The Causes of Ukrainian Famine Mortality, 1932-33*

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

May 6, 2023

We construct large, unique panel data to study the causes of Ukrainian famine mortality (*Holodomor*) during 1932-33 and document several new facts: i) Ukraine (the Soviet Union) produced enough food in 1932 to avoid famine in Ukraine (the Soviet Union); ii) mortality was increasing in the pre-famine ethnic Ukrainian population share and unrelated to food productivity across regions; iii) this pattern exists across the Soviet Union, even outside of Ukraine; iv) the pattern was similar at different administrative levels; v) migration restrictions exacerbated mortality; vi) actual and planned grain procurement were increasing, while actual and planned grain retention (production minus procurement) were decreasing in the ethnic Ukrainian population share across regions. Anti-Ukrainian bias in Soviet policy explains up to 92% of famine mortality in Ukraine and 77% in Ukraine, Russia and Belarus; approximately half of the total effect comes from bias in the centrally planned food procurement policy.

JEL: N4, P2

Keywords: Ethnic Repression, Mass Killing

*We thank Daron Acemoglu, Maxim Ananyev, Volha Charnysh, Michael Ellman, Ruben Enikolopov, Jeff Frieden, Scott Gehlbach, Mikhail Golosov, Sergei Guriev, Mark Harrison, Oleg Khlevniuk, Alexey Makarin, Victor Malein, Joel Mokyr, Cormac O'Grada, Elias Papaioannou, Steven Nafziger, Maria Petrova, Sakari Saaritsa, Torsten Santavirta, Marco Tabellini, Chris Udry and Ekaterina Zhuravskaya for their insights; the participants of the Harvard/MIT Development Workshop, Harvard Political Economy Workshop, NYU Development Workshop, LSE STICERD Applied Workshop, Northwestern Economic History Workshop, USC Development Workshop, Fudan University SOE Applied Workshop, Trinity College Applied Economics Workshop, CERGE-EI Applied Economics Workshop, EIEF Workshop, HECER (Helsinki Center of Economic Research) Labor and Public Economics seminar, Helsinki economic and social history seminar, Northwestern Economic History Lunch, Northwestern Development Lunch, George Mason Economic History Lunch, the NBER Political Economy Workshop, the Ridge Conference, the ASSA/AEA On the Role of and Culture and Community 2020 session, 2020 Summer Workshop in the Economic History and Historical Political Economy of Russia, Kiel/CEPR Conference on Geopolitics. For excellent research assistance, we thank Pavel Bacherikov, Nikita Karpov, Ivan Korolev, Sergio Lopez-Araiza Bejar, Kaman Lyu, Carlo Medici, Ludovica Mosillo. All mistakes are our own.

[†]University of Helsinki and New Economic School (Moscow, Russia), andrei.markevich@helsinki.fi, 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

Lazar Kaganovich "was fond of saying that every Ukrainian is potentially a nationalist" – from the memoirs of Nikita Khrushchev (Khrushchev, 1970).

1 Introduction

During the Great Soviet Famine (1932-33), approximately seven million people perished and forty percent of these deaths occurred in Ukraine, where mortality rates were four to six times higher than in Russia.[1] 7.5% to 11.3% of the ethnic Ukrainian peasantry died during the *Holodomor* ("to die by starvation" in Ukrainian). Levchuk et al. (2020) summarizes the controversy and intense debate about the causes of disproportionately high Ukrainian mortality rates. On the one hand, many historians believe that the famine was a "terror" intentionally waged by the Soviet government against Ukraine (e.g., Conquest, 1986). Approximately 40% of Soviet GDP was agriculture in 1928. To maximize revenues, the government forcibly procured agricultural production, especially grain, from rural areas. This was highly unpopular amongst the peasantry. The Ukrainians, who constituted the largest ethnic group in agriculturally productive regions, had a strong group identity and offered particularly stiff resistance. The Bolsheviks needed to repress Ukrainians to control agriculture (e.g., Graziosi, 2015).[2] The theory that there was inherent anti-Ukrainian bias in famine-era policy is supported by accounts of state coercion, aggressive food procurement and starvation in Ukraine (e.g., Applebaum, 2017).

On the other hand are those who argue that the high famine mortality in Ukraine was an unintended consequence of policies with no inherent anti-Ukrainian bias. Repression and starvation in agriculturally productive regions also occurred outside of Ukraine and communist ideology, in an intentional departure from the Tsarist regime, held all ethnicities to be equal. No direct documentary evidence that Stalin "ordered" a famine has been uncovered.[3] In their well-known study of the Soviet Famine, Davies and Wheatcroft (2009) document a fall in aggregate Soviet production in 1932 and use weather data from a few locations to argue that the production decline was due to exogenous factors. They reason that mortality was higher in Ukraine because it was agriculturally productive and the Soviets had no anti-Ukrainian bias *per se*.[4]

The main challenge for existing studies is the lack of representative disaggregated data, which are necessary for evaluating competing hypotheses. Most importantly, past studies have lacked systematic data on mortality, production and procurement. As such, the contentious debate on the causes of Ukrainian famine mortality reached an *impasse*.

The primary contribution of our study is to make progress on understanding the causes of Ukrainian

[1] See Section 2 for mortality estimates. Note that approximately 1 to 1.5 million famine deaths occurred in Kazakhstan, mostly among ethnic Kazakhs. We do not study Kazakh mortality because there are no reliable mortality data from Kazakhstan during the famine era.

[2] Conquest (1986), Ellman (2007), and Mace (2004) offer variations of this argument.

[3] Kondrashin (2008) points out that famine was severe in certain parts of Russia. Kotkin (2017) makes no mention of bias against Ukraine. argues that the Soviet famine was an unintended consequence, and notes the contrast between the lack of direct evidence on Stalin's intention to produce a famine for Ukrainians and the abundance of direct evidence for other killings during the Great Purge. Tauger (1991) argues that there was no policy against Ukrainians or other ethnic groups.

[4] Later, Wheatcroft (2012, 2018) argues that "... suggestions that the authorities had purposefully provoked these famines to attack any specific group does not make sense..." (Wheatcroft, 2012, p.341) and that "Applebaum (2017) [fails] to understand that there really was a shortage of grain" (Wheatcroft, 2018, p.468-9).

famine mortality in 1932-33 by constructing the largest and most comprehensive disaggregated dataset for inter-World War Soviet Union (1922-1940). Most of the data are manually digitized from archival sources made available after the fall of the USSR. The main sample used for the regression analysis includes the three largest and most populous Soviet republics: Russia, Ukraine and Belarus. We construct region-level panel data that include information about mortality, natality, ethnic composition, urbanization, weather, administrative capacity, realized grain production and procurement, planned targets for production and procurement, political alignment with Bolsheviks, and other historical economic, political and cultural variables. We are the first to digitize many of these variables and to manually harmonize changing administrative boundaries to construct a panel. Our famine mortality data are also the first to include republics other than Ukraine, which is useful for reasons that we discuss later.

We construct two panels. The first is a long province-level panel with a rich set of variables. The second is a district-level panel that is shorter and has fewer variables but nevertheless yields new insights since districts are much smaller administrative units than provinces. The data provide many advantages. First, having systematic data on pre-famine ethnic Ukrainian population share helps us overcome the lack of ethnic-specific mortality rates. We infer ethnic Ukrainian mortality from the relationship between famine mortality and the pre-famine ethnic Ukrainian population share across regions. This and the fact that our data include republics other than Ukraine allow us to investigate bias against ethnic Ukrainians throughout the Soviet Union. In contrast, past studies have mainly focused on the famine in only Ukraine, even though 25% of ethnic Ukrainians lived in other republics before the famine. Second, having multiple years of data for each administrative unit (the panel structure) allows us to accurately assess mortality due to the famine. Existing empirical analyses that we discuss later have only used one cross-section of mortality data. Since mortality rates vary across regions in normal times for reasons unrelated to famine, it is difficult to distinguish the excess mortality caused by famine from mortality due to other factors in one cross section. The panel structure also allows us to examine the dynamic effects and assess the credibility of contributing factors. For example, we would be more doubtful of the relevance of Ukrainian ethnicity for famine mortality if ethnicity and mortality are similarly correlated in famine and non-famine years. Third, the large sample size and rich set of variables allow us to simultaneously control for multiple factors and evaluate competing hypotheses. Particularly important for our study are data on grain production and procurement, which were centrally planned. We use previously classified data to correct the official production figures, which were exaggerated in the early 1930s. These variables allow us to examine whether there was anti-Ukrainian bias in government policy, which addresses the lack of documentary evidence for central directives ordering the famine. Existing empirical analyses have not had data on realized or planned production or procurement, as well as many other variables in our study.

In addition to the panel data used for the regression analysis, we also construct several other variables. We collect data on planned grain production and procurement targets, which helps us understand the government's intentions; and peasant resistance to collectivization from previously classified secret police reports, which sheds light on the degree of Ukrainian resistance to Soviet policy prior to the famine. Finally, we construct a panel of tractor allocation and use data from multiple population censuses to investigate how the key agricultural inputs, capital and labor, and ethnic composition changed after the famine.

Our analysis proceeds as follows. First, motivated by the existing discussion of the fall in agricultural production in 1932 as the main driver of famine, we ask whether the low production levels made famine inevitable for some population in the Soviet Union. We use the corrected production data to show that 1932 Ukrainian production, though lower than before, was still enough to support all of Ukraine, and production in the other republics was enough to support the rest of the Soviet population. Government procurement lowered food availability in rural Ukraine to a level that was similar to those in the famine-stricken regions of the 1921 famine. See Section 3.

Second, we investigate the contributions of anti-Ukrainian bias in Soviet policy versus a fall in agricultural production caused by factors with no bias (e.g., weather) to famine mortality. These two causes of famine are not mutually exclusive. Using the province-level panel, we regress mortality on pre-famine ethnic Ukrainian population share interacted with a famine-year dummy variable, and control for per capita grain production and its interaction with the famine dummy. The interaction coefficients reflect the relationship between famine mortality and ethnic Ukrainian population share, and the relationship between famine mortality and food production. If there was anti-Ukrainian bias, the Ukrainian share interaction effect will be positive. If production shortages caused famine mortality, then the production interaction effect will be negative. We also control for urban population share and its interaction with the famine dummy variable, which account for urban-rural differences such as food access; province fixed effects, which account for all time-invariant differences across regions (e.g., average mortality rate); and year fixed effects, which account for changes over time that affect all provinces similarly (e.g., macroeconomic changes).

Given the importance of understanding the relationship between grain production and famine, we take extra care in measuring production. As we discussed earlier, one way to address this is to use the corrected production estimates based on de-classified documents. However, these are only available for a limited number of years. Another way is to predict production with weather and geography. A third way is to directly control for weather and geography. All three methods produce similar results. Thus, the baseline estimates use predicted production, which is easier to interpret than individual weather controls and allows us to examine a longer time horizon than corrected production. We also show that the results are unchanged if we additionally control for pre-famine policies that may have lowered productivity (e.g., the forcible removal of successful farmers, *kulaks*; and the drop in livestock associated with collectivization).

The second result is that famine mortality was increasing in pre-famine ethnic Ukrainian population share, but uncorrelated with grain production. For two places that produced the same amount of grain in 1932, the place with more ethnic Ukrainians suffered higher mortality in 1933. This supports the anti-Ukrainian bias view. For two regions with the same ethnic Ukrainian population share, per capita grain production is unrelated to famine mortality. Thus, grain production and the inputs of grain production (e.g., weather and geography) cannot explain famine mortality. This contradicts the unintended consequence view. See Section 4.1

To ensure that the estimated Ukrainian bias in mortality is not spurious, we control for a large number of potential confounders, such as demographic structure, suitability for the cultivation of other crops (e.g., potatoes), political loyalty to the regime, administrative capacity and many historical economic and institutional variables. To help rule out that Ukrainian bias was due to slow-moving cultural or social features

that can affect famine intensity, we conduct a placebo test and show that there is no relationship between Ukrainian population share and famine mortality in the last large famine under the Tsarist regime in 1892. The ethnic Ukrainian-famine relationship is specific to the Soviet regime.[5] We show that our results are similar when using alternative measures of Ukrainian population share and famine severity, and conduct random permutation tests to show that they are unlikely to be driven by coincidence. See Section 4.3 for these and many other robustness exercises.

The third result is that the ethnic Ukrainian bias in famine mortality extended beyond the administrative borders of Ukraine. We re-estimate the mortality regression in restricted samples that omit Ukraine and find a similarly positive association between pre-famine ethnic Ukrainian population share and famine mortality. We also investigate whether there was bias in the famine mortality of other minorities by adding the interaction of the population share of other ethnic minorities and the famine dummy variable to the mortality regression. We find that other minorities in our sample died at similar rates as ethnic Russians.[6] Thus, Ukrainian bias is not a result of policies biased against all ethnic minorities.

To assess the magnitude of Ukrainian bias in famine mortality, we conduct a back-of-the-envelope calculation. Ukrainian bias explains up to 92% of total famine mortality in Ukraine and 77% in the three republics in our sample. Conceptually, these numbers reflect famine mortality had ethnic Ukrainians died at the same rate as other ethnicities (who are mostly Russians in our sample). The contribution of Ukrainian bias to famine mortality is higher in Ukraine because Ukraine has a higher share of ethnic Ukrainians.

