpret this measure as a proxy for state capacity in the village.

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 of the local elites to the regime or commitment to agricultural collectivization.

Next, we explore how mortality in 1933 was related to peasant resistance to the Soviet regime from January 1931 to March 1932, i.e. in the period preceding the famine. Column (6) controls for the number of anti-Soviet violent acts per 1,000 people. These acts include murders or attempted murders of local officials, arsons and destruction of collective farm property. The Soviet secret police, which registered these events, classified them as "terrorist acts." Column (7) controls for the number of mass demonstrations in the countryside per

[42] We follow Castañeda Dower and Markevich (2020) and use disaggregated district-level data on votes for the Bolsheviks from Protasov, V.V. Zhuravlev, and Shelokhaev (2014). See the Data Appendix for details.

1,000 people. Column (8) accounts for the number of anti-Soviet leaflets uncovered by the secret police per 1,000 people. Column (9) controls for the first principal component of the three resistance measures. Higher level of resistance is positively associated with famine mortality. This positive association is statistically significant and is consistent with the Soviet regime punishing regions disloyal in 1931–32 with the policies that led to a more severe famine in 1933.[43]

Panels B examines the same specifications as in panel A with natality as the dependent variable. The signs of the estimates are consistent with our interpretation of the mortality estimates.

We note that some of the additional interaction controls are statistically significant. Average effects mask important heterogeneity which we explore in the next section.

5.4 Heterogeneous Effects on Mortality and Natality

This section investigates the heterogeneous effects of policy measures and political factors on famine mortality in areas with high Ukrainian population share. Table 7 estimates the fully saturated triple interaction specification of these variables on mortality. The estimates capture the effect of the policy or political variable on famine mortality rates in a province that is 100% Ukrainian relative to a province with no Ukrainians. The interaction of Ukrainian population share and the famine dummy variable is not of primary interest in this table since it captures excess mortality rates for Ukrainians in provinces where the political variables of interest has a value of zero. To save space, we do not report the coefficients on interactions of political variables and the famine dummy variable, which capture the effects of political variables in a province with zero Ukrainian population share. To be consistent with the earlier estimates, the triple interaction specifications also control for the triple interaction

[43]Note that we do not construct a principal component measure of the policy variables because they are conceptually different. In contrast, the resistance proxies all measure peasant resistance to collectivization.

of urbanization, Ukrainian population share and famine dummy to account for the possible correlation between urbanization and the political variables, and for the triple interaction of predicted grain, Ukrainian population share and famine dummy to account for possible correlation between harvest and political variables.

Panel A examines mortality. Column (1) shows that the positive association between agricultural productivity in 1928 and the famine mortality is increasing with Ukrainian population share.

Column (2) indicates that, in contrast to an average region, Bolshevik vote share is positively associated with mortality for regions with a high Ukrainian population share.

Column (3) examines the number of Communist Party members per capita in rural areas in the years prior to the famine, which reflects the administrative capacity of the central government in the countryside. We find that rural communists increase mortality in areas with more ethnic Ukrainians. This is consistent with stronger enforcement of Soviet policies towards ethnic Ukrainian communities.[44]

Column (4) examines the effect of the number of Party Secretaries, who participated in the 1930 Party Congress and voted for collectivization. We are able to identify the ethnicity of these Party Secretaries.[45] Given historical accounts of ethnic Ukrainian Party members opposing the harsh policies during the famine (Kotkin, 2017), we distinguish the effects of the delegates who were Ukrainians themselves versus the delegates who were of other ethnicities.[46] Both triple interactions are positive and statistically significant. Moreover, the triple

[44]This result holds if we add the triple interaction of the famine dummy, Ukrainian population share and the number of urban communists per capita and all lower-level interactions. We divide communists into rural and urban because they were responsible for implementing different policies that would have had opposite effects on mortality. Rural communists were tasked with procurement, while urban communists were tasked with distributing the procured food to the urban population. During the famine, urban communists were also in charge of providing relief to urban residents, including peasants who fled to urban areas for food. Consistent with this division of labor, we find that urban communists moderated famine mortality in areas populated by more Ukrainians. These results are available upon request.

[45]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.

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

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 15% level (p-value for equality of the coefficients is 0.16). These results are consistent with a lack of formal ethnic division within the Communist Party, but also with the documentary evidence that ethnic Ukrainian Party members relented more when the famine became severe.[47]

Column (5) investigates the heterogeneous effects of peasant resistance. For brevity, we examine the first principal component of the three resistance measures.[48] Peasant resistance increases famine mortality more in regions with more Ukrainians. This is consistent with the Soviet government using the famine to respond to political opposition, and enacting its policies with more aggression in the areas with more Ukrainians.

Next, we explore the heterogeneous effects of Soviet agricultural policies on famine mortality directly. Column (6) examines the effect of collectivization, measured as average share of collectivized households in a region in 1932, and column (7) examines the effect of dekulakization campaign in 1930–31. Positive and significant coefficients on both triple interactions are in line with the view that these policies were implemented more zealously in regions with a higher share of ethnic Ukrainians that transformed into higher famine mortality.

To assess the magnitude of the estimated effects, in each column, we report predicted mortality at the mean of the political variable for two types of hypothetical provinces: for a hypothetical province with a rural share of ethnic Ukrainians at the 25th percentile (0.1% Ukrainians) and at the 75th percentile (11.4% Ukrainians). For a province with the political variable at the mean, the increase in the share of Ukrainians leads to a sharp increase in famine mortality. Note that in column (4) the magnitudes are statistically significant only for non-Ukrainian delegates.

[47]These results are also consistent with the view that Stalin often sent "outsiders" to govern to counteract local loyalties and nationalism (Gregory, 2009).

[48]The results on the heterogeneous effects of each of three peasant resistance measures are qualitatively similar and are presented in the Appendix Table A.8.

Panel B examines natality. The coefficients vary slightly in precision, but the signs mirror those for mortality. Note that in column (4) 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.

We examine the dynamic triple interaction effects by repeating the estimates in Table 7, but replacing the famine dummy variable with year dummy variables. Figure 7 shows the timing of the effects of these policy proxies. The timing is sharp. 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.[49]

5.5 Soviet Agricultural Policies

To further connect policy to famine mortality, we examine agricultural collectivization and the different components of collectivization as the dependent variables. Collectivization was the main economic policy for rural areas. It was supposed to increase procurement rate by strengthening state control over harvests, as well as boost production by increasing economies of scale and mechanization.

Table 8 presents the same heterogeneous treatment estimates as in Table 7 (with the obvious exception of the 1932 collectivization). Panel A examines the share of households in collective farms as the dependent variable. The estimates for collectivization vary slightly in precision, but the signs of the estimates are the same as those for mortality, which is consistent with collectivization being a key contributor to famine mortality.

Panel B examines the amount of tractor horsepower per capita, our proxy for mechanization.[50] The signs are opposite of those for mortality and collectivization. The estimates

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

[50]Increasing the number of tractors was central to the mechanization effort.

are statistically significant at the 1% level. They imply that the policy factors and political loyalty to the regime, which caused Ukrainian-populated regions to suffer higher mortality and more intense collectivization, also caused these regions to receive less mechanization. In other words, areas with ethnic Ukrainians systematically received less investment for mechanization, arguably the only positive aspect of collectivization.

In Panel C and D, we examine reported grain procurement, measured as a share of reported production and per capita, as the dependent variable. The results are qualitatively similar to those for collectivization and mortality, and statistically significant at the 1% level. These findings are consistent with the notion that excess procurement played an important role in causing famine mortality in Ukrainian regions.[51]

5.6 District-Level Analysis

The district-level panel consists of two years: 1928 and 1933. All data are manually collected from the former Soviet archives.[52] 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 Ukraine side of the border between Russia and Ukraine.[53] We define excess mortality as the difference between 1933 and 1928 mortality rates. Figure 8a plots 1933 excess mortality against the distance to the border between Ukraine and Russia. 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 8b, which plots the residual mortality against distance to the border. These results are consistent with our

[51] We also examine reported grain production as an outcome (but we do not control for contemporaneous predicted grain production or its interaction terms for this specification). The signs are the opposite of those for mortality, collectivization and procurement, but the estimates are imprecise. They are available upon request.

[52] See the Appendix for more details about the data.

[53] The government introduced a ban on migration from Ukraine and from North Caucasus region in January 1933 (Danilov, Manning, and Viola, 1999–2006, Vol. 3).

interpretation that the Soviet policies which led to the famine targeted ethnic Ukrainians rather than Ukraine.

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) replicates the baseline specification adding province-year fixed effects.[54,55] We estimate standard errors adjusted for spatial correlation.[56] This allows us to isolate the within province variation and account for province-specific factors (e.g., regional political competition, leadership in specific provinces). The results are robust. Finding similar patterns at different levels of bureaucracy is consistent with the presence of a systematic and centrally planned policy.

Columns (2) to (13) show that the results are robust to alternative definitions of ethnic Ukrainians, additional controls and to omitting influential observations from the sample.

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 —– if the interaction coefficient of Ukrainian population share and the famine dummy variable in equation (2) was zero. Using the estimates from equation (2), Table 3 column (4), we predict that the number of deaths in non-

[54]Unfortunately, political variables are only available at the province level so we cannot estimate heterogeneous effects at the district level.

[55]Note that we use urbanization from 1926 and 1933 because urbanization at the district-level is not available for 1928.

[56]We follow the recommendations by Colella, Lalive, Sakalli, and Thoenig (2019) and adjust for spatial correlation within 400 kilometers (the mean district width in our sample is 76 km); the distance of 400 km delivers the largest standard errors.

famine years is on average 2.70 million, and in 1933 is 4.97 million.[57] The difference, 2.27 million ($4.97 - 2.70 = 2.27$ million) is the number of 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 the number of deaths in non-famine years, 0.52 million ($3.22 - 2.70 = 0.52$ million). It follows that ethnic bias accounts for 77% ($1 - .52/2.27 = .77$) of famine deaths in our sample. See Appendix Table A.9 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 total famine deaths to be seven million, and subtract the deaths in Kazakhstan (1 to 1.5 million) and Ukraine (2.6 to 3.9 million), we are left with 1.6 to 3.4 million. This results in famine mortality rates of 14 to 30 per 1,000 for the 112 million residents of Russia. A similar calculation for Ukraine, which had a population of 32 million, yields a famine mortality rate of 81 to 122 per 1,000. Since the back-of-the-envelope exercise conceptually asks what mortality would have been had ethnic Ukrainians died at the same rate as other ethnicities in our sample (who are mostly ethnically Russian), it is then not surprising that ethnic bias contributes to such a large proportion of mortality.

We can repeat the exercise for Ukraine, Russia and Belarus separately. We find that in Ukraine during non-famine years, predicted deaths are 0.52 million. Predicted deaths in 1933 are 2.03 million. The difference, 1.50 million ($2.03 - 0.52$ million, note that there is a small discrepancy due to rounding), is the number of excess deaths due to the famine. If we assign the Ukrainian interaction coefficient a value of zero, we predict deaths in 1933 to be 0.64 million. Famine deaths without ethnic bias would have been the difference between this number and mortality in non-famine years, 0.12 million ($0.64 - 0.52 = 0.12$). Thus, ethnic bias accounts for 92% ($1 - 0.12/1.50 = 0.92$) of famine deaths in Ukraine. Since

[57] This is close to 4.81 million 1933 deaths reported in our sample.

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 Ukraine.

Repeating this exercise for Russia and Belarus, we find that ethnic bias against Ukrainians explains 48% of famine deaths in Russia and 12% of famine deaths in Belarus. Since only 8% of the population of Russia and 0.7% of Belarus were ethnic Ukrainian, these estimates again imply disproportionately high mortality rates among ethnic Ukrainians in these two republics.

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, was decimated. The empirical evidence from our paper fills some of the important gaps from the large body of historical and descriptive evidence put together by scholars over the past 35 years. We show that ethnic Ukrainians systematically suffered higher famine mortality rates, even if we compare regions that were similar along all other dimensions. Furthermore, we provide direct evidence linking policy and the zealousness in its enforcement to famine mortality.

These findings prompt several questions for future inquiry. First and foremost is why the Soviet government targeted ethnic Ukrainians. Based on the historical evidence, one possible explanation is that the Soviet government needed to repress resistance to agricultural collectivization, one of its key economic policies. Thus, it was important to control ethnic minorities, who felt less loyalty to the Russian-dominant government than ethnic Russians, and particularly those who concentrated in agriculturally productive regions. Viewed this

way, through the central planning problem, the government will repress ethnic Ukrainians because they were the second largest ethnic group after Russians, 21% of the total population, and the largest group in grain-producing regions, 45% of the population there. This political-economically motivated explanation of repression is consistent with the fact that the Bolsheviks, ironically, embraced an ideology that minimized ethnic distinctions. It is also consistent with the theoretical insights from Mitra and Ray (2014) on the origins of conflict between Hindus and Muslims, which have historically coexisted peacefully; and the predictions from Caselli and Coleman II (2013) and Esteban, Morelli, and Rohner (2015) that ethnic conflict and strategic mass killings are more likely with high levels of natural resources (agriculture in the Soviet famine context). There are potentially other explanations. This is an important avenue for future research since it will help shed light on the evolution of ethnic tensions.[58]

Second, it is important to understand the causes and impact of the famine for other groups, such as the Kazakhs. Kazakhstan, where most ethnic Kazakhs lived at the time of the famine, may have lost up to 22% of its population to famine. The historical evidence suggests that forced agricultural collectivization of this traditionally semi-nomadic groups contributed to extraordinarily high famine mortality rate. Unfortunately, we were unable to examine this group because of current data limitations.