Fourth, we ask whether the same relationship between Ukrainian population share and famine mortality exists across smaller administrative units within provinces. This is important for establishing that the Ukrainian bias inferred from the mortality regressions reflects state policy because Soviet economic policies were centrally planned and implemented top-down through the bureaucratic hierarchy and used similar formulas for assigning grain production and procurement targets at each administrative level. To investigate this, we use the district-level panel, which includes districts in Ukraine and Russia. We estimate a specification similar to the baseline, except that we add province-year fixed effects to exploit within-province variation (Ukraine is one province). The district-level patterns are similar to the province-level ones, which supports the interpretation that Ukrainian bias in famine mortality is policy driven. See Section 4.4.

Fifth, we investigate the relative importance of ethnic versus administrative boundaries. Existing studies of the famine that focus on Ukraine or compare Ukraine to other republics cannot distinguish bias against ethnic Ukrainians from bias against the republic of Ukraine since rural Ukraine, where most famine deaths occurred, is mostly ethnically Ukrainian. We address this with the district-level data. Consistent with survivor accounts, we document a discrete decline in raw famine mortality rates when crossing the border from Ukraine to Russia. We then show that the border effect disappears once we control for the ethnic Ukrainian population share in each district. Thus, administrative borders are irrelevant once ethnicity is taken into account; Ukrainian bias in famine mortality was delineated along ethnic lines. This result reinforces our earlier findings and show a clear pattern of systematic Ukrainian bias in famine mortality across the Soviet Union, province by province, district by district.

[5] We do not examine the 1921 Soviet famine because there are no disaggregated mortality data from the famine-stricken regions.

[6] Recall that our sample includes Russia, Ukraine and Belarus, but exclude Kazakhstan, which suffered high mortality rates during the famine.

The presence of the administrative border effect in the raw mortality data also sheds light on the role of migration restrictions that were put into place during the famine. If ethnic Ukrainians were allowed to freely migrate, mortality should not have discretely declined at the border. The border effect shows that migration restrictions exacerbated famine mortality by preventing Ukrainians from escaping to other regions.

The sixth question we ask is whether there is anti-Ukrainian bias in policy that affected food availability. Thus far, we have inferred anti-Ukrainian bias in Soviet policy by documenting Ukrainian bias in famine mortality, after controlling for other possible causes of famine. Here, we provide positive evidence by showing that there was anti-Ukrainian bias in the most important policy for determining food availability: centrally planned grain procurement. We estimate an equation similar to the mortality regression, except the dependent variable is grain procurement as a share of corrected production and we no longer control for production. This exercise uses the province-level panel, which contains procurement and production data. We find that procurement during the famine was increasing in pre-famine ethnic Ukrainian population shares. In other words, for two places that produced similar amounts of grain in 1932, more was procured from the place with a higher share of ethnic Ukrainians. Similarly, retention (production minus procurement) was decreasing in pre-famine ethnic Ukrainian population shares. These estimates provide positive evidence of anti-Ukrainian bias in the main government policy for determining food availability.

To quantify the contribution of anti-Ukrainian bias in food procurement on famine mortality, we estimate the relationship between food retention and mortality and combine it with the estimated Ukrainian bias in retention. We find that anti-Ukrainian bias in centrally planned procurement explains approximately half of the total effect of anti-Ukrainian bias on famine mortality. This is consistent with the importance of grain procurement in determining food availability, as well as the presence of anti-Ukrainian bias in other policies that affect famine mortality, such as mobility restrictions.[7]

The seventh question we ask is whether anti-Ukrainian bias in food procurement was centrally planned or an outcome of more decentralized decisions. To investigate this, we collect data for centrally planned grain production and procurement targets published in the First Five Year Plan in 1928, four years before the famine. We find that grain retention targets were decreasing in the share of ethnic Ukrainians across regions, which implies that anti-Ukrainian bias in grain consumption was centrally planned. However, the data do not necessarily imply that Ukrainian famine deaths were centrally planned because the production targets were very high and the retention differences implied by the target data would not have caused mortality at these levels.[8] The timing of the regime's decision to allow Ukrainians to suffer famine – e.g., 1928 or 1932 — depends on the regime's true expectations of 1932 production. See Section 4.7 for more discussion.

In addition to the main empirical findings on the causes of Ukrainian famine mortality, we provide several additional results to enrich our understanding. First, we examine the long-run patterns. We find that in the years just after the famine, areas with high pre-famine Ukrainian population share experienced a drop in grain production and rural population. Production recovered to pre-famine levels by 1936, while the rural population took much longer to recover its pre-famine size. Interestingly, we also observe an increase in the number of centrally allocated tractors just after the famine. These patterns together suggest that the regime

[7] See Section 2 for a discussion of other policies.

[8] The relationship between mortality and food consumption is strongly diminishing in consumption.

increased capital investment in famine-stricken areas to compensate for the negative impact of labor losses on production. We also provide suggestive evidence that areas with high pre-famine Ukrainian population experienced a permanent decline in the share of ethnic Ukrainian population after the famine. This goes against concerns of the "ecological fallacy" for our main mortality regressions. See Section 5.1.

Finally, we speculate about the regime's motivation and present some evidence consistent with the political economic explanation that the Bolsheviks repressed Ukrainians to control agriculture. We use declassified secret police reports on peasant resistance from the early 1930s to document that ethnic Ukrainians offered stronger resistance to Soviet agricultural policy than other groups (even outside of Ukraine); and that the steepness of the ethnic-Ukrainian-famine-mortality gradient is increasing in agricultural productivity across regions. These results are consistent with the Bolshevik's wariness of Ukrainian resistance and their objective to control agriculture. We discuss this and other complementary explanations (e.g., informational rigidities, misreporting of production) in Section 5.2. It is beyond the scope of our paper to be conclusive.

Our study is the first to systematically evaluate the causes of Ukrainian famine mortality. We add to several literatures. First, we add to studies of the causes of famine in the post-Industrial era discussed by Ó Gráda (2009). Sen (1981) famously argues that the central cause of 20th century famines is the unequal distribution of food from political elites to those who lack entitlement, and not low aggregate production. Our results strongly support this thesis by providing evidence on the detailed process that led to the second largest famine in the 20th century. Studying the Soviet context adds important insights. In market economies studied by Sen (1981), famine mortality is negatively associated with food production.[9] Our finding of no correlation between food production and famine mortality illustrates that different mechanisms are at play in the Soviet context. In this sense, we complement recent empirical analyses of the Chinese (e.g., Li and Yang, 2005; Meng et al., 2015) and Soviet famines (Davies and Wheatcroft, 2009; Naumenko, 2021). The political-economic mechanisms driving the Chinese and Soviet famines are different in several important ways. In the Chinese Famine (1959-61), rural political bases for the top leadership was hard hit by the famine and there was little ethnic delineation in mortality. [10] In their study of the Soviet Famine, Davies and Wheatcroft (2009) corrected the official production data at the aggregate Soviet level, but did not have disaggregated data for mortality, production or procurement. Their arguments rely on piecing together the aggregate data with documentary evidence and select variables from a few regions. They do not empirically evaluate their hypotheses or estimate regressions. By conducting a systematic analysis, we arrive at very different conclusions from this well-known work. Naumenko (2021) uses a cross section of districts in Ukraine to document a positive association between famine mortality and ethnic Ukrainians share. Her analysis faces several limitations. It is difficult to estimate famine mortality with one cross section of data for the reasons discussed earlier. The Ukraine-only analysis cannot meaningfully control for confounding factors such as agricultural productivity and urbanization, which is necessary for refuting alternative theories, because very few ethnic Russians live in rural Ukraine. The exclusive focus on Ukraine prevents any analysis of border

[9]For example, consider the West Bengal context studied by Sen (1981). Harvest failures caused agricultural wages to be low, such that those who did not produce food could not afford to buy surplus production from other regions. This means that famine was more severe in places that produced less food. For a more recent study of famine in colonial India, see Burgess and Donaldson (2017).

[10]The Bolshevik political base was urban workers, who suffered little famine. See Section 5.2 for more discussion.

effects.[11] This earlier work also lacks data on realized and planned production and procurement, and does not examine the mechanism underlying Ukrainian bias in mortality.

We are also related to studies of the long-run impact of early Soviet economic policy. In a study of attribution bias, Rozenas and Zhukov (2019) uses a cross-section of districts in Ukraine to document that places which had worse weather during the famine are more likely to blame Soviet policy for the famine and have stronger anti-Russian attitudes today. These results highlight the need for objective data since survivor accounts and collective memories can be biased, as well as the relevance of the famine for today. Cheremukhin et al. (2017) provides macro-calibration evidence on the effect of early Soviet Industrial policy, but exclude the cost of famine because of data limitations. The data we constructed, which we intend to make public, will facilitate future studies about this interesting and important context.

Finally, we add to the literature on ethnic conflict (e.g., Alesina and La Ferrara, 2005; Esteban et al., 2012; Montalvo and Reynal-Querol, 2005). In studying the Russian-Ukrainian relationship, we complement Korovkin and Makarin (2023), which finds that Ukrainian-Russian ethnic delineations affect firm behavior after the 2014 Russian annexation of Crimea. Recent studies of Eastern Europe also provide important evidence on the impact of group conflicts in contexts such as the abolition of serfdom (Buggle and Nafziger, 2019; Markevich and Zhuravskaya, 2018), forced migration (Bauer et al., 2013; Becker et al., 2020), peasant rebellions (Castañeda Dower et al., 2018; Finkel and Gehlbach, 2020), anti-Semitism (Grosfeld et al., 2013; Acemoglu et al., 2019) and ethnic diversity as a legacy of historical ethnic tensions (Egorov et al., 2020). Several works identify the causes of ethnic conflicts in Eastern Europe such as 19th century anti-Jewish pogroms (Grosfeld et al., 2020); the Croatian-Serbian conflict (DellaVigna et al., 2014) and ethnic divides in Nazi Germany (Adena et al., 2015).[12]

This paper is organized as follows. Section 2 summarizes the historical background. Section 3 presents the food accounting exercise. Section 4 presents the main mortality analysis. Section 5 presents additional results. Section 6 concludes.

2 Background

2.1 Ethnicity in the Soviet Union

In a departure from explicit ethnic discrimination during the Tsarist regime, Bolshevik ideology held all ethnicities to be equal. Yet, from the beginning of Soviet rule in 1917, the Bolsheviks needed to balance their reliance on the cooperation of nationalist groups with the concern that nationalist sentiments would undermine the regime. This was especially true for the Ukrainians, who were the second largest ethnic group after Russians in the Soviet Union and the largest ethnic group on agriculturally productive lands.[13]

[11] According to the 1926 Census, Only 5.6% of ethnic Russians in Ukraine live in rural areas and they are concentrated in a few districts. Thus, even though Naumenko (2021) controls for urbanization and agricultural suitability, there is insufficient support for these controls to be meaningful. In contrast, in our larger sample, 61% of rural residents are ethnic Russians and there is substantial variation across provinces and districts.

[12] There are numerous studies about the role of media on ethnic conflict in other contexts. See Petrova and Yanagizawa-Drott (2016) for a review.

[13] We present the descriptive statistics in Section 4.

The Ukrainians posed a dilemma for the Bolsheviks. On the one hand, the Bolsheviks needed the cooperation of the Ukrainians to rule. In the only free election in the Soviet era in 1917, the Bolsheviks took second place with 23% of votes, while the Ukrainian Socialist-Revolutionary Party took third place with 13% of the votes. The Bolsheviks also needed Ukrainian cooperation to achieve its economic goal of rapid industrialization, which required subsidies from agriculture, the most important source of Soviet economic production at the time. On the other hand, the Bolsheviks were aware of the danger of nationalist sentiments to the regime and the fact that the Soviet rural economic policy of "maximizing" grain extraction from rural areas would be highly unpopular for all peasants.

The strength of nationalist sentiments was apparent in the Civil War of 1918-1923.[14] To appease nationalists in 1923, the Bolsheviks launched a policy of *indigenization* (*korenizatsiya*). Indigenization encouraged schools and books in local languages, promoted native culture (e.g., national literature, theaters, museums), required local government affairs to be implemented in the native language, and promoted locals into leadership positions in the government and industry. The Ukrainian communist party was charged with the administration of Ukraine, but also viewed itself as representing ethnic Ukrainians across the Soviet Union.

The salience of ethnicity increased over time in rural areas. In the Russian empire, rural ethnic groups lived in separate communities. Ethnic delineations were inadvertently deepened during the Soviet regime, which established a hierarchy of national autonomous administrative units (republics, provinces, districts and villages) delineated along ethnic lines. A province with X% rural ethnic Ukrainians was typically a province where X% of villages were primarily ethnic Ukrainians. The residential patterns and administrative structure encouraged organization and coordination along ethnic lines, and also made it logistically easier for government to implement ethnic-specific policies.

The Bolsheviks leaders were wary of the increasing salience of ethnicity in rural areas and their concerns increased when peasants began to resist agricultural collectivization in the late 1920s and early 1930s. Resistance was particularly strong amongst Ukrainians.[15] Indigenization ended in Ukraine and other European parts of the USSR in the autumn of 1932 (Graziosi, 2015; Martin, 2001).

2.2 Soviet Economic Policy and the Famine

The Bolsheviks aimed to rapidly industrialize by subsidizing industry with agricultural surpluses. For this purpose, they aimed to maximize the expropriation of grain surpluses from the countryside. Their first attempt took place amidst the Civil War (1918-1923), which was fought mainly between the Bolshevik-led red army, the white army and separatists. *War communism* banned money and trading of foodstuffs and *prodrazverstka* aimed to extract all 'surplus' grain from peasants. The peasants resisted by not working. Sown area in 1921 was 30% lower than the 1913 level. The disruptions from the armed conflict and war communism contributed to the 1921 famine, which mostly occurred in Russia and small parts of Ukraine. Approximately five million perished. The scale of the resistance and the decline in agricultural production, which constituted over half of Soviet GDP in the early 1920s, threatened to undermine the Bolshevik

[14] The Civil War moderated in parts of the Soviet Union as early as 1920 and fully ended in all parts of the Soviet Union in 1923.

[15] As early as 1925, Stalin stated that "the national [ethnic] question [is], in essence, a peasant question" (Stalin, 15 April 1925 as quoted in Graziosi, 2015). We will return to discuss resistance to collectivization in Section 5.

economy. In 1921, Lenin retreated from War Communism and declared the New Economic Policy, which re-introduced a market economy for agriculture and small-scale manufacturing. The failure of *War communism* made clear to the Bolsheviks that they needed more control over agriculture.

In 1928, the Soviet Union was led by Stalin, who had substantially consolidated political power. The Bolsheviks, no longer distracted by armed conflict, renewed their effort to control agriculture. *Collectivization* was a bundle of policies that included the removal of private property and organizing peasants into large collective farms that the government could control directly. The government banned the trading of food and directly procured it from collective farms to feed the urban population and the small share of rural population engaged in non-agricultural production (e.g., forestry). Procured foods, which was mostly grain, were also exported and stored in centrally controlled reserves. In theory, peasants were meant to be left with enough food so that they could live and work. Production and procurement targets were set by central planners and published in 1928 in the First Five Year Plan. Production targets were mainly based on production potential, which depended on past production, geography and rural labor supply. Procurement targets were set to leave the rural population enough food to be productive laborers. Forced collectivization began in late 1929. By the summer of 1932, the share of rural households living in collective farms exceeded 60% in the USSR and was almost 70% in Ukraine (Davies and Wheatcroft, 2009).

The harvest of 1931 was lower than earlier years. News of starvation traveled to Moscow, but the government did not lower procurement. In areas where the peasants retained too little food, seed stock was consumed to make up for the deficit. This, in turn, contributed to lower production in 1932. It is widely believed that the production decline was particularly prominent in Ukraine, but the exaggerated official production data during the early 1930s made it hard to verify the magnitude of the fall in each region. We address this with the accounting exercise in Section 3.

Facing lower than expected harvests in 1932, Stalin simultaneously curtailed the initial targets and emphasized the need to maximize procurement.[16] To fulfill the remaining quota, Stalin sent his closest deputies, Vyacheslav Molotov and Lazar Kaganovich (neither of whom were Ukrainian) to Ukraine and North Caucasus, the two key grain-producing regions where most ethnic Ukrainians lived.[17] On December 14, 1932, the Politburo issued a classified decree in which Bolshevik leaders accused Ukrainian nationalists within the Communist Party and local bureaucracy of sabotaging grain procurement. The decree required regional authorities in Ukraine (as well as in the North Caucasus and the Western region) to "crush" any resistance of "counter-revolutionaries" and nationalists and fulfill procurement quotas (Danilov et al., eds, 1999-2006, Volume 3, Document 226).

[16] In a letter to his deputy, Lazar Kaganovich, from August 11, 1932, Stalin mentioned that the party district committees in about fifty districts in Ukraine had spoken out against state procurement quotas and that the Soviet government "could lose Ukraine" (Davies et al., eds, 2003). The evolving national procurement targets are reported by Davies and Wheatcroft (2009): p. 478, Table 20. Systematic data on the adjustments of the procurement targets are not available at the regional level.

[17] The central leadership claimed that shortages and famine were outcomes of intentional peasant resistance aimed to undermine agricultural collectivization, and that the peasants should be penalized for their subversion (Danilov et al., eds, 1999-2006; Davies and Wheatcroft, 2009). On May, 6 1933, during the peak famine mortality, Stalin wrote to Sholokhov, a famous writer originally from the Don region, that peasants "sabotaged" his policy and accused them of engaging in a "silent war" against the Soviet state (Murin, ed, 1997). Villages were penalized for failing to fulfill procurement targets with the seizure of all foodstuffs (not just grain), bans of imports of foodstuffs to the villages, arrests of local Communists (government bureaucrats), and the wholesale deportations peasants (Zelenin et al., 1994, p. 258, 260).

Deaths from starvation peaked in the early months of 1933. In January 1933, Moscow ordered the closure of the borders between Ukraine and the North Caucasus to prevent a mass migration of peasants out of these areas (Danilov et al., eds, 1999-2006, vol.3, p.634-5). The government sent some food aid to rural areas, but the amount was insufficient.[18] There were also other forms of famine relief, such as aid kitchens, medical assistance or housing. Most of these efforts were local, *ad hoc* and not systematically recorded. According to anecdotal accounts, rural famine victims often went to nearby urban areas to ask for food; some set up relief kitchens, while others expelled migrants back to rural areas. There are no systematic data about such relief efforts.

National mortality rates returned to trend in 1934, although some places took longer to recover. Total famine mortality estimates for the Soviet Union range from 5 to 10.8 million. Mortality was concentrated in rural areas.[19]

There are no systematic data on ethnic-specific mortality rates in the Soviet Union for the period of our study. One way to approximate ethnic Ukrainian famine mortality is to use the most cited total famine death toll of seven million for the USSR (Conquest, 1986), and 2.6 million (Meslé et al., 2013) to 3.9 million (Rudnytskyi et al., 2015) for Ukraine. If famine deaths were equally distributed between ethnic Ukrainians (80% of Ukraine) and others ethnicities in Ukraine, and no ethnic Ukrainians died outside Ukraine, then ethnic Ukrainian deaths constitute 30% ($.8 \times 2.6/7 = .3$) to 45% ($.8 \times 3.9/7 = .45$) of the total famine deaths. Thus, ethnic Ukrainians, who were 21% of total Soviet population in 1926 constituted about 40% of all famine deaths.[20] Famine mortality rates in Ukraine were four to six times higher than in Russia.[21]

In summary, it is clear that Ukrainians suffered higher mortality rates during the famine than Russians, but we do not know if this is simply because Ukrainians lived more in agricultural regions which suffered a fall in production. Nor do we know if the Ukrainians would have suffered a famine absent government procurement simply because they produced too little food. The historical evidence indicates that the Bolsheviks were concerned about Ukrainian nationalism, but does not reveal whether they intended to procure more food from the Ukrainians above and beyond other peasants in similarly productive lands. The histor-

[18]The best available data show that 176,000 tons of grain were sent to Ukraine in the spring and summer of 1933 (Davies and Wheatcroft, 2009, Table 23). This was only around 2% of 1932 Ukrainian production according to the estimates and 8% of the difference between what they retained after procurement and what was needed to survive (see Table 1). There is no systematic data for the most of other regions.

[19]Conquest (1986) estimates total famine deaths to be 7 million. Davies and Wheatcroft (2009) 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." 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). The differences in estimates are driven by data limitations, and potential underregistration of deaths during the famine. Estimates deriving excess famine deaths from a comparison of the pre-1926 and post-1937 Soviet population censuses suffer from the problem of the underregistration of infant mortality (See Davies and Wheatcroft, 2009, for a detailed discussion).

[20]These estimates likely underestimate ethnic Ukrainian deaths because ethnic Ukrainian mortality was higher than those for other groups within Ukraine, and many Ukrainians who lived in other republics also died.

[21]If total famine deaths is seven million, and we subtract the deaths in Kazakhstan (1 to 1.5 million) and Ukraine (2.6 to 3.9 million), we are left with approximately 1.6 to 3.4 million deaths for Russia (assuming no famine mortality in other republics). This implies famine mortality rates of 14 to 30 per 1,000 for the 112 million residents of Russia. A similar calculation for Ukraine, yields a famine mortality rate of 81 to 122 per 1,000, which is around four to six times larger than the implied famine mortality rate in Russia.

ical narratives focus on the tension between Moscow and Ukraine, but does not mention ethnic Ukrainians living elsewhere in the Soviet Union, who the Ukrainian Communist Party proclaimed to represent. The empirical analysis will address these and other questions about the famine.

3 Food Accounting

This section conducts a simple republic-level accounting exercise and documents that 1932 grain production in Ukraine was sufficient for avoiding famine in Ukraine and production in the rest of the Soviet Union was sufficient for avoiding famine in those republics had there been no procurement from Ukraine. Famine in Ukraine was not necessary for averting famine elsewhere and the key to understanding the causes of famine mortality is the distribution of the 1932 production. We describe the results below.

Table 1 presents grain production, procurement, retention (production minus procurement) and compare retention to rural population food needs for each year from 1927 to 1939. These data cover the entire Soviet Union. Panel I examines Ukraine. Rows (1) and (2) report total and rural populations measured at the beginning of each year. Thus, the population of 1933 does not account for the bulk of famine mortality which occurred in 1933.[22]

Row (3) reports production. The main challenge for this exercise is that the aggregate production data are widely believed to have been exaggerated during the early years of collectivization. Aggregate grain production was publicized and used as a marker of the success of Soviet economic policies. Davies and Wheatcroft (2009) provides a range for true aggregate production for the Soviet Union. We follow the spirit of this earlier study and use previously classified data to construct corrected province-level production, which we can then aggregate to the republic and aggregate levels. Specifically, we use grain procurement ratios (grain procurement stocks as a share of production) and procurement stocks to back out true production. Procurement stocks were directly observed and counted by the government. Procurement ratios were reported in Gosplan, a candid evaluation of the First Five-Year Plan shown only to the highest ranking Soviet officials. The report was classified until after the fall of the Soviet Union. The data for procurement ratios and procurement levels are widely accepted as accurate. Our correction assumes that changes in the procurement ratio reflects changes in production rather than changes in the ability to procure. Reassuringly, our corrected production measures are lower than officially reported production during the years when production is believed to have been exaggerated. At the Soviet Union level, our measures are comparable to the production estimates provided by Davies and Wheatcroft (2009). See the Appendix for more details.

Row (4) presents grain procurement, which includes urban consumption, exports and national reserves. Production and procurement are reported in millions of tons.

Row (5) reports actual rural retention, the difference between production and procurement, which includes seed stock intended for cultivation (since peasants consume seed stock during times of starvation). We report retention in kilograms per capita per day, which is more intuitive for considering subsistence

[22]The population data are reported by different official sources. The missing years are due to the lack of population data for those years. See the Data Appendix. Later, for the regressions, we address these issues by examining mortality directly and interpolating population for the missing years.

needs.[23] Row (6) reports a counterfactual retention for rural Ukrainians which assumes that no grain is procured from rural Ukraine. Row (7) reports a counterfactual retention for urban and rural Ukrainians which assumes that grain produced in Ukraine is distributed only to those living in Ukraine. This is simply production divided by the total population, converted into kilograms per capita per day units.

To understand whether retention is sufficient for avoiding famine, we calculate population food needs (row 8). We conservatively use official Soviet guidelines for maximum caloric needs for each gender and age group (e.g., heavy labor for prime-age adult males) as reported by Lositskij (1926; 1928). The guidelines vary by sex, age and occupation. Thus, we use data on sex, age and urban-rural shares as reported by the 1926 Census to adjust subsistence needs according to the demographic composition of each year.[24]

The data show that grain production in Ukraine declined from 23.2 to 16.8 million tons from 1930 to 1931, and further declined to 9.1 million tons in 1932 (row 3). Grain procurement from Ukraine remained stable at 7.7 and 7.3 million tons in 1930 and 1931; and decreased to 4.2 million tons in 1932 (row 4). The decline in procurement is consistent with the earlier historical discussion that the regime was unable to procure initial targets in the face of the production drop in 1932. Grain procurement declined with the drop in production in 1932, but not enough to prevent famine. Per capita grain retention in rural Ukraine declined from 1 kg per capita per day after the harvest of 1931 to only 0.5 kg per capita per day after the harvest of 1932 (row 5). This was 30% below the estimated food needs of 0.778 kg per capita per day (row 8) and similar to 0.582 kg per capita per day retention level in the famine-stricken areas of the 1921 famine.[25]

The famine in rural Ukraine could have been avoided, or at least, greatly moderated if there was no procurement from rural Ukraine. If no food had been procured from rural Ukrainians, rural retention would have been 1.01 kg per capita per day in 1932 (row 6), much more than the recommended level of 0.778 kg per capita per day. Moreover, Ukraine produced enough food to feed all Ukrainians. If grain was procured from rural Ukraine only to feed urban Ukrainians, average Ukrainian retention would have been 0.79 kg per capita per day (row 7).