References

ACEMOGLU, D., T. HASSAN, AND J. ROBINSON (2019): "Social Structure and Development: A Legacy of the Holocaust in Russia," *The Quarterly Journal of Economics*, 126, 895–946.

[58]Note that recent studies on the causes of ethnicity-delineated mass killings have found evidence for the role of state capacity in Rwanda (Heldring, 2020; Rogall, 2021); and more specifically, for mass media in Croatia-Serbia (DellaVigna, Enikolopov, Mironova, Petrova, and Zhuravskaya, 2014), Rwanda (Yanagizawa-Drott, 2014), Nazi Germany (Adena, Enikolopov, Petrova, Santarosa, and Zhuravskaya, 2015), and Israel-Palestine (Durante and Zhuravskaya, 2018). Petrova and Yanagizawa-Drott (2016) provides an overview of studies on the role of mass media.

ADENA, M., R. ENIKOLOPOV, M. PETROVA, V. SANTAROSA, AND E. ZHURAVSKAYA (2015): "Radio and the Rise of The Nazis in Prewar Germany," *The Quarterly Journal of Economics*, 130(4), 1885–1939.

ALLEN, R. (2003): *Farm to Factory: A Reinterpretation of the Soviet Industrial Revolution*. Princeton University Press, Princeton.

ANDREEV, E., L. DARSKIJ, AND T. KHARKOVA (1998): *Demographicheskaya Istroriya Rossii: 1927–1959*. Informatika, Moscow.

BAUER, T. K., S. BRAUN, AND M. KVASNICKA (2013): "The Economic Integration of Forced Migrants: Evidence for Post-War Germany," *The Economic Journal*, 123(571), 998–1024.

BECKER, S. O., I. GROSFELD, P. GROSJEAN, N. VOIGTLANDER, AND E. ZHURAVSKAYA (2020): "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers," *American Economic Review*, 110(5), 1430–63.

BELSLEY, D., E. KUH, AND R. WELSCH (1980): *Regression Diagnostics: Identifying Influential Data and Sources of Collinearity*. John Wiley, New York.

BERELOWITCH, A., AND V. DANILOV (eds.) (2000–2012): *Les campagnes Soviétiques vues par la Tchéka–O.G.P.U.–N.K.V.D. 1918–1939. Recueil de documents et matériaux en 4 volumes. Sovetskaya derevnya glazami VChK–OGPU–NKVD. 1918–1939 gg. Dokumenty i materialy v 4 tomakh*. Rosspen, Moscow.

BUGGLE, J., AND S. NAFZIGER (2019): "The Long-Run Consequences of Labor Coercion: Evidence from Russian Serfdom," *The Review of Economics and Statistics*, 87, 217–234.

BURGESS, R., AND D. DONALDSON (2017): "Railroads and the Demise of Famine in Colonial India," Working paper, MIT.

CASELLI, F., AND W. J. COLEMAN II (2013): "On the Theory of Ethnic Conflict," *Journal of the European Economic Association*, 11(s1), 161–192.

CASTAÑEDA DOWER, P., E. FINKEL, S. GEHLBACH, AND S. NAFZIGER (2018): "Collective Action and Representation in Autocracies: Evidence from Russia's Great Reforms," *American Political Science Review*, 112(1), 125–47.

CASTAÑEDA DOWER, P., AND A. MARKEVICH (2018): "Labor Misallocation and Mass Mobilization: Russian Agriculture during the Great War," *The Review of Economics and Statistics*, 100(2), 245–59.

——— (2020): "Democratic support for the Bolshevik Revolution: An empirical investigation of 1917 Constituent Assembly elections," *Working paper*.

CASTAÑEDA DOWER, P., A. MARKEVICH, AND S. WEBER (forthcoming): "The Value of a Statistical Life in a Dictatorship: Evidence From Stalin," *The European Economic Review*.

CHARNYSH, V., AND B. MCELROY (2020): "Explaining outgroup bias in weak states: religion and legibility in the 1891–92 Russian famine," project in progress.

CHEREMUKHIN, A., M. GOLOSOV, S. GURIEV, AND A. TSYVINSKI (2017): "The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model," *The Review of Economic Studies*, 84(2), 613–649.

COLELLA, F., R. LALIVE, S. O. SAKALLI, AND M. THOENIG (2019): "Inference with Arbitrary Clustering," .

CONQUEST, R. (1986): *The Harvest of Sorrow: Soviet Collectivization and the Terror-Famine*. Oxford University Press, Oxford, UK.

DANILOV, V., R. MANNING, AND L. VIOLA (eds.) (1999–2006): *Tragediya Sovetskoj Derevni. Kollektivizatsiya i raskulachivanie. Dokumenti i materialy v 5 tomakh, 1927–1939*. Rosspen.

DASGUPTA, P., AND D. RAY (1986): "Inequality as a Determinant of Malnutrition and Unemployment: Theory," *The Economic Journal*, 96, 1011–1034.

DAVIES, R. W., O. V. KHLEVNIUK, E. A. REES, L. P. KOSHELEVA, AND L. A. ROGOVAYA (eds.) (2003): *The Stalin-Kaganovich Correspondence, 1931–36*. Yale University Press, New Haven.

DAVIES, R. W., AND S. G. WHEATCROFT (2004): *The Years of Hunger: Soviet Agriculture, 1931–1933*, vol. 5 of *The Industrialization of Soviet Russia*. Palgrave Macmillan, Basingstoke, UK.

DELLAVIGNA, S., R. ENIKOLOPOV, V. MIRONOVA, M. PETROVA, AND E. ZHURAVSKAYA (2014): "Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia," *American Economic Journal: Applied Economics*, 6(3), 103–132.

DURANTE, R., AND J. BUGGLE (forthcoming): "Climatic Risk, Cooperation, and the Co-Evolution of Culture and Institutions," *Economic Journal*.

DURANTE, R., AND E. ZHURAVSKAYA (2018): "Attack When the World Is Not Watching? US News and the Israeli-Palestinian Conflict," *Journal of Political Economy*, 126(3), 1085–1133.

DYSON, T., AND C. Ó GRÁDA (eds.) (2002): *Famine Demography. Perspectives from the Past and Present*. Oxford University Press, Oxford, UK.

EGOROV, G., R. ENIKOLOPOV, A. MAKARIN, AND M. PETROVA (2020): "Divided We Stay Home: Social Distancing and Ethnic Diversity," NBER Working Papers 27277, National Bureau of Economic Research, Inc.

EGOROV, G., AND K. SONIN (2011): "Dictators and Their Viziers: Endogenizing the Loyalty-Competence Trade-off," *Journal of the European Economic Association*, 9(5), 903–930.

——— (2020): "The Political Economics of Non-democracy," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, (2020-142).

ELLMAN, M. (2005): "The Role of Leadership Perceptions and of Intent in the Soviet Famine of 1931–1934," *Europe-Asia Studies*, 57(6), 823–841.

——— (2007): "Stalin and the Soviet famine of 1932–33 Revisited," *Europe-Asia Studies*, 59(4), 663–693.

ESTEBAN, J., M. MORELLI, AND D. ROHNER (2015): "Strategic Mass Killings," *Journal of Political Economy*, 123(5), 1087–1132.

FINKEL, E., AND S. GEHLBACH (2020): *Reform and Rebellion in Weak States*, Elements in Political Economy. Cambridge University Press.

FINKEL, E., S. GEHLBACH, AND D. KOFANOV (2017): "(Good) Land and Freedom (for Former Serfs): Determinants of Peasant Unrest in European Russia, March-October 1917," *Slavic Review*, 76(3), 710–721.

GEHLBACH, S., K. SONIN, AND M. W. SVOLIK (2016): "Formal Models of Nondemocratic Politics," *Annual Review of Political Science*, 19(1), 565–584.

GOKMEN, G., AND D. KOFANOV (2020): "Career incentives in political hierarchy: evidence from Imperial Russia," *European Review of Economic History*, 24(2), 264–287.

GORODNICHENKO, Y., AND G. ROLAND (2017): "Culture, Institutions, and the Wealth of Nations," *The Review of Economics and Statistics*, 99(3), 402–416.

GRAZIOSI, A. (2015): "The uses of hunger Stalin's solution of the peasant and national questions in Soviet Ukraine, 1932 to 1933," in *Famines in European History. The Last Great European Famines Reconsidered*, ed. by L. L. Declan Curran, and A. G. Newby, pp. 223–260. Routledge.

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

GREGORY, P. R., P. J. SCHRÖDER, AND K. SONIN (2011): "Rational dictators and the killing of innocents: Data from Stalin's archives," *Journal of Comparative Economics*, 39(1), 34–42.

GROSFELD, I., A. RODNYANSKY, AND E. ZHURAVSKAYA (2013): "Persistent Antimarket Culture: A Legacy of the Pale of Settlement after the Holocaust," *American Economic Journal: Economic Policy*, 5(3), 189–226.

GROSFELD, I., S. O. SAKALLI, AND E. ZHURAVSKAYA (2020): "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire," *The Review of Economic Studies*, 87(1), 289–342.

HELDRING, L. (2020): "The Origins of Violence in Rwanda," *The Review of Economic Studies*, rdaa028.

KONDRASHIN, V. (2008): *Golod 1932–1933 godov: Tragediya Sovetskoj derevni*. Rosspen, Moscow.

KOROVKIN, V., AND A. MAKARIN (2019): "Conflict and Inter-Group Trade: Evidence from the 2014 Russia-Ukraine Crisis," *EIEF Working paper*.

KOTKIN, S. (2017): *Stalin. Waiting for Hitler, 1929–1941*. Penguin Press, New York.

LI, W., AND D. YANG (2005): "The Great Leap Forward: Anatomy of a Central Planning Disaster," *Journal of Political Economy*, 113(4), 840–877.

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

——— (1926): *Sostoyanie pitaniya gorodskogo naseleniya SSSR, 1919–1924 [The State of Nutrition of Urban Population of the USSR, 1919–1924]*. TsSU.

——— (ed.) (1928): *Sostoyanie pitaniya selskogo naseleniya SSSR v 1920–1924 [The State of Nutrition of Rural Population of the USSR in 1920–1924]*. TsSU.

MACE, J. E. (2004): "Soviet Man-Made Famine in Ukraine," in *Century of Genocide: Critical Essays and Eyewitness Accounts.*, ed. by W. S. P. Totten Samuel, and I. W. Charny. Routledge.

MARKEVICH, A. (2019): "A Regional Perspective on the Economic Development of the late Russian Empire," *Working paper*.

MARKEVICH, A., AND E. ZHURAVSKAYA (2018): "The Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire," *American Economic Review*, 108(4–5), 1074–1117.

MARTIN, T. (2001): *The Affirmative Action Empire: Nations and Nationalism in the Soviet Union, 1923–1939*. Cornell University Press.

MATSUURA, K., AND C. WILLMOTT (2014): "Terrestrial Air Temperature and Precipitation: 1900–2014 Gridded Monthly Time Series (Version 4.01)," *National Oceanic and Atmospheric Administration*.

MENG, X., N. QIAN, AND P. YARED (2015): "The Institutional Causes of China's Great Famine, 1959–1961," *The Review of Economic Studies*, 82(4), 1568–1611.

MESLÉ, F., J. VALLIN, AND E. ANDREEV (2013): "Demographic Consequences of the Great Famine: Then and Now," in *After the Holodomor: The Enduring Impact of the Great Famine on Ukraine*, ed. by A. Graziosi, L. A. Hajda, and H. Hryn. Harvard University Press.

MITRA, A., AND D. RAY (2014): "Implications of an Economic Theory of Conflict: Hindu-Muslim Violence in India," *Journal of Political Economy*, 122(4), 719–765.

NAUMENKO, N. (2021): "The Political Economy of Famine: the Ukrainian Famine of 1933," *The Journal of Economic History*, 81(1), 156–197.

NOVE, A. (1992): *An Economic History of the U.S.S.R*. The Penguin Press.

Ó GRÁDA, C. (1999): *Black '47 and Beyond: The Great Irish Famine in History, Economy, and Memory*. Princeton University Press, Princeton, NJ.

——— (2009): *Famine: A Short History*. Princeton University Press, Princeton, NJ.

PETROVA, M., AND D. YANAGIZAWA-DROTT (2016): "Media Persuasion, Ethnic Hatred, and Mass Violence: a Brief Overview of Recent Research Advances," in *Economics of Genocide, Mass Killing, and Their Prevention*, ed. by C. Anderton, and J. Brauer, chap. 12. Oxford University Press, Oxford.

POLYAKOV, Y., AND V. ZHIROMSKAYA (eds.) (2000): *Naselenie Rossii v XX veke: istoricheskie ocherky. Vol. 1 1900–1939*. Rosspen.

PROTASOV, L., A. R. V.V. ZHURAVLEV, AND V. SHELOKHAEV (2014): *Vserossijskoe uchreditelnoe sobranie. Entsiklopedia. (All-Russian Constitutional Assembly. Encyclopedia)*. Rosspen, Moscow.

ROGALL, T. (2021): "Mobilizing the Masses for Genocide," *American Economic Review*, 111(1), 41–72.

ROLAND, G. (2010): "The Long-Run Weight of Communism or the Weight of Long-Run History?," WIDER Working Paper Series wp-2010-083, World Institute for Development Economic Research (UNU-WIDER).

ROZENAS, A., AND Y. M. ZHUKOV (2019): "Mass Repression and Political Loyalty: Evidence from Stalin's 'Terror by Hunger'," *American Political Science Review*, 113(2), 569–583.

RUDNYTSKYI, O., N. LEVCHUK, O. WOLOWYNA, P. SHEVCHUK, AND A. KOVBASIUK (2015): "Demography of a man-made human catastrophe: The case of massive famine in Ukraine 1932–1933," *Canadian Studies in Population*, 42(1–2), 53–80.

SEN, A. (1981): *Poverty and Famines. An Essay on Entitlement and Deprivation*. Oxford University Press, Oxford, UK.

SNYDER, T. (2010): *Bloodlands: Europe Between Hitler and Stalin*. Basic Books, New York City, NY.

STALIN, I. (15 April 1925): "Concerning the National Question in Yugoslavia," *Bol'shevik*, 7.

STATE DUMA (2008): "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"," .

TALIBOVA, R., AND Y. M. ZHUKOV (2018): "Stalin's terror and the long-term political effects of mass repression," *Journal of Peace Research*, 55(2), 267–283.

TAUGER, M. (2001): "Natural Disaster and Human Actions in the Soviet Famine of 1931–1933," *The Carl Beck Papers in Russian and East European Studies*.

VIOLA, L. (1996): *Peasant Rebels Under Stalin: Collectivization and the Culture of Peasant Resistance*. Oxford University Press, Oxford, UK.

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

WHEATCROFT, S., AND R. DAVIES (1994): "Agriculture," in *The Economic Transformation of the Soviet Union, 1913–1945*, ed. by R. Davies, M. Harrison, and S. Wheatcroft. Cambridge University Press.

WHEATCROFT, S. G. (2001): "O zernovikh balansakh i otsenkakh urozhajnosti v SSSR v 1931–1933," in *Tragediya Sovetskoj Derevni. Kollektivizatsiya i raskulachivanie. Dokumenti i materialy v 5 tomakh, 1927–1939. Vol. 3. Konetz 1930–1933*, ed. by V. Danilov, R. Manning, and L. Viola. Rosspen.

YANAGIZAWA-DROTT, D. (2014): "Propaganda and Conflict: Evidence from the Rwandan Genocide," *The Quarterly Journal of Economics*, 129(4), 1947–1994.

Table 1: Ethnic Composition and Peasant Resistance

	Dependent Variable: Peasant Resistance		
	(1)	(2)	(3)
Ukrainians	0.065		-0.841***
	(0.213)		(0.231)
Collectivization 1931–32		0.490**	0.795***
		(0.183)	(0.189)
Ukrainians × Collectivization			0.788***
			(0.261)
Urbanization 1931–32	-0.358*	-0.378*	-0.473**
	(0.190)	(0.179)	(0.184)
Constant	0.000	0.000	-0.338
	(0.226)	(0.193)	(0.238)
Observations	19	19	19
R-squared	0.134	0.369	0.471

Notes: Observations are at the province level. The dependent variable is the first principal component of the peasant resistance: terrorist acts, mass demonstrations and leaflets. Terrorist acts is the number of violent acts against government officials from January 1931 to March 1932 per 1,000 people. Mass demonstrations is the number of mass demonstrations from January 1931 to March 1932 per 1,000 people. Leaflets is the number of episodes of anti-government leaflets being distributed among the population from January 1931 to March 1932 per 1,000 people. The explanatory variables are the Ukrainian population share in 1926, the average share of households in collective farms during 1931–32, the interaction of these two variables, and the average urbanization rate during 1931–32. The dependent variable and the uninteracted explanatory variables are normalized to have mean equal to zero and standard deviation equal to one. Huber-White robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 2: Per Capita Food Production and Requirements for the U.S.S.R.

	1927	1928	1929	1930	**1931**	**1932**	1933	1937	1939
	A. Production and Procurement								
(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 (mt)	11.1	10.8	16.1	22.2	**22.8**	**19.0**	23.7	31.9	30.9
(6) Procurement rate	14.9%	14.7%	22.4%	26.6%	**32.9%**	**27.2%**	26.3%	26.5%	30.7%
	B. Implied Food Availability								
(7) Production pc (kg per year)	504	487	465	530	**433**	**428**	542	742	609
(8) Production pc (cal. per day)	4,329	4,183	3,994	4,555	**3,716**	**3,675**	4,651	6,373	5,231
(9) DW (cal. per day)	.	.	.	4,089	**3,263**	**3,024**	3,807	.	.
(10) Rural retention pc (kg per year)	523	503	440	479	**364**	**400**	521	803	632
(11) Rural retention pc (cal. per day)	4,486	4,318	3,775	4,113	**3,122**	**3,436**	4,469	6,890	5,425
(12) *Cal. needs pc – heavy labor*	*2,450*	*2,453*	*2,450*	*2,446*	***2,439***	***2,427***	*2,421*	*2,375*	*2,369*
(13) *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 (2004). Conversion from grain to calories are based on estimates from Lositskij (1920). 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 light work (urban) – 2,750 per day (Lositskij 1928), for relative caloric needs of other groups are based on Lositskij (1926). Caloric needs for avoiding mortality are based on the 900 caloreis per day for prime age adult males (Dasgupta and Ray, 1986).

Table 3: Famine Mortality in Ethnic Ukrainian Areas

	Dependent Variable: Mortality in Year t+1						
	Reported Grain	Predicted Grain	Reported Grain	Predicted Grain (Baseline)			Omit Province FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ukrainians × Famine			0.047***	0.051***	0.045***	0.053***	0.051***
			(0.002)	(0.006)	(0.004)	(0.005)	(0.005)
Standardized Coef.			*0.766*	*0.825*	*0.733*	*0.869*	*0.831*
Ukrainians							-0.007***
							(0.002)
Grain	0.004	0.001	0.007	0.003*	0.003*	0.003	-0.001
	(0.006)	(0.002)	(0.005)	(0.002)	(0.002)	(0.002)	(0.004)
Grain × Famine	0.120***	0.079	0.072*	0.003	-0.032**	-0.027	-0.002
	(0.042)	(0.061)	(0.039)	(0.030)	(0.013)	(0.035)	(0.027)
Standardized Coef.	*0.483*	*0.416*	*0.289*	*0.017*	*-0.167*	*-0.140*	*-0.009*
Urbanization	-0.011**	-0.012**	-0.009**	-0.011**	-0.011**	-0.011**	-0.013***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)
Urbanization × Famine	0.014	0.010	0.011	0.005	0.001	-0.002	0.008
	(0.019)	(0.011)	(0.012)	(0.007)	(0.008)	(0.006)	(0.008)
Pre-famine Collectivization × Famine					0.042***		
					(0.011)		
Pre-famine Exiled Kulak HHs × Famine						1.771	
						(1.371)	
Observations	337	337	337	337	337	337	337
R-squared	0.671	0.658	0.804	0.785	0.824	0.799	0.433

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 and zero otherwise. Reported and predicted grain are measured in 10s of kilograms per person per day. Estimates in column 5 control for the share of collectivized rural households in 1932 interacted with the famine indicator. Estimates in column 6 control for the number of kulak households exiled during 1930–31 per 1930 population interacted with the famine indicator. Estimates in columns 1–6 control for province fixed effects. All estimates control for year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Famine Mortality and Natality in Ethnic Ukrainian Areas — Sensitivity to Alternative Measures

	Baseline (1)	Urban Mortality/ Natality (2)	Rural Mortality/ Natality (3)	Total Ukrainians (4)	Urban Ukrainians (5)	Mother Tongue Ukrainian 1926 (6)	Mother Tongue Ukrainian 1897 (7)	Famine = 1931, 32 (8)
	A. Dependent Variable: Mortality in Year t+1							
Ukrainians × Famine	0.051***	0.013	0.063***	0.055***	0.090***	0.056***	0.058***	0.032***
	(0.006)	(0.008)	(0.005)	(0.006)	(0.010)	(0.007)	(0.007)	(0.004)
Standardized Coef.	*0.825*	*0.249*	*0.870*	*0.819*	*0.785*	*0.747*	*0.820*	*0.521*
Observations	337	285	285	337	337	337	337	337
R-squared	0.785	0.746	0.792	0.784	0.780	0.773	0.789	0.737
	B. Dependent Variable: Natality in Year t+1							
Ukrainians × Famine	-0.014***	-0.002***	-0.012***	-0.015***	-0.025***	-0.015***	-0.016***	-0.010***
	(0.003)	(0.0005)	(0.003)	(0.003)	(0.005)	(0.003)	(0.003)	(0.002)
Standardized Coef.	*-0.432*	*-0.149*	*-0.347*	*-0.429*	*-0.409*	*-0.380*	*-0.425*	*-0.317*
Observations	337	285	285	337	337	337	337	337
R-squared	0.835	0.924	0.909	0.835	0.835	0.834	0.835	0.836
Ukrainians								
Mean	0.104	0.104	0.104	0.095	0.055	0.074	0.085	0.104
Std. Dev.	0.216	0.216	0.216	0.197	0.116	0.178	0.188	0.216

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. In columns 2 and 3, the dependent variables are urban mortality/natality, and rural mortality/natality, respectively. In columns 4 and 5, Ukrainian population share is the share of ethnic Ukrainians in the total population and the urban population, respectively. In columns 6 and 7, Ukrainian population share is the share of people whose mother tongue is Ukrainian according to the 1926 and 1897 Censuses, respectively. In columns 1–7, famine is an indicator that equals one in 1932 and zero otherwise; in column 8, famine equals one in 1931 and 1932. All regressions control for urbanization, urbanization × famine, predicted grain, predicted grain × famine, and province and year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table 5: Famine Mortality and Natality in Ethnic Ukrainian Areas — Sensitivity to Demographic and Geographic Characteristics, and Omitting Influential Observations

	Baseline (1)	Control for Demographic Structure x Famine (2)	Control for Latitude x Longitude x Famine (3)	Weight by Population (4)	Omit Outliers by Cook's D (5)	Omit Ukraine (6)	Omit Ukraine, Lower Volga, North Caucasus, West Siberia (7)	1892 Famine (8)
	A. Dependent Variable: Mortality in Year t+1							
Ukrainians × Famine	0.051***	0.048***	0.060***	0.045***	0.048***	0.086***	0.058***	-0.000
	(0.006)	(0.007)	(0.004)	(0.004)	(0.007)	(0.007)	(0.006)	(0.003)
Standardized Coef.	*0.825*	*0.786*	*0.971*	*0.739*	*0.177*	*0.839*	*0.405*	*-0.004*
Observations	337	337	337	337	319	319	268	1,297
R-squared	0.785	0.791	0.814	0.880	0.836	0.767	0.815	0.864
	B. Dependent Variable: Natality in Year t+1							
Ukrainians × Famine	-0.014***	-0.015***	-0.013***	-0.012***	-0.035***	-0.030***	-0.032***	0.004**
	(0.003)	(0.005)	(0.002)	(0.002)	(0.008)	(0.003)	(0.009)	(0.002)
Standardized Coef.	*-0.432*	*-0.457*	*-0.406*	*-0.351*	*-0.284*	*-0.519*	*-0.273*	*0.102*
Observations	337	337	337	337	314	319	268	1,297
R-squared	0.835	0.835	0.837	0.880	0.887	0.830	0.832	0.934
Ukrainians								
Mean	0.104	0.104	0.104	0.104	0.021	0.061	0.028	0.172
Std. Dev.	0.216	0.216	0.216	0.216	0.049	0.113	0.048	0.299

Notes: Observations are at the province and year level. All regressions control for urbanization, urbanization × famine, predicted grain, predicted grain × famine, and province and year fixed effects. Column 2 also controls for the share of people aged 10 and younger and the male/female ratio, each interacted with the famine dummy variable. Column 3 also coentrols for latitude × longitude × famine, and all lower-order interactions. Column 4 estimates the baseline specification, but weighted by population. Column 5 omits observations that have Cook's D higher than 4/337. Column 6 omits the Republic of Ukraine. Column 7 omits the Republic of Ukraine and the provinces of Lower Volga, North Caucasus and West Siberia. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Famine Mortality and Natality in Ethnic Ukrainian Areas — Robustness to Political Controls

		Pre-famine Factors							
	Baseline	X = Grain 1928	X = Bolshevik Votes 1917	X = Rural Communists 1922, 27, 31	X = Delegates in 1930 Congress	X = "Terrorist Acts" 1931–32	X = Mass Demonstrations 1931–32	X = Leaflets 1931–32	X = Resistance Principal Component
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	A. Dependent Variable: Mortality in Year t+1								
Ukrainians × Famine	0.051***	0.055***	0.050***	0.053***	0.050***	0.053***	0.048***	0.052***	0.052***
	(0.006)	(0.006)	(0.006)	(0.004)	(0.006)	(0.006)	(0.003)	(0.004)	(0.003)
X (see column heading) × Famine		0.092*	-0.013**	0.003	0.033	0.057*	0.314***	0.440**	0.005***
		(0.047)	(0.005)	(0.002)	(0.021)	(0.030)	(0.105)	(0.212)	(0.001)
Observations	337	337	337	337	337	337	337	337	337
R-squared	0.785	0.806	0.789	0.796	0.792	0.790	0.815	0.815	0.826
	B. Dependent Variable: Natality in Year t+1								
Ukrainians × Famine	-0.014***	-0.016***	-0.014***	-0.015***	-0.014***	-0.014***	-0.013***	-0.014***	-0.014***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
X (see column heading) × Famine		-0.035***	-0.001	-0.001	-0.012*	-0.007	-0.080**	-0.113*	-0.001**
		(0.011)	(0.003)	(0.001)	(0.007)	(0.011)	(0.035)	(0.061)	(0.001)
Observations	337	337	337	337	337	337	337	337	337
R-squared	0.835	0.837	0.835	0.836	0.835	0.835	0.836	0.836	0.836