Procurement from Ukraine was not necessary to avoid famine elsewhere. Panel II examines all other republics of the USSR. Row (11) shows that urban and rural grain retention would have been 1 kg per capita per day in 1932 if the production of other republics was distributed equally across their population.[26]

There are several important points to keep in mind for interpreting the accounting results. First, we focus on grain because it is the most important agricultural commodity for the Soviet economy and the main staple for consumption. We do not have systematic disaggregated data for the production of other crops. Later, in the regression estimates, we account for this by controlling for the suitability for cultivating other staple and cash crops. The availability of other food would strengthen our point that food production was sufficient for avoiding famine and that over-procurement is a necessary condition for high famine mortality. Second,

[23] To obtain this measure, we multiply the difference between production and procurement (rural retention in millions of tons) by 1,000 to convert the unit of measurement into kilograms, and divide by the rural population size and 365.

[24] The Soviet guidelines are more generous than international standards. For example, for a prime-age adult male, it is assumed that 3,750 calories are needed for heavy labor. We convert calories to kilograms of grain using the conversion offered by Lositskij (1920), which is based on the typical grain consumption of Russian laborers.

[25] During the 1921-22 famine, per capita retention was 0.582 kg per day in Russia and Ukraine, where the famine was most intense. Data are not available for the whole Soviet Union for these years. Approximately 5 million people died in this earlier famine.

[26] See the Appendix Table A.2 for food accounting for Russia and the entire Soviet Union.

our estimates do not take into account losses during transportation or storage because of data limitations. Lositskij (1920) estimates such losses to be around 5%. Our threshold for food needs is very generous since the amount for heavy labor is much more than the amount needed to avoid mortality. Thus, taking food losses into account should not overturn the result that famine could have been largely avoided if no food were taken out of Ukraine. Third, we do not account for food that the central government distributes back to regions after procurement. These data are not systematically available and they are unlikely to affect our results because the transfers were mostly sent to urban areas and rural workers not involved in grain production (e.g., forestry workers). The accounting exercise also excludes *ad hoc* famine relief efforts that we discussed in the Background Section because of data limitations. The small quantities of such relief mean that they are unlikely to overturn our result.

Finally, the estimates in this section do not take into account the fact that some of the population living in Ukraine are not ethnically Ukrainian, or that many ethnic Ukrainians lived outside Ukraine. We address this in the next section by inferring ethnic-specific mortality rates from the regression estimates .

4 The Causes of Ukrainian Famine Mortality

This section uses province and district-level panel data to investigate the causes of Ukrainian famine mortality.

4.1 Ukrainian Bias, Production Fall and Other Factors

Motivated by the current debate on the causes of famine in Ukraine, we begin by examining the contribution of Ukrainian bias and the level of food production. Our main measure of famine severity is mortality (deaths per 1,000). Province-level mortality data are available for each year in our sample for nineteen provinces from the three most populous Soviet republics: Belarus, Russia and Ukraine. Unless otherwise stated, the subsequent analysis is restricted to these three republics, which include 84% of the 1926 Soviet population and 88% of the 1928 Soviet grain production. For brevity, we refer to these three republics as the Soviet Union or the USSR in the remainder of the paper. The average province in our sample has 6.5 million people in 1926. All data are mapped to their 1932 province borders.[27]

Figure 1a plots mortality from 1923 to 1940 for Ukraine and the other two republics. It shows that during non-famine years, mortality rates are lower in Ukraine (18 per 1,000) than elsewhere (22 per 1,000) and stable over time. However, mortality sharply increases during the famine. In Belarus and Russia, mortality increases in 1933 to approximately 30 per 1,000. In Ukraine, the increase begins slightly earlier in 1932 to approximately 22 per 1,000 and then spikes in 1933 to approximately 60 per 1,000. The earlier timing and larger magnitude of the total mortality increase in Ukraine are consistent with historical accounts that starvation began earlier in Ukraine and the greater intensity of the famine peak in 1933. There is significant variation in famine mortality across and within provinces.[28]

[27] See the Data Appendix for details.

[28] Appendix Figure A.1a plots mean mortality and the cross-province normalized standard deviation in mortality for the full sample. It shows that the sharp rise in average famine morality coincides with a sharp rise in the variance. Appendix Figure A.2c maps excess mortality at the district level. Appendix Figure A.2e maps the excess mortality after demeaning by province fixed

The data on ethnic Ukrainian population share is reported by the 1926 Population Census (the last census before the famine). In the 1926 Census, 23.2 million ethnic Ukrainians lived in Ukraine and an additional 7.9 million lived in Russia and Belarus. Ukrainians constituted 21% of total Soviet population and made up the second largest ethnic group. Russians, the largest group, constituted 53% of the population. 89% of ethnic Ukrainians lived in rural areas. In regions that produced large food surpluses that the government designated as "grain-surplus" areas, Ukrainians were the largest group (43.8%) and Russians were a close second (41.9%).[29] Ethnicity is self-reported in the census and these data are widely accepted to be accurate. There was little reason to intentionally misreport. Later, we show that our results are similar when using alternative measures of ethnicity.

In the 1926 Census, 23.2 million Ukrainians lived in Ukraine and 7.9 million lived in the other two Soviet republics of our sample.[30] Since most non-Ukrainians in our sample are ethnic Russians, we sometime refer to the reference group in the regressions as "Russians".

The following equation characterizes the relationship between famine mortality, pre-famine ethnic Ukrainian population share and grain production.

$$mortality_{i,t+1} = \alpha + \beta Ukrainians_i \times Famine_t + \gamma prod_{i,t} \times Famine_t + \Gamma X_{it} + \eta_i + \delta_t + \varepsilon_{it}. \quad (1)$$

Mortality in province i during year $t+1$ is a function of: the interaction of the rural ethnic Ukrainian population share in province i in 1926, $Ukrainians_i$, and a dummy variable that equals one in the famine year, $Famine_t$; the interaction of per capita grain production, $prod_{i,t}$, and the famine dummy; and a vector of additional controls, X_{it}, which include the uninteracted production variable, urban population share and its interaction with the famine year dummy; province fixed effects η_i; and year fixed effects δ_t. Since $Ukrainians_i$ is a time-invariant measure, the uninteracted term is absorbed by the province fixed effects. We define the famine dummy, $Famine_t$, to equal one in 1932 because 1933 was the year with the highest mortality rates and we assume that grain production in year t is used to sustain the population in year $t+1$. Our results are qualitatively similar if the famine variable takes the value of one for 1931 and 1932. This is clear when we present the dynamic estimates later in the paper.

Given the nature of food production and famine, we allow the correlation of the standard errors to gradually decay across space to account for the movement of people, information and other factors across space. The province-level results are similar if we cluster the standard errors at the province level.[31] Later,

effects.

[29] The 1926 Population Census is commonly viewed as one of the highest quality Soviet censuses (Andreev et al., 1998). It is the last census before agricultural collectivization. After Ukrainians, the next largest ethnic group was an order of a magnitude smaller: Belorussians were 3.2% of total Soviet Union population and 3.5% of our sample. Appendix Table A.1 Panel A lists the three largest ethnic groups in the entire Soviet Union, Panel B lists the three largest ethnic groups in our sample, and Panel C lists the three largest ethnic groups in the grain-surplus provinces of our sample.

[30] Appendix Figure A.2b maps the share of ethnic Ukrainians in the rural population for each province as reported in the 1926 Census. Grain-surplus regions are shaded in crosses. Appendix Figure A.2d maps Ukrainians at the district level. Appendix Figure A.2f maps Ukrainians at the district level after demeaning by province fixed effects. It shows that there is significant variation within provinces.

[31] We follow the recommendations by Colella et al. (2019) in adjusting for spatial correlation within 1,500 kilometers (the mean province width in our sample is 1,300 km. Appendix Table A.4 shows that standard errors of the estimate of β are very similar for different levels of spatial and time correlation. If we cluster at the province level and estimate wild bootstrapped standard errors to account for the small number of clusters (Cameron et al., 2008), the p-value is 0.030. We do not focus on the clustered standard

when we use geographically smaller districts as units of observation, we estimate both standard errors that are spatially corrected and clustered at the district level.

β is the difference in the correlation of Ukrainian share and mortality between famine and non-famine years. If ethnic Ukrainians died at higher rates during the famine, then $\beta > 0$. γ captures the relationship between per capita grain production and famine mortality rates. If famine mortality was caused by low grain production, then $\gamma < 0$. We focus on the interaction coefficient, γ, because the relationship between production and mortality is likely to differ between famine and normal years, when production levels were higher. This is because the positive relationship between food consumption and survival exhibits strongly diminishing returns.

To account for urban-rural differences in food policies and other factors that affect famine mortality, X_{it} includes urban population share in province i during year t, and its interaction with $Famine_t$. We show that the results are similar if we control for 1926 urban share (or lag urban share) interacted with the famine dummy.

Table 2 presents the results. We measure grain production in several ways to address the concern that official figures from the early 1930s are exaggerated. In column (1), we use the corrected estimates based on previously classified documents as in the accounting exercise. However, the raw data we use for the correction are only available for a few years (1928 to 1933) and the correction is subject to assumptions discussed in Section 3. Alternatively, column (2) presents estimates that control for production predicted by weather and geography. These data are available for a longer time horizon and allows us to avoid the assumptions needed for the corrected production measure. Predicted grain can be interpreted as a parsimonious way of controlling for weather and geography.[32] The main concern for predicting production with weather is the possibility that the returns to inputs changed in the early Soviet era. We address this after we present the baseline estimates. Later, we also directly control for weather and the results are very similar.

In both columns (1) and (2), the interaction coefficient of Ukrainian population share and the famine dummy variable is positive and statistically significant at the 1% level. For two provinces with the same per capita production in 1932 (and urbanization), the one with a higher share ethnic Ukrainian population prior to the famine experienced higher famine mortality. This is consistent with anti-Ukrainian bias in famine-era policies contributing to famine mortality. In contrast, the interaction coefficient for grain and the famine dummy variable is positive. Because the positive interaction coefficient varies in precision and magnitude with the two measures of production (and the slightly different samples), we will cautiously interpret it as a null result. These estimates imply that for two places with the same ethnic Ukrainian population share (and urbanization), famine mortality was uncorrelated to production in 1932. This is inconsistent with the drop in production (and therefore, the drivers of production, such as earlier policies or weather) contributing to famine mortality. To maximize sample size, we henceforth focus our discussion and the remaining mortality regressions in this section on the estimates using predicted grain.

errors given the recent work by Canay et al. (2021), which argues that asymptotic validity for Cameron et al. (2008) requires strong assumptions of strict homogeneity between clusters.

[32] We use pre-Soviet province-level grain production, monthly temperature and precipitation and data for other inputs (e.g., log province area) to estimate the relationship between weather and grain for the years 1901 to 1915. We then use data for these variables from 1922 to 1940 to predict grain output. See the Appendix Section B for details.

The implied magnitude of Ukrainian bias in famine mortality is large. Taken literally, column (2) implies that for two provinces that experienced the same weather, and which have the same degree of urbanization, famine mortality rate was higher by 51 per 1,000 in a province with 100% Ukrainians than in a province with no Ukrainians. Mean mortality rates are 21 per 1,000 during non-famine years and 31 per 1,000 during the famine. Another way to assess the magnitude is to examine the standardized coefficient, which is presented in italics. During the famine, increasing Ukrainian population share by one standard deviation would result in a 0.826 standard deviation increase in mortality relative to normal years.

Conceptually, controlling for predicted grain production controls for production inputs such as weather and the geographic and climatic suitability for cultivating grain. However, if the relationship between these factors and production changes between the pre-Soviet era and the Soviet era, then the predicted estimates may be misleading. In our context, one may be concerned that earlier Soviet policies reduced the returns to production inputs so that predicted production overstates true production in Ukrainian regions.

We address this in two ways. First, we control for predicted grain as well as the policies that were most likely to have reduced agricultural productivity. In columns (3) and (4), we examine the sensitivity of the ethnic Ukrainian interaction coefficient to controlling for the two main policies that would have reduced the returns to normal agricultural inputs: *dekulakization* and the loss in livestock that occurred just prior to the famine.

In the *dekulakization* campaign, approximately two million peasants (*kulaks*) were exiled to Siberia and other remote regions for actively resisting collectivization (Viola, 2007). *Kulaks* were often the relatively more productive peasants and their removal could have reduced the returns to the inputs we use to predict production. Between 1929 and 1932, the number of horses declined by 42% and cattle by 40% (Viola, 1996, p. 70). When peasants lost the property rights to their livestock, they responded by slaughtering, eating or simply neglecting the newly collectivized animals. Livestock was the main source of horsepower and manure was an important input for crop cultivation. The loss of livestock could have reduced productivity. Also, the depletion of livestock meant that the traditional means of avoiding famine — slaughtering and eating the animals — were unavailable to Soviet peasants.

We control for the number of *kulak* households exiled from each region in 1930-31 divided by the 1930 population (column 3) and the drop in per capita livestock between 1929 and July, 1931 (column 4). Since these variables are time invariant, we control for their interactions with the famine indicator. The interactions of Ukrainian population share and the famine dummy variable are similar to the baseline. The estimates are similarly robust when controlling for alternative measures of the Soviet *dekulakization* campaign (Appendix Table A.6).[33] Thus, the baseline result is unlikely to be confounded by systematic mis-measurement of production.

Column (5) controls for 1928 mortality rates interacted with the famine indicator. This addresses potential mechanical mean reversion in mortality rates. Again, the ethnic Ukrainian interaction coefficient is nearly identical to the baseline.

In column (6), we address the concern that province fixed effects absorb meaningful variation in ethnic

[33]The results are similarly robust if we control for different normalizations of the drop in livestock, such as livestock decline divided by sown area. These are not reported for brevity and available upon request.

Ukrainian share and famine mortality by including the uninteracted Ukrainian effect instead. The ethnic Ukrainian interaction coefficient in column (6) is nearly identical to the baseline in column (2). The uninteracted Ukrainian coefficient is -0.007 and statistically significant at the 1% level. This is similar to the raw mortality figures in showing that in non-famine years, ethnic Ukrainian population share is *negatively* associated with mortality. It is only during the famine that mortality is *positively* associated with ethnic Ukrainian population share. The sum of the interaction and uninteracted coefficients presented at the bottom of the table is positive and statistically significant at the 1% level.

4.2 Ukrainian Bias Outside of Ukraine, Other Ethnic Minorities

In 1926, 25% of ethnic Ukrainians in our sample lived outside of Ukraine, which is one province in our sample. Column (7) examines the Ukrainian-mortality gradient in other provinces by omitting Ukraine. Column (8) additionally omits three other regions where agricultural productivity was particularly high and many Ukrainians lived (Lower Volga, North Caucasus, and West Siberia). The Ukrainian interaction coefficient is positive and statistically significant in both cases. Thus, the bias against ethnic Ukrainians in famine mortality extends across the Soviet Union, beyond Ukraine. In column (8), the standardized coefficient is smaller because the variation in Ukrainian population share is lower in the restricted subsample.

The results in columns (7) and (8) go against the concern that our results are driven by outliers or a few influential observations. Another way to demonstrate that our results are unlikely to be driven by coincidence is to conduct random permutation tests. We alternately permute Ukrainian population shares and mortality rates across provinces (while preserving the mean and variance of the sample distribution) and estimate the same equation using randomly assigned measures. In each permutation, we find that the probability that our main result is due to coincidence is 1% or less.[34]

One limitation with the data is that population is not reported for every year and we interpolate population for the missing years.[35] Column (9) shows that our results are similar if we alternatively use population data from the 1926 census instead of the interpolated data. The dependent variable is the number of deaths in province i and year t divided by total population in province i in 1926 (instead of year t), and we control for urban population share in 1926 (instead of year t) interacted with famine. The robustness of these estimates also alleviates any concerns about controlling for contemporaneous urban population, which may be endogenous to the famine, in the baseline.

In column (10), we add an interaction of the pre-famine share of non-Ukrainian ethnic minorities with the famine dummy variable. If bias against Ukrainians was part of a larger policy of bias against all ethnic minorities, we should find that this interaction coefficient is also large, positive and statistically significant. Instead, we find that it is negative, negligible in size and only marginally significant. This interaction coefficient becomes insignificant in the district-level estimates presented later in the paper. Thus, we will interpret this as a statistical zero. These results show that bias in famine mortality was exclusive to Ukrainians.[36]

[34] This is done for 10,000 iterations. See Appendix Figure A.4.

[35] See Data Appendix.

[36] Other ethnic minorities include Belorussians, Tatars, Mordvins, Chuvashs, Germans, Bashkirs, Jews, Poles and all other non-Ukrainians and non-Russians. When we examine each minority group separately, we find that ethnic Germans suffered higher

Back-of-Envelope Calculation

We quantify the total contribution of Ukrainian bias in famine mortality with a simple back-of-the-envelope calculation using the estimates in Table 2 column (2). The regression predicts that the number of deaths is on average 2.72 million in non-famine years and 4.97 million in 1933.[37] The number of excess deaths due to the famine is the difference between mortality during famine and non-famine years: $4.97 - 2.72 = 2.26$ million (with rounding error). We obtain the counterfactual famine mortality in a world with no anti-Ukrainian bias by setting the interaction coefficient of ethnic Ukrainian population share and the famine dummy variable in equation (1) to zero. When we do this, predicted deaths in 1933 is 3.23 million. The number of famine deaths without bias against ethnic Ukrainians is the difference between this number and the number of deaths in non-famine years, 0.51 million ($3.23 - 2.72 = 0.51$ million). Thus, bias against Ukrainians accounts for 77% ($1 - 0.51/2.26 = 0.77$) of famine deaths in our sample. Since most non-Ukrainians in our sample are Russians, who suffered much lower famine mortality rates, our results imply that total famine mortality would have been 77% lower if ethnic Ukrainians died at similar rates as ethnic Russians.

We repeat the exercise for only Ukraine and find that bias against ethnic Ukrainians accounts for 92% ($1 - 0.12/1.51 = 0.92$) of famine deaths in Ukraine. The larger magnitude of the contribution for anti-Ukrainian bias to famine mortality in Ukraine is due to the higher share of ethnic Ukrainians living in Ukraine.

Our preferred interpretation is that Ukrainian bias in famine mortality is due to anti-Ukrainian bias in policies that contributed to mortality. It is important to note that the data used in our analysis were collected by the Soviet government and available to central planners.[38] Our interpretation assumes that the estimate of Ukrainian bias is not confounded by spurious factors (e.g., variables that are correlated with pre-famine ethnic Ukrainian share and famine mortality, but unrelated to anti-Ukrainian bias in famine-era policy) and that anti-Ukrainian bias in policy led to famine mortality. We provide evidence for both of these assumptions in the subsequent analysis. We show that the Ukrainian bias in famine mortality is not confounded by omitted variables, and provide positive evidence that there is anti-Ukrainian bias in food procurement policy.

4.3 Robustness

4.3.1 Dynamic Estimates

To investigate the dynamic patterns of the mortality-Ukrainian gradient and provide evidence against spurious correlations, we estimate an equation similar to before, except that we interact ethnic Ukrainian population share, grain production, and urbanization with dummy variables for each year instead of only 1932. The reference year is 1923. Figure 1b plots the interaction coefficients and their 95% confidence intervals and shows a sharp temporal pattern that goes against the concern that the baseline estimate is confounded by

famine mortality. However, because of the concentration of Germans in a few areas, we do not have adequate support in the data to meaningfully control for other factors such as urbanization and production. Recall that Kazakhstan, which suffered tremendous mortality, is not in our sample.

[37] This is comparable to the 4.81 million deaths in 1933 in the raw data.

[38] The one exception is the suitability data.

spurious correlations.[39] The correlation between Ukrainian population share and mortality becomes positive in 1932 and peaks in 1933. There is no correlation in other years.

4.3.2 Sensitivity to Alternative Measurements

The following exercises address the fact that our main dependent and explanatory variables may be measured with error and examines the sensitivity of our estimates to alternative measures.

Famine Severity We infer famine severity from the increase in mortality during the famine relative to other years. However, one may be concerned that many famine deaths were not registered either because of reduced bureaucratic capacity or because of intentional understating. If undercounting is increasing with famine severity (i.e., ethnic Ukrainian population share), then our estimates will understate anti-Ukrainian bias. We address this by examining two alternative measures of famine severity. The first is natality. Crude birth rate data also face the problem of under-registration, but the bias goes in the opposite direction as mortality. Live births should be decreasing in famine severity since starvation is negatively associated with the probability of a healthy pregnancy or birth, and is positively associated with the probability of miscarriage and stillbirths (Dyson and Ó Gráda, eds, 2002). If there is more under-registration of births in places with more severe famine (i.e., regions with higher ethnic Ukrainian population share), then our estimates will overstate anti-Ukrainian bias. Thus, qualitatively similar findings between the natality and mortality estimates will lessen the concern that anti-Ukrainian bias in famine mortality is an artifact of the under-registration of mortality.

Second, we use birth cohort size measured in the 1939 Census as the dependent variable. We use the place of residence and age in 1939 to create a synthetic panel of province-specific birth cohort sizes. This measure does not face the problem of under-registration (although it does introduce measurement error from migration).[40] Nevertheless, finding consistent results with this alternative data will be reassuring since they do not face the same measurement error problems as the official vital statistics data. The birth cohort size sample has more observations because provinces in the 1939 Census are smaller than those we use in the earlier samples. We do not have annual population data for these smaller units and therefore normalize cohort size with 1939 total province population. We control for 1926 urban population share interacted with the famine instead of time varying urban population share and its interaction with the famine.[41]

Table 3 rows (1) and (2) present the natality and birth cohort size estimates from the baseline specification. The interaction coefficients are all negative and statistically significant at the 1% level. A back-of-the-envelope exercise using the natality and birth cohort size estimates show that anti-Ukrainian bias explains 54% of missing births and 50% of missing survivors in 1939 in Ukraine, and 26% of missing births and 14% of missing survivors in 1939 in Belarus, Russia and Ukraine. The dynamic estimates for the two outcomes shown in Figures 1c and 1d exhibit similar temporal patterns. We focus the discussion on the more precise

[39] The coefficients and their standard errors are presented in Appendix Table A.5.

[40] Using birth cohort size to proxy for famine severity in a given province assumes that in 1939 people live in the same places as in 1932 and ignores the cross-province migration during the seven interim years.

[41] Alternatively, we can normalize birth cohort size with the average size of all birth cohorts in each province as reported in the 1939 Census. The results are nearly identical.

estimates of birth cohort size. The interaction coefficients begin to become statistically negative in 1926, which means that during the famine, those who were around eight years of age or younger were more likely to die from famine than older cohorts. Provinces with high ethnic Ukrainian population share experienced the largest decline in births and child survival after the harvest of 1932.[42] The estimates for natality and birth cohort sizes show that the main results are unlikely to be an artifact of mis-measurement in the mortality data.

Weather Given the prominence of weather in studies of famine, we are cautious and control for it in several different ways in addition to the baseline control of weather-predicted production. Table 3 Panel A directly controls for weather. Row (4) controls for spring and summer temperature and precipitation and the following year's winter temperature and precipitation.[43] Row (5) controls for monthly temperature and precipitation and their squared terms. Row (6) controls for monthly temperature and precipitation and their interactions. Row (7) controls for monthly weather following the standard in the literature, where the weather shock indicator is equal to one if the month's temperature or precipitation is more than one standard deviation away from the long-term (1900–50) mean. Row (8) controls for the deviations from the long-term mean of monthly temperature and precipitation for the 12 months of year t. Finally, as in Rozenas and Zhukov (2019), row (9) controls for the monthly deviations of temperature and precipitation for years $t-1$ and t. These additional controls cause little change in the interaction coefficient of interest relative to the baseline in row (3).

Ukrainian Population Share and Mortality Rates The baseline measure of famine severity is total mortality rate because this is the standard measure in the famine literature, and in our context, this variable is available for a longer time horizon than rural or urban mortality rates.[44] We address the fact that most mortality was rural by controlling for urban population share and its interaction with the famine dummy variable. Here, we show that the results are robust to using alternative measures of mortality. Table 3 Panel B presents the estimates. Rows (10) and (11) replace total mortality with urban and rural mortality as the dependent variable. The estimate for urban mortality is small in magnitude and statistically insignificant. The estimate for rural mortality is large and statistically significant at the 1% level. The results confirm that higher famine mortality in regions with a larger share of ethnic Ukrainians was mostly a rural phenomenon.

Next, we use different measures of Ukrainian population share as the explanatory variable. Instead of the the baseline measure of rural ethnic Ukrainian share, we alternately use the total population or urban population share of ethic Ukrainians. The standardized coefficients presented in italics in rows (12) and (13) are similar. Another proxy for ethnicity is the share of people whose mother tongue is Ukrainian, which we observe in the 1926 and 1897 censuses. Using the 1897 measure is an important robustness check for the concern that people may have had incentives to misreport ethnicity in the Soviet Regime. Rows (14) and (15) show that the results are similar with these alternative measures of Ukrainian share.

[42] The coefficients and standard errors are presented in Appendix Table A.5. See Appendix Section D for a detailed discussion of the interpretation for and descriptive statistics of natality and birth cohort sizes.

[43] Weather in the spring and summer of 1931 and 1932 is discussed most often as a cause of poor harvests during the famine (e.g., Davies and Wheatcroft, 2009; Tauger, 1991). 1933 winter conditions could have contributed to famine mortality.

[44] Rural and urban mortality are available starting in 1926, while total mortality is available starting in 1923.

4.3.3 Omitted Variables

The following exercises address concerns that pre-famine Ukrainian population share may be correlated with omitted variables that influence famine severity.

Demographic Structure To address the concern that higher mortality in ethnic Ukrainian areas may be due to differences in the demographic composition across regions (i.e., young children are more vulnerable), Table 3 row (16) 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. Appendix Table A.6 Panel B also shows that the results are similar if we control for the share of the elderly and a large number of alternative controls for the age and gender structure.

Other Food Production We follow existing studies of the Soviet famine and focus on the production of grain because it was the main agricultural commodity and source of calories for the population, and because there are no systematic disaggregated data for the production of other foods. If other food production was systematically different in Ukrainian areas, then our interpretation of the mortality-Ukrainian gradient would be confounded. To address this, we control for the suitability of each province for the cultivation of other crops. Since suitability is mostly determined by geography, we first address this by controlling for the interaction of latitude, longitude and the famine dummy, and the lower order interaction terms. Then, we control for the suitability for each of the other staple and cash crops — potato, wheat, rye, sugar beets, sunflowers, and flax — as predicted by the FAO GAEZ model interacted with the famine dummy.[45] Table 3 rows (17) and (18) present the estimates. The Ukrainian interaction coefficient is robust.