Notes: Observations are at the province and year level. 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. Rural communists is the average over 1922, 1927, and 1931 of the number of rural Communist Party members and candidates per 1,000 people. Delegates in 1930 Congress 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. "Terrorist acts" is the number of violent acts against government officials from January 1931 to March 1932 per 1,000 people. Mass demonstrations is the number of mass demonstrations from January 1931 to March 1932 per 1,000 people. Leaflets is the number of times anti-government leaflets were distributed among the population from January 1931 to March 1932 per 1,000 people. The resistance principal component is the first principal component of the three resistance measures (terrorist acts, mass demonstrations, and leaflets). All estimates control for urbanization, urbanization × famine, predicted grain, predicted grain × famine, province and year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Heterogeneous Effects of Political Factors on Famine Mortality and Natality in Ethnic Ukrainian Areas

Pre-famine Factors:	X1 = Grain 1928	X1 = Bolshevik Votes 1917	X1 = Rural Communists 1922, 27, 31	X1 (X2) = Ethnic Ukrainian (non-Ukrainian) Delegates in 1930 Congress	X1 = Resistance Principal Component	X1 = Pre-famine Collectivization	X1 = Pre-famine Exiled Kulak HHs
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	A. Dependent Variable: Mortality in Year t+1						
X1 × Ukrainians × Famine [1]	2.974***	1.038***	0.058***	1.802***	0.049***	0.969***	58.605***
	(0.519)	(0.390)	(0.016)	(0.542)	(0.007)	(0.328)	(14.107)
X2 × Ukrainians × Famine [2]				2.229***			
				(0.336)			
Ukrainians × Famine [3]	-0.172***	0.002	0.232***	-0.035	0.310***	-0.453***	0.192***
	(0.036)	(0.025)	(0.087)	(0.033)	(0.073)	(0.151)	(0.062)
Observations	337	337	337	337	337	337	337
R-squared	0.852	0.821	0.829	0.855	0.854	0.851	0.833
[1] - [2]: p-value				0.161			
[1] + [3] (X1 = Mean, Ukrainians = 25th Percentile)	0.0003***	0.0003**	0.0005***	-0.00003	0.0004***	0.0002**	0.0005***
[1] + [3] (X1 = Mean, Ukrainians = 75th Percentile)	0.026***	0.026**	0.043***	-0.002	0.035***	0.014**	0.043***
[2] + [3] (X2 = Mean, Ukrainians = 25th Percentile)				0.0004***			
[2] + [3] (X2 = Mean, Ukrainians = 75th Percentile)				0.034***			
	B. Dependent Variable: Natality in Year t+1						
X1 × Ukrainians × Famine [1]	-0.798***	-0.430***	-0.023***	-0.079	-0.019***	-0.242**	-15.847***
	(0.266)	(0.090)	(0.004)	(0.338)	(0.003)	(0.103)	(5.316)
X2 × Ukrainians × Famine [2]				-0.437**			
				(0.177)			
Ukrainians × Famine [3]	0.046**	0.000	-0.091***	-0.036	-0.110***	0.112**	-0.055***
	(0.020)	(0.011)	(0.022)	(0.026)	(0.017)	(0.055)	(0.017)
Observations	337	337	337	337	337	337	337
R-squared	0.838	0.838	0.838	0.838	0.839	0.838	0.838
[1] - [2]: p-value				0.039			
[1] + [3] (X1 = Mean, Ukrainians = 25th Percentile)	-0.0001**	-0.0001***	-0.0002***	-0.0001	-0.0002***	-0.00004**	-0.0001***
[1] + [3] (X1 = Mean, Ukrainians = 75th Percentile)	-0.007**	-0.011***	-0.017***	-0.004	-0.012***	-0.004**	-0.012***
[2] + [3] (X2 = Mean, Ukrainians = 25th Percentile)				-0.0001***			
[2] + [3] (X2 = Mean, Ukrainians = 75th Percentile)				-0.012***			

Notes: Observations are at the province and year level. The table presents the triple interaction of Ukrainian population share, the famine year dummy variable, and the variable stated in the column heading. The regressions control for all lower-order interaction terms; Ukrainians × Predicted grain × Famine, Ukrainians × Urbanization × Famine, and their lower-order interactions; and province and year fixed effects. 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. Rural communists is the average over 1922, 1927, and 1931 of the number of rural Communist Party members and candidates per 1,000 people. Ethnic Ukrainian (non-Ukrainian) Delegates in 1930 Congress is the number of province-, county-, city-, and borough-level Party secretaries of Ukrainian (non-Ukrainian) ethnicity that participated and had a right to vote in the 1930 Party Congress per 100,000 people. The resistance principal component is the first principal component of the three resistance measures (terrorist acts, mass demonstrations, and leaflets). Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Heterogeneous Effects of Political Factors on Collectivization, Mechanization and Grain Procurement for Ethnic Ukrainian Areas

Pre-famine Factors:	X1 = Grain 1928	X1 = Bolshevik Votes 1917	X1 = Rural Communists 1922, 27, 31	X1 (X2) = Ethnic Ukrainian (non-Ukrainian) Delegates in 1930 Congress	X1 = Resistance Principal Component	X1 = Pre-famine Exiled Kulak HHs	X1 = Grain 1928	X1 = Bolshevik Votes 1917	X1 = Rural Communists 1922, 27, 31	X1 (X2) = Ethnic Ukrainian (non-Ukrainian) Delegates in 1930 Congress	X1 = Resistance Principal Component
	Dependent Variables										
	A. Collectivization						B. Tractors' Horse Power				
	(1)	(2)	(3)	(4)	(5)	(6)	(8)	(9)	(10)	(11)	(12)
X1 × Ukrainians × Famine [1]	13.293***	7.847***	0.234***	7.529	0.197**	207.761*	-2.654***	-0.969***	-0.047***	-2.125***	-0.042***
	(4.759)	(1.437)	(0.090)	(5.281)	(0.082)	(119.553)	(0.418)	(0.364)	(0.015)	(0.463)	(0.009)
X2 × Ukrainians × Famine [2]				11.864***						-2.216***	
				(3.585)						(0.254)	
Ukrainians × Famine [3]	-0.792*	0.07	1.410**	0.572	1.589**	0.799	0.132***	-0.024	-0.225***	0.045**	-0.294***
	(0.445)	(0.147)	(0.565)	(0.370)	(0.691)	(0.497)	(0.028)	(0.024)	(0.074)	(0.020)	(0.063)
Observations	228	228	228	228	228	228	247	247	247	247	247
R-squared	0.968	0.969	0.969	0.968	0.968	0.968	0.790	0.787	0.788	0.790	0.789
	C. Procurement Rate (Procurement/Production)						D. Procurement per Capita (Procurement/Rural Population)				
X1 × Ukrainians × Famine [1]	10.508***	5.501***	0.328***	14.983***	0.235***	150.279***	2.912***	1.457***	0.081***	3.389***	0.056***
	(3.280)	(1.656)	(0.097)	(4.374)	(0.056)	(51.695)	(0.405)	(0.405)	(0.020)	(0.851)	(0.008)
X2 × Ukrainians × Famine [2]				12.161***						3.257***	
				(2.618)						(0.472)	
Ukrainians × Famine [3]	-0.244	0.401**	1.811***	0.385	1.875***	1.042***	-0.147***	0.032	0.374***	-0.013	0.413***
	(0.258)	(0.192)	(0.366)	(0.364)	(0.312)	(0.213)	(0.032)	(0.042)	(0.088)	(0.059)	(0.041)
Observations	186	186	186	186	186	186	186	186	186	186	186
R-squared	0.883	0.876	0.880	0.880	0.878	0.880	0.814	0.810	0.811	0.811	0.812

Notes: Observations are at the province and year level. The dependent variables are the share of rural households in collective farms (Panel A), tractors' horse power per capita (Panel B), procure production (Panel C), and procurement/rural population (Panel D). The table presents the triple interaction of Ukrainian population share, the famine year dummy variable, and the variable stated heading. The regressions control for all lower-order interaction terms; Ukrainians × Urbanization × Famine, Ukrainians × Predicted grain × Famine and and their lower-order interactions; and pro fixed effects. 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. Rural com average over 1922, 1927, and 1931 of the number of rural Communist Party members and candidates per 1,000 people. Ethnic Ukrainian (non-Ukrainian) Delegates in 1930 Congress is the numbe county-, city-, and borough-level Party secretaries of Ukrainian (non-Ukrainian) ethnicity that participated and had a right to vote in the 1930 Party Congress per 100,000 people. The resist component is the first principal component of the three resistance measures (terrorist acts, mass demonstrations, and leaflets). Standard errors in parentheses are adjusted for spatial correlation wi *** p<0.01, ** p<0.05, * p<0.1

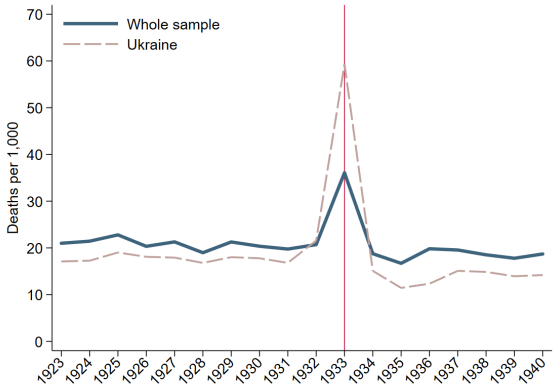
Table 9: Famine Mortality in Ethnic Ukrainian Areas — District-level Analysis

	Dependent Variable: Mortality												
	Baseline with Province-Year FE (1)	Control for Collectivization 1930 × Famine (2)	Urban Mortality (3)	Rural Mortality (4)	Total Ukrainians (5)	Urban Ukrainians (6)	Mother Tongue Ukrainian (7)	Control for gender ratio (8)	Control for Latitude × Longitude × Famine (9)	Weighted by population (10)	Omit Outliers by Cook's D (11)	Omit Ukraine (12)	Omit Ukraine, Lower Volga, North Caucasus, West Siberia (13)
Ukrainians × Famine	0.039***	0.035***	0.008	0.029***	0.044***	0.043***	0.047***	0.040***	0.040***	0.041***	0.034***	0.022***	0.027***
	(0.010)	(0.010)	(0.006)	(0.007)	(0.011)	(0.008)	(0.010)	(0.010)	(0.009)	(0.011)	(0.005)	(0.007)	(0.010)
Standardized Coef.	*0.506*	*0.449*	*0.150*	*0.353*	*0.545*	*0.423*	*0.592*	*0.506*	*0.509*	*0.519*	*0.575*	*0.190*	*0.231*
Observations	3,513	3,076	1,743	3,047	3,515	2,052	3,505	3,513	3,513	3,513	2,871	2,734	2,181
R-squared	0.770	0.792	0.900	0.901	0.773	0.786	0.777	0.770	0.773	0.783	0.905	0.742	0.693
Provinces	18	18	18	18	18	18	18	18	18	18	18	17	14
Mortality 1933													
Mean	0.038	0.038	0.031	0.038	0.038	0.041	0.038	0.038	0.038	0.038	0.030	0.030	0.027
Std. Dev.	0.028	0.029	0.019	0.030	0.028	0.029	0.028	0.028	0.028	0.028	0.015	0.019	0.014
Ukrainians													
Mean	0.237	0.246	0.216	0.239	0.227	0.215	0.219	0.237	0.237	0.237	0.116	0.069	0.034
Std. Dev.	0.363	0.368	0.358	0.363	0.348	0.290	0.358	0.363	0.363	0.363	0.250	0.165	0.119

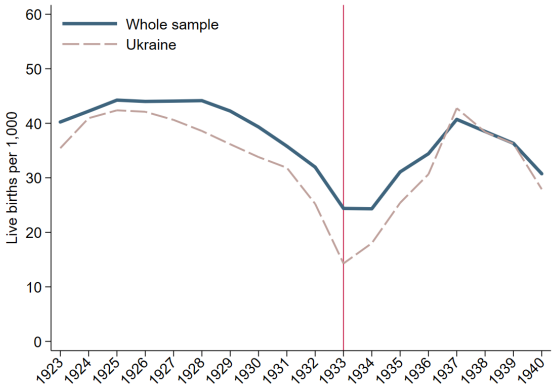
Notes: Observations are at the district and year level (1928, 1933). Mortality is the number of deaths divided by population, measured as total mortality in the baseline, urban mortality in column 3 and rural mortality in column 4. Ukrainians is the share of ethnic Ukrainians in the rural population in the baseline, in the total population in column 5, urban population in column 6, or the share of people whose mother tongue is Ukrainian in column 7. Column 8 controls for the male/female ratio × famine. Column 9 controls for latitude × longitude × famine, and lower-order interactions. Column 10 is weighted by population. Column 11 omits observations that have Cook's D higher than 4/3,513. Column 12 omits Ukraine. Column 13 omits Ukraine and the provinces of Lower Volga, North Caucasus and West Siberia. All regressions control for urbanization, urbanization × famine, grain suitability × famine, and district and province-year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 400 km. *** p<0.01, ** p<0.05, * p<0.1

Figure 1: Mortality and Natality Rates over Time

(a) Mortality



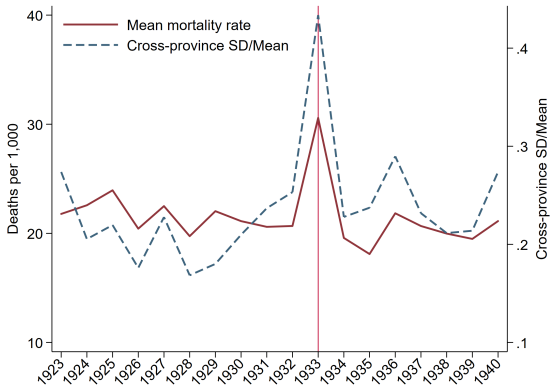
(b) Natality



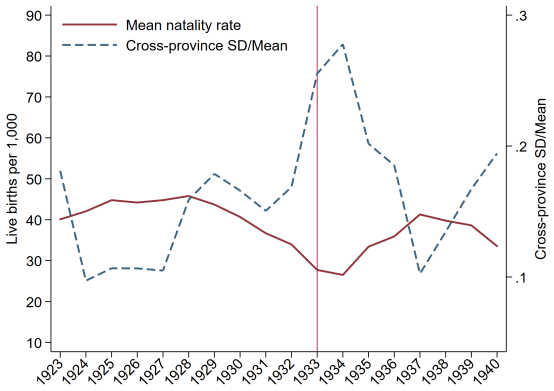
Notes: Mortality is the number of deaths per 1,000 individuals. Natality is the number of live births per 1,000 individuals. *Source:* See the Data Appendix.

Figure 2: Cross-Province Mean and Standard Deviation of Mortality and Natality Rates

(a) Mortality



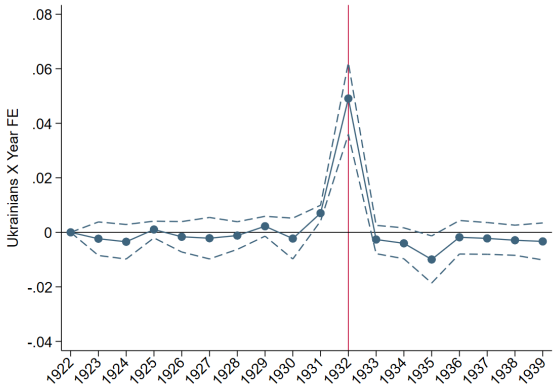
(b) Natality



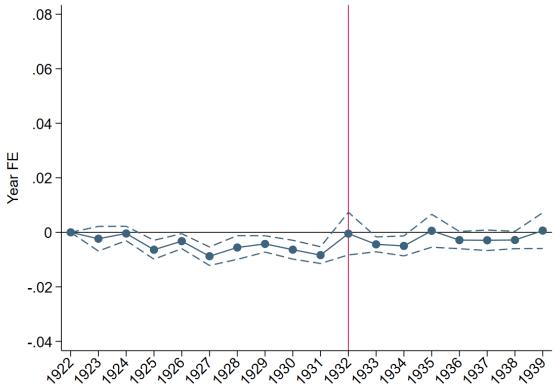
Notes: Mean mortality (natality) rate is the average mortality (natality) rate across provinces in each year. Cross-province SD/Mean is the standard deviation in mortality (natality) rates across provinces in year t divided by the mean mortality (natality) rate in year t. *Source:* See the Data Appendix.

Figure 3: The Dynamic Effect of Ukrainian Population Share on 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 adjusted for spatial correlation within 1,500 km. *Source:* See the Data Appendix.

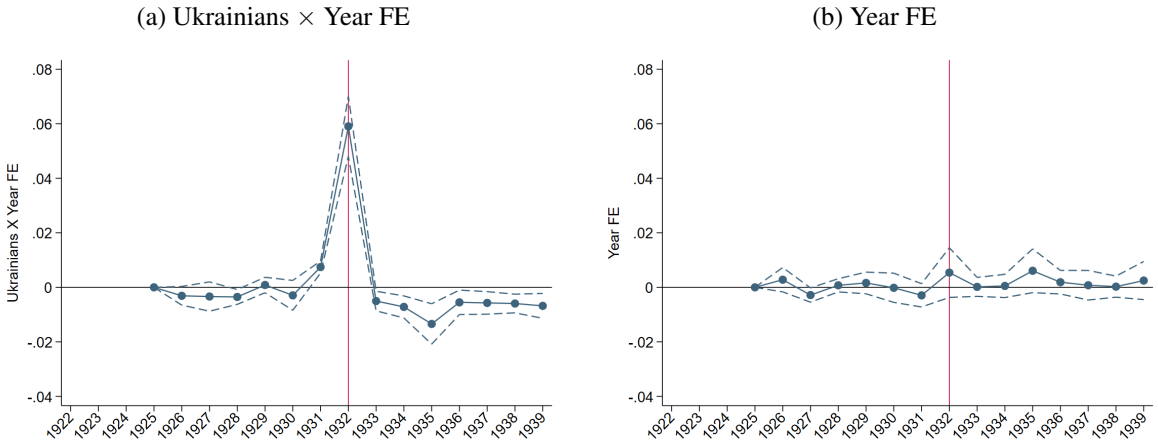
Figure 4: The Dynamic Effect of Ukrainian Population Share on Natality

(a) Ukrainians × Year FE

(b) Year FE

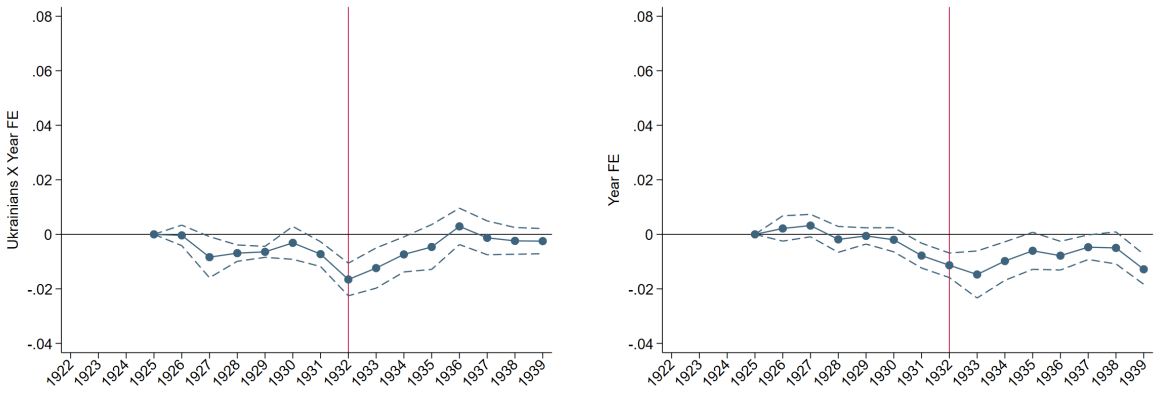
Notes: The figures show regression coefficients with their 95% confidence intervals from regressing natality in year $t + 1$ on the rural share of ethnic Ukrainians interacted with year indicators, urbanization interacted with year indicators, predicted grain interacted with year indicators, year indicators (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 adjusted for spatial correlation within 1,500 km. *Source:* See the Data Appendix.

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



Notes: The figures show regression coefficients with their 95% confidence intervals from regressing rural 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 1925 year indicator is omitted for comparison. The coefficients plotted in figures (a) and (b) are estimated from one regression. Standard errors are adjusted for spatial correlation within 1,500 km. *Source:* See the Data Appendix.

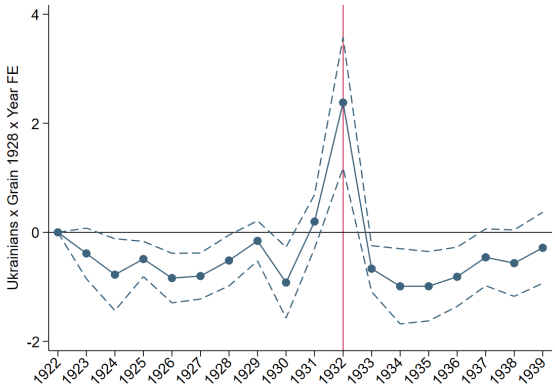
Figure 6: The Dynamic Effect of Ukrainian Population Share on Rural Natality



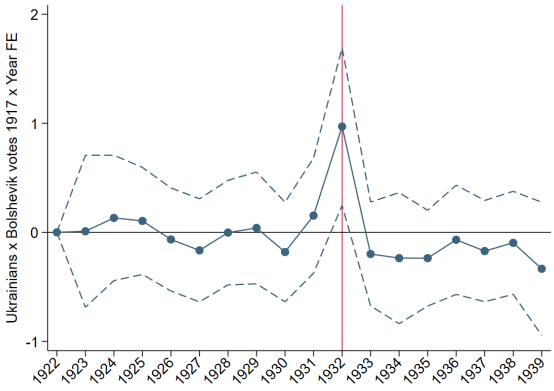
Notes: The figures show regression coefficients with their 95% confidence intervals from regressing rural natality in year $t+1$ on the rural share of ethnic Ukrainians interacted with year indicators, urbanization interacted with year indicators, predicted grain interacted with year indicators, year indicators (fixed effects), and province fixed effects. The 1925 year indicator is omitted for comparison. The coefficients plotted in figures (a) and (b) are estimated from one regression. Standard errors are adjusted for spatial correlation within 1,500 km. *Source:* See the Data Appendix.

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

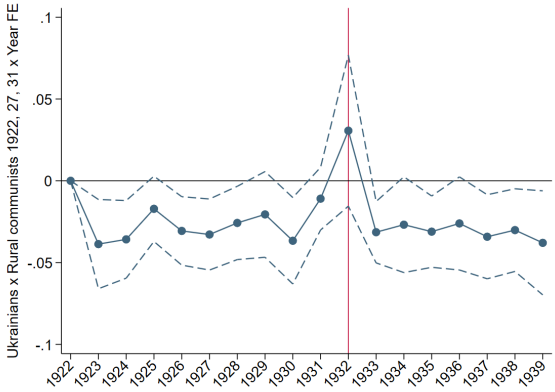
(a) Ukrainians × Grain 1928 × Year FE



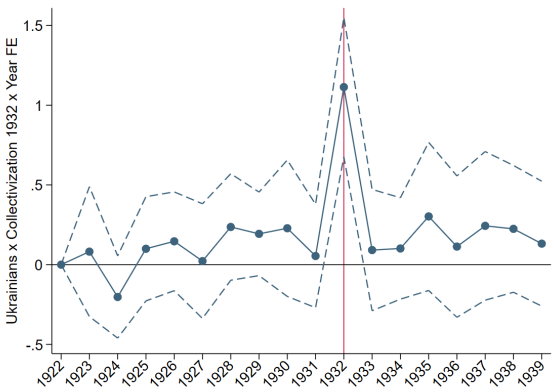
(b) Ukrainians × Bolshevik votes 1917 × Year FE



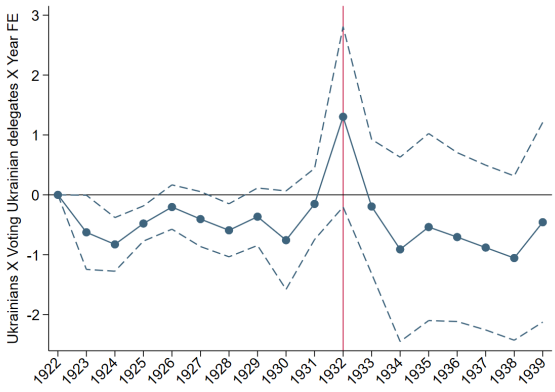
(c) Ukrainians × Rural Communists 1922, 27, 31 × Year FE



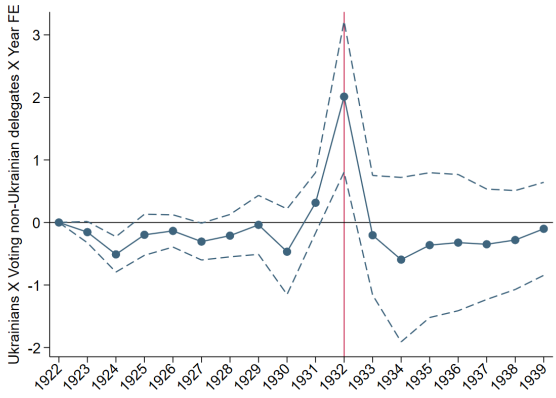
(d) Ukrainians × Pre-famine Collectivization × Year FE