Social Capital and Other Slow-Moving, Historical Factors A recent study by Buggle and Durante (2021) argues that social capital can play an important role for surviving famines. We investigate the role of Ukrainian-specific cultural norms or social networks by examining the 1892 famine. This was the last large famine in the Russian empire, during which approximately 500,000 people died. We obtain province-level mortality data from 1885 to 1913.[46] Table 3 row (19) presents the baseline specification for the 1892 famine and show that famine mortality is not associated with ethnic Ukrainian population share. Thus, the Ukrainian-mortality gradient is specific to the Soviet famine of 1932-33 and unlikely to be explained by slow-moving features of Ukrainian culture or institutions.

We also directly control for historical factors that could have affected famine severity. Table 3 Panel D controls for the interactions of various proxies of pre-Soviet regional wealth: the nominal and real regional per capita incomes in 1897, regional labor productivity in 1897, regional rural labor productivity in 1897, the value of agricultural equipment in 1910 and livestock in 1916. It also controls for proxies of cultural

[45] Cotton was another important cash crop for the Soviet Union, but was mostly cultivated in regions outside of our sample.

[46] We thank Volha Charnysh for sharing 1885–1896 mortality and natality data from Charnysh (2022). The sample includes the fifty European provinces of the Russian Empire. Buggle and Durante (2021) argues that harsh winters encourage the formation of social capital, which is needed to avert and survive famines. The fact that our results are robust to controlling for the interaction of latitude, longitude and the famine dummy also suggests that our interpretation of the main results is not confounded by this slow-moving outcome of geography and history.

norms, institutions and inequality: 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 Population Census, the share of peasant households in repartition communes and the land Gini estimated from the 1905 Land Census, and per capita peasant revolts from 1895 to 1914. The main result is robust.[47]

Local Differences in Administrative Capacity or Political Zealousness This section addresses the concern that Ukrainian population share may be correlated with local administrative capacity or political zealousness to enforce Bolshevik policies. Table 3 row (34) controls for the share of votes for the Bolshevik Party in the 1917 Constituency Assembly election, the first and only election until the end of the Bolshevik rule. We interpret this as local suport for Bolshevik policies.[48] Row (35) controls for the size of the Communist Party in each province just before the famine: the number of Communist Party Members (averaged over 1922, 1927 and 1931) per 1,000 individuals in each province. Party members were the key enforcers of state policy in the countryside and were responsible for grain procurement. Row (36) controls for the number of delegates who attended the 1930 Party Congress. The Congress was a showcase of support where all delegates voted in the affirmative for comprehensive collectivization. Each province was encouraged to send a large number of delegates. We interpret the number of communists and 1930 Congress delegates as proxies for the administrative capacity and political zealousness. All three variables are time invariant and we control for the interaction of each and the famine dummy variable. The Ukrainian interaction coefficient changes little with these additional controls. This goes against the concern that anti-Ukrainian bias in famine mortality is due to differences in local administrative capacity or political alignment with Moscow.

LASSO In Table 3 row (37), we follow Chernozhukov et al. (2018) to Lasso all potential controls. We include all of the additional controls that are available for the whole sample and presented in Tables 2, 3 and A.6: *dekulakization* and drop in livestock, demographic structure, latitude, longitude and their interaction, suitabilities for various crops, historical wealth and cultural norms and institutions, political controls and weather shocks. The Ukrainian interaction is similar to the baseline.

4.4 Within-Province Patterns

Soviet policies were centrally planned, implemented top-down by the bureaucracy and usually used similar allocation rules at various administrative levels. If the patterns we observe are driven by centrally planned policies, then we would expect to see positive famine-mortality-Ukrainian-share gradient across districts within provinces as well as across provinces. Our district-level panel consists of two years: 1928 and 1933. Almost all of the data are manually collected from former Soviet archives. District-level mortality data are

[47] Rural labor productivity in 1897 is measured as the upper and lower bound estimates from Markevich (2019). The value of agricultural equipment in 1910 and livestock in 1916 is from Castañeda Dower and Markevich (2018). Living in a repartition commune required a more cooperative behavior (than in a hereditary commune), and according to the 1905 Land Census, repartition communes were less widespread in Ukrainian-populated regions than in Russian-populated regions. If the values of cooperation were transmitted down generations, this difference could contribute to the difference in mortality between the two ethnicities.

[48] Approximately 60% of the eligible voters turned out to vote. We follow Castañeda Dower and Markevich (2021) and use disaggregated district-level data on votes for the Bolsheviks from Protasov et al. (2014).

only available for Russia and Ukraine and we have fewer variables at this more disaggregated level (e.g., there are no data for production or procurement).

Table 4 column (1) replicates the baseline specification with district and year fixed effects instead of province and year fixed effects. Conceptually, famine mortality in this regression is the difference between 1933 and 1928 mortality rates. Since we do not observe production at the district level, the baseline controls for the suitability for grain cultivation from the FAO GAEZ database and its interaction with the famine year dummy; and weather (monthly temperature and precipitation in years t and $t-1$).[49]

Column (2) and all subsequent columns control for province-year fixed effects, which isolate the within-province variation and control for factors that vary by province and year (e.g., regional political competition, leadership differences across provinces).[50] The spatial patterns are similar to the province-level estimates. Within provinces, famine mortality is increasing with ethnic Ukrainian population share.

Column (3) controls for an alternative measure of weather: the deviations from long-term (1900-1950) means of monthly temperature and precipitation in years t and $t-1$. The results are similar. In column (4), we omit Ukraine and show that the patterns are similar in Russian provinces. In column (5), we additionally omit Lower Volga, North Caucasus and West Siberia. The patterns are still present, meaning that Ukrainian bias is systematic even in provinces with a low share of Ukrainians overall.

In column (6), we show that there is no bias for other ethnic minorities. The coefficient is small in magnitude and statistically insignificant.

The results are consistent with the presence of a systematic and centrally planned policy that targets ethnic Ukrainians.

We show that the results are qualitatively similar when we subject the district-level estimates to the same sensitivity checks as the province-level estimates.[51]

4.5 Administrative vs. Ethnic Boundaries

Past discussions about Ukrainian famine mortality have focused on the difference between Ukraine and other republics. Yet, our estimates show that Ukrainian bias in mortality exists outside of Ukraine. In this section, we use the district-level data to directly investigate the importance of administrative borders by examining the change in district-level famine mortality as one crosses from Ukraine to Russia. We plot raw famine mortality, the difference between 1933 and 1928 mortality rates, against the distance to the border

[49] We can alternatively measure weather as the deviations from long-term (1900-1950) means of monthly temperature and precipitation in years t and t-1. The results are nearly identical. Thus, we do not present the latter for brevity. They are available upon request. Also, we use urbanization from 1926 and 1933 because urbanization at the district level is not available for 1928.

[50] When using the district-level panel, we estimate standard errors that are adjusted for spatial correlation. We follow the recommendations by Colella et al. (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 (most conservative) standard errors. We also cluster the standard errors at the district level. They are very similar and presented in brackets in the table.

[51] We do not observe age structure at the district level. Thus, we control for the gender ratio of the whole population to address potential differences in demographic structure. See Appendix Table A.7. We also conduct a random permutation exercise with the district-level data and alternately permute ethnic Ukrainian population share and 1933 mortality across districts. Like the province-level results, we find that our estimates are unlikely to be driven by coincidence. We estimate the specification as in Table 4 column (2) using the randomly assigned ethnic Ukrainian population share or 1933 mortality. This is done for 10,000 iterations. Comparing the distribution of the coefficients from the permutations to the baseline estimate (indicated by the vertical red line) in Appendix Figure A.4, we find that the probability that the latter is due to coincidence is less than 1%.

between Ukraine and Russia, together with the fitted lines and their 95% confidence interval. Figure 2a shows that there is discrete decline in famine mortality rates as one crosses the border from Ukraine to Russia. Famine mortality rates are lower in Russia by 2.5 to 3.6 percentage-points.[52] This is consistent with survivor accounts of notably lower mortality across the border (e.g., Applebaum, 2017, Ch. 10, 11).

The border effect on raw mortality rates also sheds light on the contribution of the migration ban on Ukraine imposed in January, 1933. To see this, consider the hypothetical scenario of free mobility. In that case, we should not observe discrete changes in famine mortality rates along any administrative border. Note that there is more variance in famine mortality rates in Ukraine than in Russia. This is because the variance in mortality increases with mean mortality (Appendix Figure A.1a).

Figure 2b plots the mortality residuals from a regression controlling for 1926 Ukrainian population share against distance to the border. There is no border effect once we control for the ethnic Ukrainians rural population share of each district. Conceptually, this accounts for the decline in ethnic Ukrainian share when crossing the border (see Figure 2e).[53] Thus, anti-Ukrainian bias was delineated along ethnic and not administrative lines.

To illustrate the importance of ethnic Ukrainian bias versus other factors such as agricultural suitability or urbanization, we present mortality residuals from a regression that only controls for pre-famine urban share (Figure 2c), or agricultural suitability and weather (Figure 2d). Controlling for these variables, unlike controlling for Ukrainian population share, does not diminish the border effect.[54] These results emphasize the importance of ethnic Ukrainian share and the unimportance of the other factors in explaining famine mortality.

4.6 Realized Procurement and Retention

Ukrainian bias in famine mortality is consistent with the presence of anti-Ukrainian bias in policy. In this section, we investigate whether there is bias in the most important policy for food distribution: centrally planned grain procurement. Positive statistical evidence addresses the lack of conclusive documentary evidence.

Table 5 documents the relationship between Ukrainian population share, centrally planned procurement, retention and famine mortality. The sample size is smaller than the main analysis because of the limited availability of procurement and production data. In column (1), the dependent variable is realized procurement as a share of realized production. We use our corrected grain production. The estimate is similar to equation (1), except that we no longer control for predicted grain and its interaction with famine because grain production is the denominator of the dependent variable. The Ukrainian interaction coefficient is positive, 0.185, and statistically significant at the 1% level. This implies that, all else equal, the share of production taken away from a province that was 100% Ukrainian was 18.5 percentage-points higher than a province that had no Ukrainians.

[52] We estimate the size of the decline in mortality by regressing excess mortality on a dummy variable that equals one if the district is in Russia. We do this with samples that include districts within 100 to 250 kilometers on either side of the border and find that raw famine mortality rates were on average 25 to 36 per 1,000 lower in Russia. See Appendix Table A.8 Panel A.

[53] The difference in residualized mortality on the two sides of the border is statistically zero (see Appendix Table A.8 Panel B).

[54] For border effect estimates, see Appendix Table A.8 Panels C and D.

Column (2) examines per capita retention (production minus procurement) as the dependent variable. It shows that all else equal, a province with 100% rural Ukrainian population share retained 1.085 kilograms per capita per day less than a province with no Ukrainians. The Ukrainian interaction coefficient is positive and statistically significant at the 1% level. To examine the timing of anti-Ukrainian bias in food retention, we replace the interaction of rural ethnic Ukrainian population share and the famine dummy with the interaction of rural ethnic Ukrainian population share and year dummy variables. 1928 is the omitted reference group. Figure 1e shows that Ukrainian population share is uncorrelated with food retention for most years, but negatively correlated during the famine.[55] These estimates show that there was anti-Ukrainian bias in centrally planned food procurement during the famine.

Columns (3) and (4) document the relationship between mortality and grain retention. In column (3), we regress mortality on realized grain retention and its squared term. Consistent with the positive and concave relationship between food consumption and mortality, we find that the coefficient for grain retention is negative and the coefficient for the squared term is positive. Both estimates are statistically significant at the 5% level. In column (4), we add the interactions of these variables with the famine dummy. The interactions are statistically zero, while the uninteracted terms are similar to column (3). This is a sanity check and reflects the fact that the biological relationship between food consumption and mortality is stable over time. The estimates support the interpretation that the Soviets procured a higher share of 1932 production from Ukrainians, which resulted in lower food retention and higher famine mortality.

To understand the importance of anti-Ukrainian bias in grain procurement policy for famine mortality, we conduct a simple quantification exercise that compares famine mortality when there is bias in retention to the counterfactual of no bias.

First, we estimate the relationship between mortality and grain retention so that we can predict mortality at different levels of retention.

$$mortality_{i,t+1} = F(retention_{it}) + \Gamma X_{it} + \eta_i + \delta_t + \varepsilon_{it}. \tag{2}$$

Equation (2) is conceptually similar to the estimates shown in Table 5, but with a more flexible functional form and fewer controls for parsimony.[56]

Second, we estimate famine mortality with anti-Ukrainian bias in grain retention. This is the difference between the number of deaths in 1933 predicted by realized 1932 retention and the number of deaths in non-famine years. The number of deaths predicted by realized retention can be obtained by applying 1932 realized grain retention to the estimates from equation (2): predicted 1933 mortality is 1.42 million deaths in Ukraine and 4.41 million deaths in the full sample.[57] Average predicted mortality for non-famine years

[55] Appendix Table A.9 columns (7)-(9) report the coefficients and standard errors.