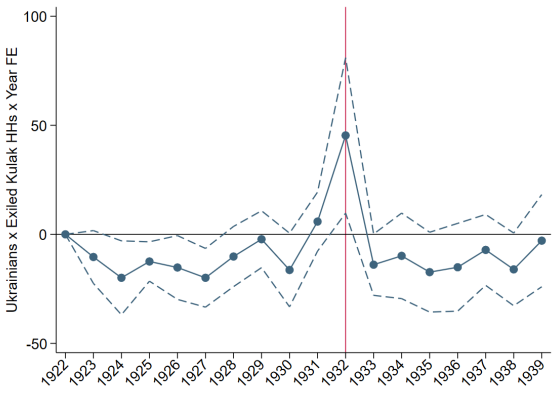
(e) Ukrainians × Delegates in 1930 Congress – ethnic Ukrainian × Year FE



(f) Ukrainians × Delegates in 1930 Congress – other ethnicities × Year FE



(g) Ukrainians × Resistance Principal Component × Year FE

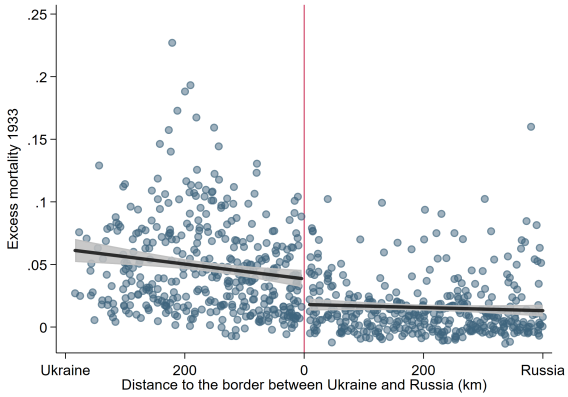


(h) Ukrainians × Pre-famine Exiled Kulak HHs × Year FE

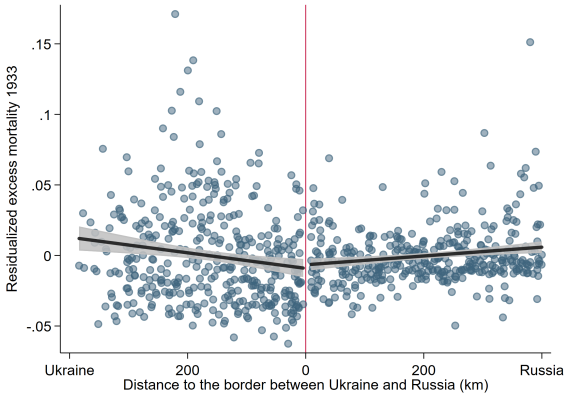
Notes: Figures (a), (b), (c), (d), (g) and (h) plot coefficients and their 95% confidence intervals estimated from separate regressions. Figures (e) and (f) are estimated from one regression. The dependent variable in each regression is mortality in year $t + 1$. The right hand side variables include the triple interaction stated in the sub-figure heading, all lower-order interaction terms, urbanization interacted with the rural share of ethnic Ukrainians and with year indicators and all lower-order interaction terms, predicted grain interacted with the rural share of ethnic Ukrainians and with year indicators and all lower-order interaction terms, and year and province fixed effects. 1922 is the omitted reference year. Standard errors are adjusted for spatial correlation within 1,500 km. *Source:* See the Data Appendix.

Figure 8: 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 Province-level Sample

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. Appendix Figure A.1a shows our provinces on the map (omitted territories are in white). See the Data Appendix for data sources.

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.3 shows the estimated grain production function. We then use this production function to predict grain harvest from 1922 to 1940. The predicted grain and actual grain are closely correlated; the two exceptions are Karelia and Eastern Siberia provinces, both are likely a result of errors in our matching procedure. In-sample R-squared is 0.90; out-of-sample R-squared is 0.77 (Appendix Figure A.3). This is consistent with the lack of major technological changes in the Soviet agriculture before the 1930s argued by historians (Allen, 2003).

C *Dekulakization*

Appendix Table A.4 shows that our main result is robust to different ways of controlling for the extent of *dekulakization* in the region. These measures are the number of exiled kulak households during 1930–31 according to Davies and Wheatcroft (2004, Table 28), the number of exiled kulak households during 1930–31 according to a secret police report in Berelowitch and Danilov (2000–2012, Document 253), *ex ante* 1930 quotas for kulak exile, secret police estimates of total number of kulaks in countryside, and the number of arrested peasants.

D District-Level Data

There is substantial variation in famine mortality across districts, even those within the same province. Appendix Figure A.1b shows a map of excess mortality in 1933 for each district where data are available. Appendix Figure A.5 presents the mean and normalized standard deviation in district-level mortality for Russia and Ukraine.[59] It shows 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 across districts within republics. Thus, there is inequality in famine intensity across smaller administrative units. Appendix Figure A.2 maps the shares of ethnic Russians and ethnic Ukrainians at the district-level before the famine in 1926.

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

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

	1926 census						1939 census	
	Total		Urban		Rural		Total	
	mil.	%	mil.	%	mil.	%	mil.	%
	A. USSR							
Russians	77.8	53.1	16.6	63.5	61.2	50.8	99.6	58.4
Ukrainians	**31.2**	**21.3**	**3.3**	**12.6**	**27.9**	**23.2**	**28.1**	**16.5**
Belorussians	4.7	3.2	0.5	1.9	4.2	3.5	5.3	3.1
Kazakhs	4.0	2.7	0.1	0.3	3.9	3.2	3.1	1.8
Uzbeks	3.9	2.7	0.7	2.8	3.2	2.6	4.8	2.8
Tatars	2.9	2.0	0.5	1.7	2.5	2.1	4.3	2.5
Jews	2.6	1.8	2.1	8.2	0.5	0.4	3.0	1.8
Georgians	1.8	1.2	0.3	1.2	1.5	1.3	2.2	1.3
Azerbaijanis	1.7	1.2	0.3	1.0	1.4	1.2	2.3	1.3
Armenians	1.6	1.1	0.6	2.1	1.0	0.8	2.2	1.3
	B. Regression Sample							
Russians	77.1	57.2	16.1	67.9	61.1	54.9	94.8	65.0
Ukrainians	**31.1**	**23.1**	**3.2**	**13.7**	**27.9**	**25.1**	**27.1**	**18.6**
Belorussians	4.7	3.5	0.5	2.0	4.2	3.8	5.2	3.6
Tatars	2.9	2.1	0.4	1.8	2.5	2.2	4.0	2.7
Jews	2.5	1.9	2.1	8.8	0.5	0.4	2.9	2.0
Mordvins	1.3	1.0	0.0	0.1	1.3	1.2	1.4	1.0
Germans	1.2	0.9	0.2	0.7	1.0	0.9	1.3	0.9
Chuvashs	1.1	0.8	0.0	0.1	1.1	1.0	1.4	0.9
Poles	0.8	0.6	0.2	1.0	0.5	0.5	0.6	0.4
Bashkirs	0.7	0.6	0.0	0.1	0.7	0.7	0.8	0.6
	C. "Grain-producing" Provinces							
Ukrainians	**28.5**	**45.5**	**3.0**	**29.9**	**25.4**	**48.5**	**25.2**	**38.8**
Russians	26.2	41.9	5.0	49.3	21.2	40.4	31.5	48.5
Tatars	1.7	2.8	0.2	1.8	1.6	3.0	2.2	3.3
Jews	1.7	2.7	1.3	13.1	0.4	0.7	1.7	2.6
Mordvins	1.0	1.6	0.0	0.1	1.0	1.9	0.9	1.4
Germans	1.0	1.6	0.1	1.0	0.9	1.7	1.0	1.6
Poles	0.5	0.8	0.1	1.0	0.4	0.7	0.4	0.6
Chuvashs	0.3	0.5	0.0	0.0	0.3	0.6	0.3	0.5
Moldovans	0.3	0.4	0.0	0.1	0.3	0.5	0.2	0.4
Armenians	0.2	0.3	0.1	0.9	0.1	0.1	0.2	0.3

Notes: These data are reported by the 1926 and 1939 Population Censuses. Panel C includes Central Black-Earth region, Crimea, Lower Volga, Middle Volga, North Caucasus, Tatar ASSR, and Ukraine.

Table A.2: Per Capita Food Production and Requirements for Ukraine

	1927	1928	1929	1930	**1931**	**1932**	1933	1937	1939
	A. Production and Procurement								
(1) Total population (mil.)	29.0	29.6	30.3	30.8	**31.3**	**31.7**	31.9	28.4	29.6
(2) Rural population (mil.)	23.6	24.6	24.9	25.1	**25.0**	**24.8**	25.0	18.8	18.7
(3) Production (mt)	18.6	13.9	18.7	22.7	**18.3**	**14.5**	22.0	22.5	23.8
(4) Procurement (mt)		1.9	5.3	7.7	**7.3**	**4.2**	6.1		
(5) Procurement rate	22.9%	13.6%	28.3%	33.8%	**39.5%**	**29.2%**	27.8%	.	.
	B. Implied Food Availability								
(6) Production pc (kg per year)	641	469	618	739	**587**	**457**	689	792	805
(7) Production pc (cal. per day)	5,506	4,023	5,307	6,342	**5,039**	**3,927**	5,919	6,797	6,912
(8) Rural retention pc (kg per year)	607	488	538	600	**444**	**415**	637	.	.
(9) Rural retention pc (cal. per day)	5,213	4,189	4,620	5,149	**3,815**	**3,559**	5,467	.	.
(10) *Cal. needs pc – heavy labor*	*2,455*	*2,462*	*2,459*	*2,455*	***2,446***	***2,437***	*2,437*	*2,374*	*2,357*
(11) *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 Lositskij (1920). 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 light work (urban) – 2,750 per day (Lositskij 1928), for relative caloric needs of other groups are based on Lositskij (1926). Caloric needs for avoiding mortality are based on the 900 caloreis per day for prime age adult males (Dasgupta and Ray, 1986).

Table A.3: The Effect of Weather and Natural Conditions on Grain Production

	Dependent Variable: Log Per Capita Grain Production		
	(1)		(2)
Log area	0.352***	Fall temperature × Fall precipitation	0.000*
	(0.067)		(0.000)
Log grain suitability	-4.643***	Winter temperature × Winter precipitation	0.001*
	(0.640)		(0.000)
Log area × Log grain suitability	0.278***	Spring temperature × Spring precipitation	0.000
	(0.023)		(0.000)
Fall temperature	0.015	Summer temperature × Summer precipitation	0.001***
	(0.037)		(0.000)
Winter temperature	0.027	Fall temperature^2	0.004**
	(0.043)		(0.002)
Spring temperature	-0.169**	Winter temperature^2	0.000
	(0.079)		(0.002)
Summer temperature	-0.978***	Spring temperature^2	-0.001
	(0.194)		(0.003)
Fall precipitation	-0.006	Summer temperature^2	0.028***
	(0.006)		(0.005)
Winter precipitation	-0.005	Fall precipitation^2	0.000
	(0.007)		(0.000)
Spring precipitation	0.010	Winter precipitation^2	0.000
	(0.007)		(0.000)
Summer precipitation	-0.025***	Spring precipitation^2	-0.000**
	(0.009)		(0.000)
		Summer precipitation^2	0.000**
			(0.000)
Observations	220		
R-squared	0.998		

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

Table A.4: The Effect of Ukrainian Population Share on Famine Mortality – Robustness to Controlling for *Dekulakization*

	Dependent Variable: Mortality in Year t+1								
	Baseline	Exiled kulaks (DW) × Famine	Exiled kulaks (OGPU) × Famine	Planned kulaks 1930 × Famine	Baseline with info on total kulaks	Total kulaks (OGPU estimate) × Famine	Baseline with info on arrested kulaks	Arrested kulaks 1930 × Famine	First principal component of Kulak variables × Famine
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Ukrainians × Famine	0.051***	0.054***	0.053***	0.044***	0.050***	0.058***	0.050***	0.051***	0.053***
	(0.006)	(0.005)	(0.006)	(0.006)	(0.008)	(0.007)	(0.006)	(0.006)	(0.005)
Observations	337	337	337	337	267	267	302	302	249
R-squared	0.785	0.8	0.797	0.815	0.756	0.788	0.794	0.795	0.808

Notes: Observations are at the province and year level. Exiled kulaks (DW) is the number of dekulakized and exiled households in 1930–31 per 1930 population according to Davies and Wheatcroft (2004): Table 28. Exiled kulaks (OGPU) is the number of dekulakized and exiled households in 1930–31 according to an OGPU report (Soviet Countryside from the Perspective of VChK-OGPU-NKVD, Document 253) per 1930 population. Planned kulaks are the planned number of "de-kulakizations" per capita in February, 1930 (the average between lower and upper bounds). Total kulaks (OGPU estimate) is the total number of kulaks in the rural population according to the OGPU estimate (Soviet Countryside from the Perspective of VChK-OGPU-NKVD, Document 253). Arrested kulaks is the number of peasants processed by "troiki" per capita in 1930 according to the OGPU estimate (Soviet Countryside from the Perspective of VChK-OGPU-NKVD, Document 279). All regressions control for urbanization, urbanization × famine, predicted grain, predicted grain × famine, and province and year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table A.5: The Effect of Ukrainian Population Share on Famine Mortality – Robustness to Controlling for Drop in Livestock

	Dependent Variable: Mortality in Year t+1					
	Baseline	Drop in horses × Famine	Drop in work horses × Famine	Drop in cattle × Famine	Drop in cows × Famine	Drop in all livestock (horses + cattle) × Famine
	(1)	(2)	(3)	(4)	(5)	(6)
Ukrainians × Famine	0.051***	0.051***	0.051***	0.052***	0.052***	0.053***
	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)
Observations	337	337	337	337	337	337
R-squared	0.785	0.785	0.785	0.8	0.786	0.794

Notes: Observations are at the province and year level. Drop in horses, work horses, cattle, cows, and all livestock is the difference between the 1929 level and the 1932 level per capita. Work horses is a subcategory of all horses that might be particularly important for production. Cattle includes oxen, bulls, cows, and calfs. Cows is a subcategory of cattle that might be particularly important during the famine by providing milk and dairy products. All livestock is horses plus cattle. Column 2 controls for the drop in horses interacted with the famine indicator. Column 3 controls for the drop in work horses interacted with the famine indicator. Column 4 controls for the drop in cattle interacted with the famine indicator. Column 5 controls for the drop in cows interacted with the famine indicator. Column 6 controls for the drop in all livestock interacted with the famine indicator. All estimates control for Urbanization, Urbanization × Famine, Predicted grain, Predicted grain × Famine, and province and year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table A.6: The Effect of Ukrainian Population Share on Famine Mortality – Robustness to Controlling for Wealth

	Dependent Variable: Mortality in Year t+1										
	Baseline	Nominal income 1897 × Famine	Real income 1897 × Famine	Labor productivity 1897 × Famine	Rural labor productivity 1897 (lower bound) × Famine	Rural labor productivity 1897 (upper bound) × Famine	Value of agricultural equipment 1910 × Famine	Horses 1916 × Famine	Cattle 1916 × Famine	Livestock 1916 × Famine	First principal component of wealth variables × Famine
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Ukrainians × Famine	0.051***	0.052***	0.051***	0.051***	0.050***	0.050***	0.046***	0.050***	0.051***	0.051***	0.050***
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)	(0.004)	(0.006)	(0.005)	(0.005)	(0.005)
Observations	337	337	337	337	337	337	337	337	337	337	337
R-squared	0.785	0.788	0.785	0.785	0.785	0.786	0.817	0.785	0.786	0.785	0.786

Notes: Observations are at the province and year level. All income proxies are measured in per capita terms. Column 2 controls for the nominal income 1897 interacted with the famine indicator. Column 3 controls for the real income 1897 interacted with the famine indicator. Column 4 controls for the labor productivity 1897 interacted with the famine indicator. Columns 5 and 6 control for the rural labor productivity 1897 (lower and upper bounds) interacted with the famine indicator. Column 7 controls for the value of agricultural equipment 1910 interacted with the famine indicator. Columns 8, 9 and 10 control for the horses, cattle, and total livestock 1916 interacted with the famine indicator. Column 11 controls for the first principal component of all wealth measures interacted with the famine indicator. All estimates control for Urbanization, Urbanization × Famine, Predicted grain, Predicted grain × Famine, and province and year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table A.7: The Effect of Ukrainian Population Share on Famine Mortality – Robustness to Controlling for Historical Institutions

	Dependent Variable: Mortality in Year t+1							
	Baseline	Share of catholics 1897 × Famine, Share of orthodox christians 1897 × Famine	Share of serfs 1858 × Famine	Peasant revolts 1895–1914 × Famine	Baseline with info on land 1905	Share of peasant land in repartition commune 1905 × Famine	Share of peasant households in repartition commune 1905 × Famine	Peasant and private land gini 1905 × Famine
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Ukrainians × Famine	0.051***	0.053***	0.052***	0.050***	0.043***	0.053***	0.057***	0.040***
	(0.006)	(0.006)	(0.005)	(0.006)	(0.005)	(0.007)	(0.008)	(0.006)
Observations	337	337	337	337	286	286	286	286
R-squared	0.785	0.791	0.787	0.805	0.796	0.803	0.805	0.803

Notes: Observations are at the province and year level. Column 2 controls 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. Column 3 controls for the share of serfs in the rural population in 1858 interacted with the famine indicator. Column 4 controls for the number of peasant revols from 1895 to 1914 per capita interacted with the famine indicator. Column 5 replicates 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). Column 6 controls for the share of peasant land in repartition commmune according to the 1905 Land Census interacted with the famine indicator. Column 7 controls for the share of peasant households that belonged to a repartition commune according to the 1905 Land Census interacted with the famine indicator. Column 8 controls for the gini coefficient for land ownership from the 1905 Land Census interacted with the famine indicator. All regressions control for urbanization, urbanization × famine, predicted grain, predicted grain × famine, and province and year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table A.8: Heterogeneous Effects of Peasant Resistance on Famine Mortality, Natality, Collectivization, Mechanization and Grain Procurement in Ethnic Ukrainian Areas

	Dependent Variables								
	A. Mortality in Year t+1			B. Natality in Year t+1			C. Collectivization		
Ukrainian × Famine	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
× Terrorist acts 1931–32	5.688***			-1.810***			27.327***		
	(0.698)			(0.170)			(4.470)		
× Mass demonstrations 1931–32		3.783***			-1.575***			-2.176	
		(1.468)			(0.351)			(10.944)	
× Leaflets 1931–32			6.724***			-2.390***			44.789***
			(0.804)			(0.383)			(11.150)
Observations	337	337	337	337	337	337	228	228	228
R-squared	0.846	0.833	0.851	0.838	0.838	0.839	0.968	0.969	0.968
	D. Tractors' Horse Power			E. Procurement Share			F. Procurement per Capita		
× Terrorist acts 1931–32	-5.933***			28.747***			7.722***		
	(0.378)			(6.612)			(1.032)		
× Mass demonstrations 1931–32		-2.777*			17.285***			3.931**	
		(1.487)			(5.722)			(1.653)	
× Leaflets 1931–32			-5.518***			28.805***			7.917***
			(0.815)			(7.257)			(1.008)
Observations	247	247	247	186	186	186	186	186	186
R-squared	0.790	0.787	0.790	0.878	0.876	0.879	0.811	0.81	0.811

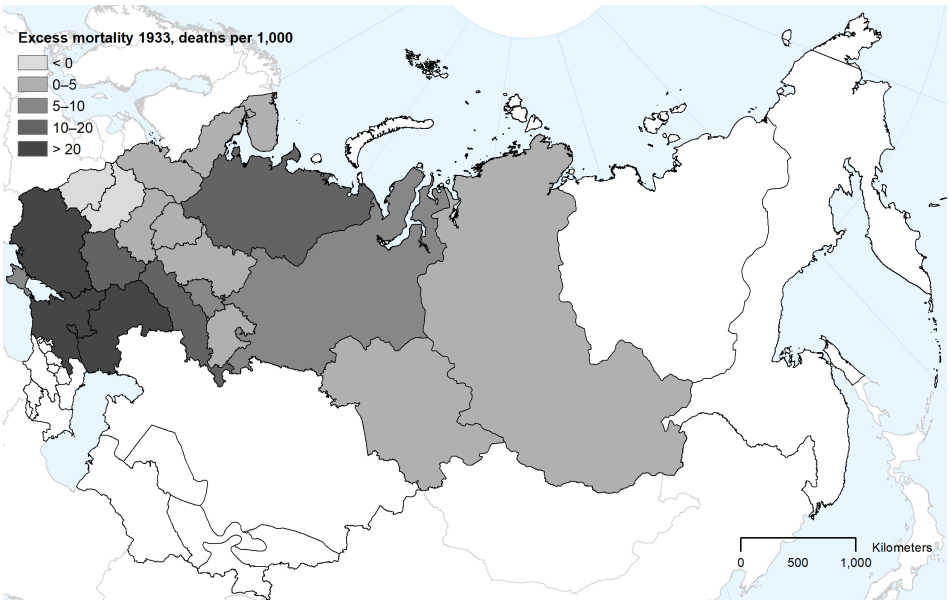
Notes: Observations are at the province and year level. The dependent variables are mortality in year t+1 (Panel A), natality in year t+1 (Panel B), the share of rural households in collective farms (Panel C), tractors' horse power per capita (Panel D), procurement/reported production (Panel E), and procurement/rural population (Panel F). The table presents the triple interaction of Ukrainian population share, the famine year dummy variable, and the variable stated in each row. The regressions control for all lower-order interaction terms; Ukrainians × Predicted grain × Famine, Ukrainians × Urbanization × Famine, and their lower-order interactions; and province and year fixed effects. Standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table A.9: Back-of-the-Envelope Calculation

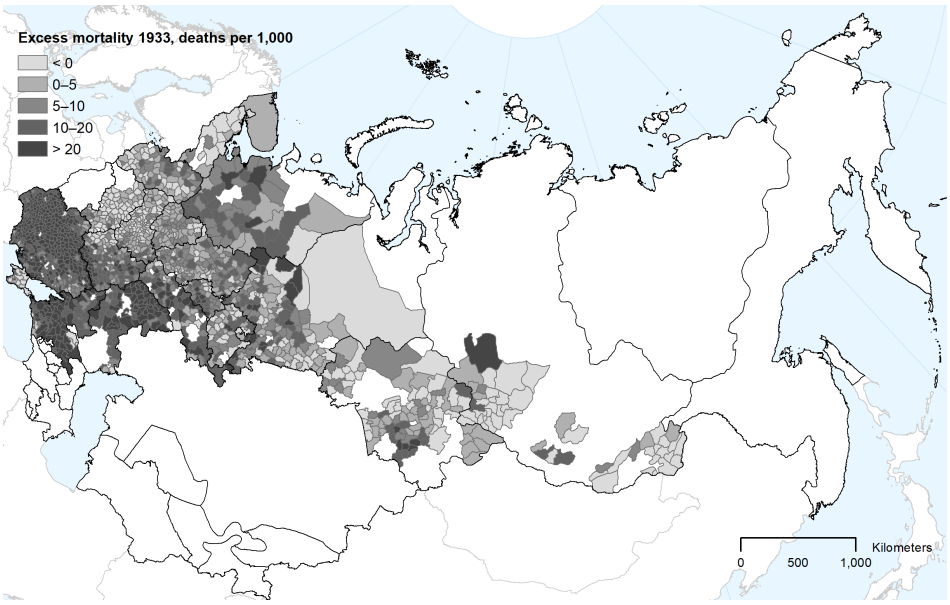
	Full Sample	Ukraine	Russia	Belarus
	A. Deaths			
(1) 1933 deaths if no famine (famine dummy = 0), mln	2.70	0.52	2.09	0.08
(2) Reported 1933 deaths, mln	4.81	1.86	2.88	0.07
(3) Predicted 1933 deaths (famine dummy = 1, Ukrainian = as reported), mln	4.97	2.03	2.84	0.10
(4) if no bias (famine dummy = 1, Ukrainian = 0), mln	3.22	0.64	2.48	0.10
(5) Total famine deaths: (3) - (1), mln	2.27	1.50	0.74	0.02
(6) if no bias: (4) - (1), mln	0.52	0.12	0.38	0.02
(7) Famine deaths due to bias: 1 - (6) / (5)	0.77	0.92	0.48	0.11
	B. Births			
(1) 1933 births if no famine (famine dummy = 0), mln	4.81	1.02	3.61	0.19
(2) Reported 1933 births, mln	3.25	0.45	2.68	0.12
(3) Predicted 1933 births (famine dummy = 1, Ukrainian = as reported), mln	3.22	0.38	2.70	0.14
(4) if no bias (famine dummy = 1, Ukrainian = 0), mln	3.71	0.76	2.80	0.14
(5) Total missing births: (1) - (3), mln	1.60	0.65	0.90	0.05
(6) if no bias: (1) - (4), mln	1.11	0.26	0.80	0.05
(7) Missing births due to bias: 1 - (6) / (5)	0.31	0.60	0.11	0.02