[56] $F(retention_{it})$ is a flexible step function defined over 0.5 kilogram per capita per day intervals of grain retention, X_{it} includes province characteristics that may affect mortality (e.g., urbanization), and η_i and δ_t are province and year fixed effects. We estimate this regression without a constant. Lower retention is associated with higher mortality. Appendix Figure A.5a plots the estimated step function. Appendix Table A.9 presents the coefficients and standard errors. The results are robust if in addition to urbanization, we control for urbanization interacted with the famine indicator and the official 1928 grain production interacted with the famine indicator.

[57] These numbers are slightly below the mortality numbers in the raw data: 4.81 and 1.86 million deaths for the full sample and Ukraine because there are few observations with exceptionally high mortality in the data for which our estimated function under-predicts mortality.

is 0.52 million for Ukraine and 2.72 million for the full sample.[58] Thus, famine mortality with bias is 0.9 million for Ukraine ($1.42 - 0.52 = 0.9$) and 1.7 million for the whole sample ($4.41 - 2.72 = 1.7$).

Next, we estimate famine mortality for the counterfactual of no bias in grain retention. This is the difference between the number of 1933 deaths predicted by the counterfactual of no bias in retention and the number of deaths in non-famine years. To calculate the counterfactual 1933 mortality, we first predict counterfactual 1932 grain retention. We use the estimates of the relationship between retention and the interactions of Ukrainian population share and year dummy variables shown in Figure 1e. We set Ukrainian share to zero in 1932 to predict counterfactual grain retention for the famine. Then, we apply the counterfactual grain retention to the estimates from equation (2) to predict the counterfactual 1933 mortality: predicted 1933 deaths are 0.9 million in Ukraine and 3.84 million in the full sample. Thus, the counterfactual famine mortality is 0.38 million for Ukraine ($0.9 - 0.52 = 0.38$) and 1.12 million for the whole sample ($3.84 - 2.72 = 1.12$).

Finally, we estimate the contribution of bias in grain retention on famine mortality by comparing famine mortality when there is anti-Ukrainian bias to the counterfactual famine mortality of no bias in retention. It follows that anti-Ukrainian bias in grain retention explains 58% of excess deaths in Ukraine ($1 - 0.38/0.9 = 0.58$) and 34% of all excess deaths in our sample ($1 - 1.12/1.7 = 0.34$). Since overall anti-Ukrainian bias explains 77% and 92% of famine mortality in the whole sample and Ukraine, the estimates imply that approximately half of the total anti-Ukrainian bias effect on famine mortality takes place through bias in food retention.[59]

The magnitudes are consistent with the importance of grain procurement for determining food availability, but also leaves room for other policies to contribute to food availability, such as migration restrictions or aid relief (recall earlier discussions in Sections 4.4 and 2).

4.7 Production and Procurement Targets

To understand whether anti-Ukrainian bias in grain procurement was centrally planned and intentional, we examine grain production and procurement targets published before the famine. The First Five-Year Plan, published in 1928, laid out production and procurement targets for each province for the years from 1928 to 1933.[60] In Table 6 column (1), we regress per capita production targets on the rural share of ethnic Ukrainians. We control for year fixed effects to account for the fact that the planners assumed a high rate of growth in all regions, and officially reported 1928 per capita grain production, which was the measure used by planners to account for regional agricultural productivity. We do not control for province fixed effects because of the limited variation in a five-year panel. The coefficients for Ukrainians and grain 1928 are both positive and statistically significant. For two places with the same observed production in 1928, the one with more ethnic Ukrainians was expected to produce more. This could reflect the regime's desire to be harsher

[58] These are predicted using the estimates from Table 2 column (2).

[59] We conduct a similar exercise for natality and find that anti-Ukrainian bias in grain retention explains 13% of missing births in our sample and 26% of missing births in Ukraine. Figure A.5b shows a similar estimate for natality. As expected, the relationship is increasing. As with mortality, the bias in grain retention accounts for half of the total contribution of anti-Ukrainian bias to reduced natality during the famine.

[60] See the Data Appendix for details.

with Ukrainians or a belief that Ukrainian peasants were less productive than other peasants living in similar natural conditions and had more room to expand production.

Column (2) examines per capita procurement targets as the dependent variable, while controlling for production targets, which captures differences in perceived production capacities across regions. For two provinces assigned the same production target, the one with more Ukrainians is assigned a higher procurement target. Column (3) shows a similar pattern when we examine procurement as a share of production as the dependent variable. Thus, the regime intended to take more grain from Ukrainian areas after conditioning for factors such as production capacity. Since we control for production targets, the Ukrainian bias in grain procurement targets cannot be driven by the belief that Ukrainian grain production had more room to grow.

In column (4), we replace procurement ratios with per capita grain retention targets (the difference between the production and procurement targets) as the dependent variable. The coefficient for ethnic Ukrainian population share is -0.275 and statistically significant at the 1% level. For two places facing the same production targets, central planners intended for the one with 100% ethnic Ukrainians to have 0.275 kilograms of grain per capita per day less than the one with no Ukrainians. The magnitude of the discrepancy is sizable: it is approximately one-third of the official Soviet food requirement for heavy labor (see Section 3).[61]

The target data show that as early as 1928, the regime had planned for Ukrainian areas to retain less grain than other regions with the same level of grain production. However, the data do not imply that the regime planned to kill Ukrainians because the production targets are much higher than actual grain production for all regions and the degree of unequal grain retention implied by the target data would not result in famine mortality at high levels of grain retention. Thus, the results leave open the question of the timing of the decision to let Ukrainians die. In one extreme scenario, the regime aimed to reduce the Ukrainian population size *ex ante* and set procurement targets in 1928 that would lead to high Ukrainian famine mortality. This assumes that the regime secretly knew what true 1932 production would be, but planned to enforce the procurement targets. In an alternative scenario, the regime aimed to penalize Ukrainians for being troublesome by leaving them with less surplus without the intention to cause famine. Stalin was known to have advocated using the over-procurement of food to discipline peasants (Danilov et al., eds, 1999-2006; Davies and Wheatcroft, 2009), and to have rewarded loyal ethnic groups and penalized disloyal ones (Polyan, 2001). In this scenario, the regime intended for Ukrainians to retain less food than others, but was sincerely overoptimistic about future production and planned for Ukrainians to retain enough food for subsistence; when production fell in 1932, the state made the decision to enforce higher procurement from Ukrainians. The truth may also be somewhere in between: the leadership knew that production targets were too optimistic, but did not fully predict the severity of the fall in production in 1932; and faced with the choice of giving up procurement versus letting Ukrainians die, they chose the latter.

[61] Note that the estimated effect on retention targets is not directly comparable to the estimated effect on realized procurement because of slight differences in sample and controls (the latter controls for province fixed effects to be comparable to the mortality regressions).

5 Additional Results

5.1 Rural population and Grain Production after the Famine

This section examines long-run economic and demographic patterns in areas that had a higher share of Ukrainian population and suffered disproportionately high famine mortality rates.

First, we use all available population censuses from 1897 to 2002 to examine total rural population.[62] We regress log rural population size on the interaction of ethnic Ukrainian share and (census) year fixed effects, controlling for province and year fixed effects. Figure 3a plots the interaction coefficients and its 95% confidence intervals.[63] The figure shows that relative to areas with no Ukrainians, the rural populations in Ukrainian areas declined in size immediately after the famine and had not recovered by 1939. WWII disrupted data collection after 1939. Rural population had recovered by 1959 in the first population census after the war.

In Figure 3b, we examine the rural population share of ethnic Ukrainians and ethnic Russians.[64] The estimates are imprecise and should be interpreted as suggestive. Nevertheless, the patterns are striking: the share of ethnic Ukrainians declined after the famine and the decline was permanent. This goes against the concern that famine mortality in regions with high ethnic Ukrainian population share was driven by the deaths of other ethnicities and not that of ethnic Ukrainians ("ecological fallacy"). The statistical evidence is consistent with the fact that there are no known accounts of such mortality patterns.

Second, we examine grain production. We regress log total grain production for each province and year on the interaction of pre-famine ethnic Ukrainian population share and year fixed effects, controlling for the time-invariant measure of suitability for grain cultivation interacted with year fixed effects, 48 monthly temperature and precipitation variables for years t and $t-1$, and province and year fixed effects. Figure 3c plots the Ukrainian interaction coefficients. Consistent with the republic-level accounting exercise, we observe a larger production drop in Ukrainian areas. In addition, we observe that production in these areas remained lower than others for a total of four years and had recovered by 1936.[65] Thus, production in famine-stricken areas recovered before the labor supply.

Agriculture in the early 1930s still mostly relied on traditional means of production and a key feature of Soviet economic policy was to mechanize agriculture. The most important mode of mechanization was the adoption of tractors, which were centrally allocated by Moscow. Thus, we examine the allocation of tractors to understand whether the regime increased mechanization in famine-stricken regions to boost production. We collect archival data on province-level tractors allocation for 1927 to 1939. In normal times, tractors were allocated based on sown area and the importance of a location for agricultural production. Thus, we regress

[62] For 1939, we use revised estimates for rural population provided by Russian demographers (Bogoyavlensky, 2013) because the official figures have been inflated. The revised estimates are not available by ethnicity. Each observation is a province in a census year. The data are harmonized to use 1932 province borders. Population size in the census reflects mortality, fertility and migration.

[63] Appendix Table A.11 column (1) reports the coefficients and their standard errors.

[64] Figure 3b examines ethnic composition until 1989 because the 2002 Population Census only reports ethnic composition for total provincial population and not rural population. We find that the ethnic composition results persist into 2002 if we examine total provincial population as the outcome variable. These results are available upon request.

[65] As with the accounting exercise, we use corrected grain figures for the years 1928-1933 and official grain figures for all other years. The coefficients and standard errors are shown in Appendix Table A.10. The results are similar without controlling for weather. They are available upon request.

total tractor horse power per hectare of 1928 sown area on the interaction of year dummy variables and the pre-famine Ukrainian population share, controlling for the time-invariant measure of suitability for grain production interacted with year fixed effects, and province and year fixed effects. Figure 3d shows small increases in the number of tractors allocated to regions that had a high pre-famine Ukrainian population share during 1927 to 1932, and a large increase during 1933 to 1939. These estimates together with the results on rural population and production imply that tractors were used to moderate the loss of labor and boost production in famine-stricken regions.[66]

5.2 Motivation of Anti-Ukrainian Bias

As we discussed in the Introduction, the primary political economic explanation of Ukrainian famine mortality is that Ukrainians were repressed because of their importance to Soviet agriculture and their strong resistance to Soviet rural economic policy. It is beyond the scope of this paper to provide conclusive evidence on the motivations of anti-Ukrainian bias. This section provides a speculative discussion to help shed light on this important question and provide some clues for future research.

Ukrainians had a strong group identity that included their own language and culture, which facilitated collective action. The Ukrainian communist party was the largest national branch of the Soviet Communist Party and viewed itself as representing the interests of ethnic Ukrainians across the Soviet Union, including those who lived outside the boundaries of the republic. Strong political opposition from Ukrainian nationalists during the Civil War was central to the Bolsheviks' "national question" (Graziosi, 2015). The Bolsheviks were concerned about controlling all peasants, but especially concerned about controlling the Ukrainian peasantry.

We provide two stylized facts consistent with this political-economic motive. The first is to validate the claim that ethnic Ukrainians offered stronger resistance to collectivization than other ethnic groups. We collect data on de-classified secret police reports about peasant resistance to collectivization.[67] We find that the positive slope between collectivization (the share of households that have been collectivized) and resistance is steeper in the ten provinces with Ukrainian population share above the sample median than in the nine provinces with Ukrainian population share below the median, controlling for official 1928 grain production and urban population share. We plot the fitted lines and residuals in Figures 4a and 4b.[68] They show that all else equal, Ukrainian peasants resisted collectivization more intensely. The residuals in Figure 4b show that this is true for ethnic Ukrainians outside of Ukraine. This is consistent with the Bolsheviks'

[66] The coefficients and standard errors are shown in Appendix Table A.10. Note that tractors allocated in 1933 are used to produce food for 1934. Also, the allocation of tractors may have been accompanied by other inputs such as better fertilizers, which was not systematically recorded. Thus, one can interpret tractors as broadly reflecting the use of advanced inputs or the mechanization of agriculture.

[67] The reports are from January 1931 to March 1932, just before the peak of the famine. They include three types of resistance. The first is the number of anti-Soviet "violent acts" per 1,000 people, which include murders or attempted murders of local officials, arsons and the destruction of collective farm or state property. The second is the number of mass demonstrations in the countryside. The third is the number of episodes when anti-Soviet leaflets were distributed in the countryside. For brevity, we examine the first principal component of the three indicators on a cross-section of provinces. The results are similar if we examine each measure separately.

[68] In the subsample of provinces with low (high) Ukrainian population share, the mean and standard deviation of collectivization are 0.437 and 0.111 (0.642 and 0.147). The p-value for the statistical difference between the two slopes is 0.09.

fear that political resistance was stronger from Ukrainians, as well as our finding that anti-Ukrainian bias extended to ethnic Ukrainians across the Soviet Union.