Figure A.1: Excess Mortality 1933

(a) Province Map



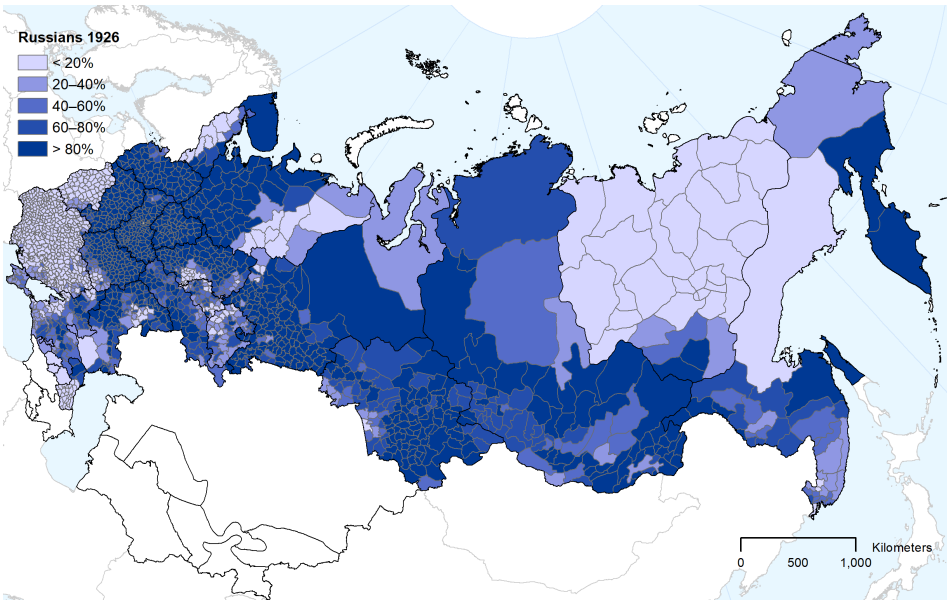
(b) District Map



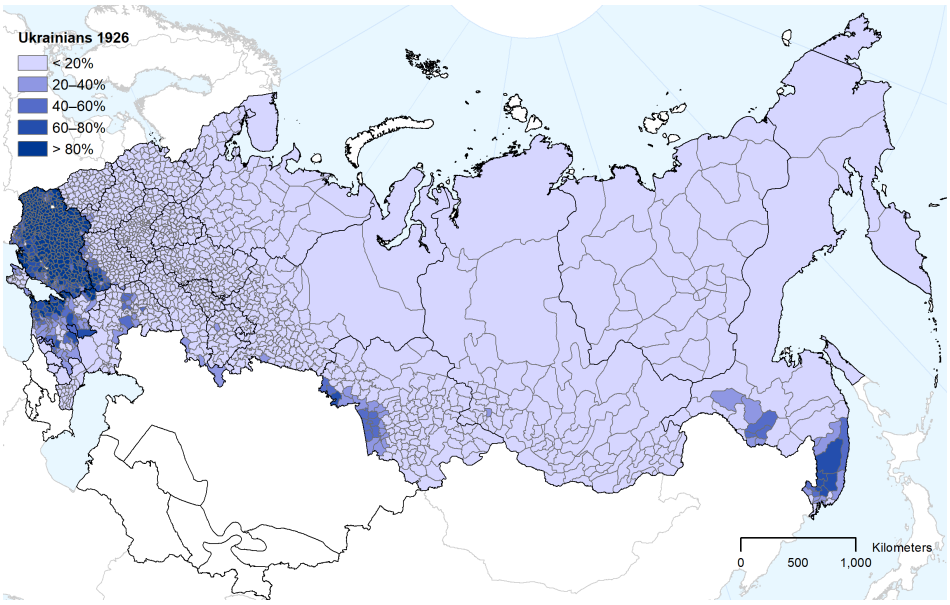
Notes: Excess mortality 1933 is mortality in 1933 minus mortality in 1928. *Source:* See the Data Appendix.

Figure A.2: Rural Ethnic Composition 1926

(a) Russians



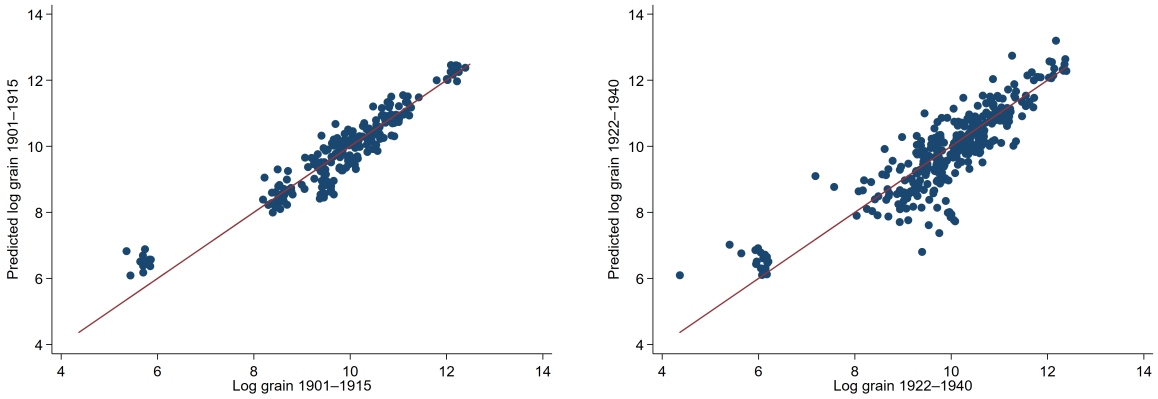
(b) Ukrainians



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

Figure A.3: Reported and Predicted Grain

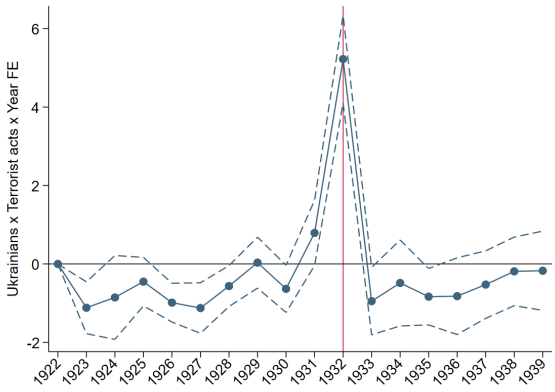
(a) In-Sample Fit, $R^2 = 0.90$ (b) Out-of-Sample Fit, $R^2 = 0.77$



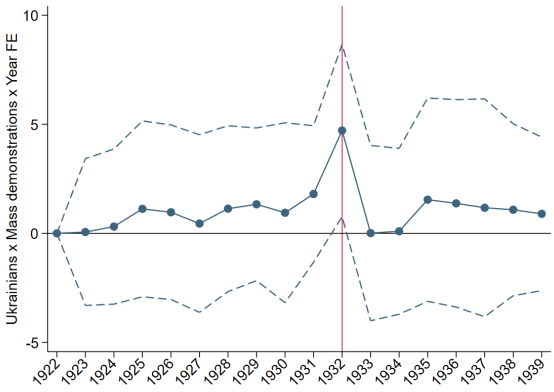
Notes: The figures show logs of reported and predicted grain with a 45-degree line; (a) for 1901–1915, a sample on which grain production function is estimated (in-sample fit), and (b) for 1922–1940 (out-of-sample fit); see Appendix section B for details., The Data Appendix presents the source of every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

Figure A.4: Heterogenous Effects of Ukrainian Population Share and Peasant Resistance on Mortality

(a) Ukrainians × Terrorist Acts 1931–32 × Year FE



(b) Ukrainians × Mass Demonstrations 1931–32 × Year FE

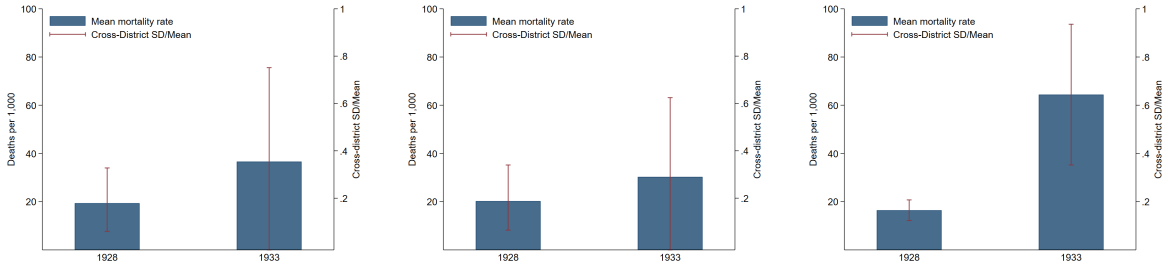


(c) Ukrainians × Leaflets 1931–32 × Year FE

Notes: The figures plot coefficients and their 95% confidence intervals estimated from separate regressions. 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 lower-order interaction terms, urbanization interacted with the rural share of ethnic Ukrainians and with year indicators and all lower-order interaction terms, predicted grain interacted with the rural share of ethnic Ukrainians and with year indicators and all lower-order interaction terms, and year and province fixed effects. 1922 is the omitted reference year. Standard errors are adjusted for spatial correlation within 1,500 km. *Source:* See the Data Appendix.

Figure A.5: 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-district 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.

Data Appendix (Not for Publication)

1932-33 Famine

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 S.S.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. Rural-urban decomposition of deaths and births is available since 1926.

Natality and Mortality

Natality is the number of live births divided by population (\crude birth rate). Mortality is the total number of deaths divided by population (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 1897 and hand-created district-level 1926 maps. The 1897 map is from Kessler, Gijs and Andrei Markevich, Electronic Repository of Russian Historical Statistics, 18th - 21st centuries, https://ristat.org/, Version I (2020).

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.

Dekulakization

The baseline measure of kulak households exiled during 1930--31 per 1930 population is estimated as the average between Exiled kulaks (DW) and Exiled kulaks (OGPU) defined below.

Exiled kulaks (DW) is the number of dekulakized and exiled households in Category II of kulaks in 1930--31 according to Davies and Wheatcroft (2004) (Table 28) per 1930 population.

Exiled kulaks (OGPU) is the number of dekulakized and exiled households of all categories between 01.01.1930 and 01.07.1931 according to an OGPU (secret police) 1931 report per 1930 population. The report is published in Berelovich A. and V. Danilov (2003). *"Sovetskaya derevnya glazami VChk-OGPU-NKVD. 1918—1939. Documents i materialy"* [Soviet Countryside from the Perspective of VChK-OGPU-NKVD]. Moscow: Rosspen. Vol. 3 "1930—1934 gg.", Book 1. "1930—1931 gg.", document 253.

Planned kulaks (lower bound) and Planned kulaks (upper bound) is the OGPU (secret police) planned number of dekulakizations by as of February, 1930 per 1930 population. The planned figures are published in Danilov, Victor, Robert Manning and Lynne Viola (Eds.). (1999-2006). *"Tragediya Sovetskoj Derevni. Kollektivizatsiya i raskulachivanie. Dokumenti i materialy v 5 tomakh, 1927-1939"* [Tragedy of the Soviet Countryside. Collectivization and Dekulakization. Documents and Materials. 5 volumes]. Moscow: Rosspen. Volume 2 "November 1929 — December 1930", Document 69.

Total kulaks (OGPU estimate) is the total number of kulaks in the rural population according to the OGPU (secret police) estimate published in Berelovich A. and V. Danilov (2003). *"Sovetskaya derevnya glazami VChk-OGPU-NKVD. 1918—1939. Documents i materialy"* [Soviet Countryside from the Perspective of VChK-OGPU-NKVD]. Moscow: Rosspen. Vol. 3 "1930—1934 gg.", Book 1. "1930—1931 gg.", document 253.

Arrested kulaks 1930 is the number of peasants processed by "troiki" in 1930 per 1930 population according to the OGPU (secret police) estimate published in Berelovich A. and V. Danilov (2003). *"Sovetskaya derevnya glazami VChk-OGPU-NKVD. 1918—1939. Documents i materialy"* [Soviet Countryside from the Perspective of VChK-OGPU-NKVD]. Moscow: Rosspen. Vol. 3 "1930—1934 gg.", Book 1. "1930—1931 gg.", document 279.

Peasant resistance to the Soviet regime

"Terrorist acts", unrest demonstrations, and anti-Soviet leaflets registered by the OGPU (secret police) between 01.01.1932 and 01.04.1932 per 1,000 1930 population are according to two OGPU reports. The reports are published in Berelovich A. and V. Danilov (2003). *"Sovetskaya derevnya glazami VChk-OGPU-NKVD. 1918—1939. Documents i materialy"* [Soviet Countryside from the Perspective of VChK-OGPU-NKVD]. Moscow: Rosspen. Vol. 3 "1930—1934 gg.", Book 1. "1930—1931 gg.", document 272, and Danilov, Victor, Robert Manning and Lynne Viola (Eds.). (1999-2006). *"Tragediya Sovetskoj Derevni. Kollektivizatsiya i raskulachivanie. Dokumenti i materialy v 5 tomakh, 1927-1939"* [Tragedy of the Soviet Countryside. Collectivization and Dekulakization. Documents and Materials. 5 volumes]. Moscow: Rosspen. Volume 3 "Late 1930 — 1933", Document 118.

Peasant resistance to the Tsarist regime

Peasant revolts in 1895—1914 are from Gokmen and Kofanov (2020).

Bolshevik votes 1917

Bolshevik vote share is from Protasov, V.V. Zhuravlev, and Shelokhaev (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 Kessler, Gijs and Andrei Markevich, Electronic Repository of Russian Historical Statistics, 18th - 21st centuries, https://ristat.org/, Version I (2020).

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 taken from Castañeda Dower and Markevich (2018), using manually constructed ArcGIS district (*uezd*)-level maps.

Pre-Soviet wealth measures

Nominal regional income per capita in 1897, real regional income per capita in 1897, regional labor productivity in 1897, regional rural labor productivity in 1897 (upper and lower estimates) are calculated from corresponding measures for imperial provinces, using hand-created ArcGIS district (*uezd*)-level maps. Imperial province estimates are from Markevich (2019).

We estimate the value of agricultural machines by multiplying the number of agricultural machines of different types by their prices and taking the sum. Agricultural machines data are originally from the 1910 census of agricultural machines. We calculate province shares from district (*uezd*)-level figures taken from Castañeda Dower and Markevich (2018), using hand-created ArcGIS district (*uezd*)-level maps. Prices are from Ministerstvo Zemledeliya [Ministry of Agriculture] (1917). *"Sbornik statistiko-ekonomicheskikh svedenij po sel'skomu khozyajstvu Rossii i inostrannikh gosudarstv. [A collection of statistical and economic information about agriculture in Russian and foerign countries]"*, Volume X.

Horses, cows, and livestock in 1916 are originally from the 1916 agricultural census. We calculate province shares from district (*uezd*)-level figures taken from Castañeda Dower and Markevich (2018), using hand-created ArcGIS district (*uezd*)-level maps.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 A.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.

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]"*.