Second, we document that the famine mortality-Ukrainian gradient is steeper in places that were more important for agricultural production. This implies that there was more intense targeting of Ukrainians in agriculturally productive places. We regress mortality on the triple interaction of Ukrainian population share, the famine dummy variable, and the importance of a region for rural economic production as perceived by the state (measured with official 1928 grain production). Table 7 shows that the triple interaction coefficient with 1928 grain production is positive and statistically significant at the 1% level. In column (2), we control for a parsimonious measure of administrative capacity and political zealousness (the first principal component of the variables that we examined earlier in Table 3 Panel E). The triple interaction coefficient with 1928 grain production is robust to the additional control, while the triple interaction coefficient with the administrative capacity proxy is small in magnitude and statistically insignificant. Thus, the Ukrainian-mortality gradient is steeper in places that are more important for agriculture, but does not vary with these other factors that are unrelated to agriculture. Figure 1f plots the dynamic triple interaction coefficients. The timing is very sharp. The triple interaction is zero in all years except during the famine, when it spikes up. This goes against concerns of spurious correlations.[69] Column (3) repeats the estimate with district-level data.[70] The triple interaction of Ukrainian population share, suitability for grain cultivation and the famine dummy is positive, large in magnitude and statistically significant at the 1% level. As we discussed earlier in the paper, finding similar patterns at different administrative levels supports the interpretation that the patterns we observe are due to centrally planned policy.

Additional Mechanisms There are several complementary hypotheses for why the Soviets repressed Ukrainians for political economic reasons. A variation of the hypothesis discussed above stems from the well-known insights of Horowitz (1985). The logic is as follows. Given the need to control agriculture, the regime targeted ethnic Ukrainians because it lacked more precise information on the likelihood of subversion. Thus, ethnic Ukrainian population share was used as a crude marker on which central planners conditioned policies.[71]

Another hypothesis is that anti-Ukrainian bias is partly driven by asymmetric information and rigidities in the command economy. Meng et al. (2015) first made this argument for the Chinese famine, which occurred in a centrally planned procurement system modeled after the Soviet one. They argue that the famine was partly due to the government's inability to adjust procurement because it did not trust local information. Neither farmers nor bureaucrats were incentivized to report truthfully.[72] Thus, the government

[69] The coefficients and their standard errors are presented in Appendix Table A.5.

[70] We do not observe administrative capacity or grain production at the district level. Thus, we control for the interaction of suitability for grain production and the famine dummy instead of predicted grain production and its interaction with famine, and do not control for administrative capacity variables because we do not observe them at the district level.

[71] The Horowitz (1985) hypothesis is unlikely to explain all of the anti-Ukrainian bias in famine mortality. The Ukrainian-famine mortality gradient is still positive and statistically significant, although slightly smaller in magnitude than the baseline, if we control for resistance to collectivization interacted with the famine dummy in the baseline mortality regression. The result is not presented, but available upon request. The Horowitz (1985) hypothesis is consistent with the positive triple interaction effect of grain production in 1928, ethnic Ukrainian population share and the famine dummy on mortality if the regime believed that the Ukrainian-subversion gradient was steeper in agricultural regions.

[72] Farmers are not the residual claimants of production and therefore incentivized to under-report production or shirk. Local

will not lower procurement until they can verify that the production drop is not caused by peasants shirking (or if famine would be too politically costly). Information rigidities can explain the anti-Ukrainian bias in Soviet policy and famine mortality if the Bolsheviks were particularly distrustful of Ukrainians.[73] This theory of famine is consistent with the Bolsheviks' primary objective to control agricultural production and under-reporting production can be viewed as a form of Ukrainian resistance that threatens the Bolsheviks.[74]

The variants of the political economic motivation discussed in this section are not mutually exclusive; nor do they preclude other motivations for anti-Ukrainian bias.

6 Conclusion

The deaths of approximately 2.1 to 3.15 million ethnic Ukrainians during the 1932-33 Soviet Famine is one of the most controversial human tragedies of the 20th century. Ethnic Ukrainians in the Soviet Union declined 21.3% to 16.5% of the total population between 1926 and 1939. In areas that the Bolshevik regime marked as important for grain production, ethnic Russians replaced ethnic Ukrainians as the largest ethnic group.

This paper provides the first systematic evidence on the causes of Ukrainian famine mortality. We show that disproportionately high Ukrainian famine deaths were not an unintended consequence, but an outcome of anti-Ukrainian bias in Soviet policy. The Bolsheviks systematically and intentionally over-procured food from ethnic Ukrainians, province by province, district by district, within and outside of Ukraine.

Our study demonstrates the importance of grain distribution policy and migration restrictions on famine mortality. An interesting subject for future research is to better understand the other contributing factors. At a more macro level, our results raise the question of the political-economic tradeoffs of repressing Ukrainians for the regime. On the one hand, rural production recovered within a few years of the famine. On the other hand, the famine had long-run political (Rozenas and Zhukov, 2019) and economic consequences (Korovkin and Makarin, 2022, 2023). The fact that Russian officials removed the commemorative memorial of the victims of the *Holodomor* to combat "political disinformation" when they occupied Mariupol in 2022 highlights the importance of the famine for people today.[75] We hope that the rigorous empirical evidence provided in this paper will establish a ground truth of facts for future research and public discussions.

---

bureaucrats may over-report production to curry favor with the Party or under-report to build a local power base. Both are bad for the regime. Over-reporting can lead to over-procurement and famine, which can lower future production and be politically destabilizing. Under-reporting lowers revenues.

[73] This explanation is consistent with the positive triple interaction effect of grain production in 1928, ethnic Ukrainian population share and the famine dummy on mortality if distrust in Ukrainians is increasing in agricultural productivity.

[74] Note that there are important differences in the political economic backgrounds of the two famines. The Chinese famine occurred in the rural political power base of the Chinese Communist Party, including the home provinces of the top Party leaders. In contrast, by the time of the Ukrainian famine, there had been a history of political conflict between the Ukrainian peasantry and the Bolsheviks, whose political power base lie in urban areas. Agricultural production was also much more important to the Chinese economy in 1959 (approximately 80% of GDP) than the Soviet economy in 1928 (approximately 50% of GDP).

[75] See RIA Novosti channel in telegram, October 19, 2022, https://t.me/rian_ru/182454. Accessed October 24, 2022.

References

Acemoglu, Daron, Tarek Hassan, and James Robinson, "Social Structure and Development: A Legacy of the Holocaust in Russia," *The Quarterly Journal of Economics*, 2019, *126*, 895–946.

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya, "Radio and the Rise of The Nazis in Prewar Germany," *The Quarterly Journal of Economics*, 07 2015, *130* (4), 1885–1939.

Alesina, Alberto and Eliana La Ferrara, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, September 2005, *43* (3), 762–800.

Andreev, E.M., L.E. Darskij, and T.L. Kharkova, *Demographicheskaya Istroriya Rossii: 1927-1959*, Moscow: Informatika, 1998.

Applebaum, Anne, *Red Famine: Stalin's War on Ukraine*, New York: Penguin Press, 2017.

Bauer, Thomas K., Sebastian Braun, and Michael Kvasnicka, "The Economic Integration of Forced Migrants: Evidence for Post-War Germany," *The Economic Journal*, 2013, *123* (571), 998–1024.

Becker, Sascha O., Irena Grosfeld, Pauline Grosjean, Nico Voigtlander, and Ekaterina Zhuravskaya, "Forced Migration and Human Capital: Evidence from Post-WWII Population Transfers," *American Economic Review*, May 2020, *110* (5), 1430–63.

Berelowitch, A. and V. Danilov, eds, *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*, Moscow: Rosspen, 2000-2012.

Bogoyavlensky, Dmitry, "O pripiskakh v perepisi 1939 g [On the centralized additions in the 1939 census]," *Demoscope.ru*, 2013.

Buggle, Johannes and Stephen Nafziger, "The Long-Run Consequences of Labor Coercion: Evidence from Russian Serfdom," *The Review of Economics and Statistics*, 2019, *87*, 217–234.

Buggle, Johannes C and Ruben Durante, "Climate Risk, Cooperation and the Co-Evolution of Culture and Institutions," *The Economic Journal*, 01 2021, *131* (637), 1947–1987.

Burgess, Robin and Dave Donaldson, "Railroads and the Demise of Famine in Colonial India," Working Paper, MIT 2017.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller, "Bootstrap-Based Improvements for Inference with Clustered Errors," *The Review of Economics and Statistics*, 08 2008, *90* (3), 414–427.

Canay, Ivan A., Andres Santos, and Azeem M. Shaikh, "The Wild Bootstrap with a "Small" Number of "Large" Clusters," *The Review of Economics and Statistics*, 05 2021, *103* (2), 346–363.

Castañeda Dower, Paul and Andrei Markevich, "Labor Misallocation and Mass Mobilization: Russian Agriculture during the Great War," *The Review of Economics and Statistics*, 2018, *100(2)*, 245–59.

_ and _ , "Did Industrialization Increase Support for the Radical Left? Evidence from the 1917 Russian Revolution," *Working paper*, 2021.

_ , **Evgeny Finkel, Scott Gehlbach, and Steven Nafziger**, "Collective Action and Representation in Autocracies: Evidence from Russia's Great Reforms," *American Political Science Review*, 2018, *112* (1), 125–47.

Charnysh, Volha, "Explaining Out-Group Bias in Weak States: Religion and Legibility in the 1891/1892 Russian Famine," *World Politics*, 2022, *74* (2), 205–248.

Cheremukhin, Anton, Mikhail Golosov, Sergei Guriev, and Aleh Tsyvinski, "The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model," *The Review of Economic Studies*, 2017, *84* (2), 613–649.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins, "Double/debiased machine learning for treatment and structural parameters," *The Econometrics Journal*, 01 2018, *21* (1), C1–C68.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig, "Inference with Arbitrary Clustering," 2019.

Conquest, Robert, *The Harvest of Sorrow: Soviet Collectivization and the Terror-Famine*, Oxford, UK: Oxford University Press, 1986.

Danilov, Victor, Robert Manning, and Lynne Viola, eds, *Tragediya Sovetskoj Derevni. Kollektivizatsiya i raskulachivanie. Dokumenti i materialy v 5 tomakh, 1927-1939*, Rosspen, 1999-2006.

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

Davies, Robert W. and Stephen G. Wheatcroft, *The Years of Hunger: Soviet Agriculture, 1931-1933. Second edition.*, Vol. 5 of *The Industrialization of Soviet Russia*, Basingstoke, UK: Palgrave Macmillan, 2009.

DellaVigna, Stefano, Ruben Enikolopov, Vera Mironova, Maria Petrova, and Ekaterina Zhuravskaya, "Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia," *American Economic Journal: Applied Economics*, 2014, *6* (3), 103–132.

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

Egorov, Georgy, Ruben Enikolopov, Alexey Makarin, and Maria Petrova, "Divided We Stay Home: Social Distancing and Ethnic Diversity," NBER Working Papers 27277, National Bureau of Economic Research, Inc May 2020.

Ellman, Michael, "The Role of Leadership Perceptions and of Intent in the Soviet Famine of 1931-1934," *Europe-Asia Studies*, 09 2005, *57* (6), 823–841.

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

Esteban, Joan, Laura Mayoral, and Debraj Ray, "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, June 2012, *102* (4), 1310–42.

Finkel, E. and S. Gehlbach, *Reform and Rebellion in Weak States* Elements in Political Economy, Cambridge University Press, 2020.

Gokmen, Gunes and Dmitrii Kofanov, "Career incentives in political hierarchy: evidence from Imperial Russia," *European Review of Economic History*, 2020, *24(2)*, 264–287.

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

Grosfeld, Irena, Alexander Rodnyansky, and Ekaterina Zhuravskaya, "Persistent Antimarket Culture: A Legacy of the Pale of Settlement after the Holocaust," *American Economic Journal: Economic Policy*, August 2013, *5* (3), 189–226.

_ , Seyhun Orcan Sakalli, and Ekaterina Zhuravskaya, "Middleman Minorities and Ethnic Violence: Anti-Jewish Pogroms in the Russian Empire," *The Review of Economic Studies*, 01 2020, *87* (1), 289–342.

Horowitz, Donald L., *Ethnic groups in conflict*, Berkeley: University of California Press, 1985.

Khrushchev, N.S., *Nikita S. Khrushchev: Vremia, Liudi, Vlast* 1970.

Kondrashin, Viktor, *Golod 1932-1933 godov: Tragediya Sovetskoj derevni*, Moscow: Rosspen, 2008.

Korovkin, Vasily and Alexey Makarin, "Production Networks and War: Evidence from Ukraine," *SSRN Working paper*, 2022.

_ and _ , "Conflict and Intergroup Trade: Evidence from the 2014 Russia-Ukraine Crisis," *American Economic Review*, January 2023, *113* (1), 34–70.

Kotkin, Stephen, *Stalin. Waiting for Hitler, 1929-1941*, New York: Penguin Press, 2017.

Levchuk, Nataliia, Oleh Wolowyna, Omelian Rudnytskyi, Alla Kovbasiuk, and Natalia Kulyk, "Regional 1932-1933 Famine Losses: A Comparative Analysis of Ukraine and Russia," *Nationalities Papers*, 2020, *48* (3), 492–512.

Li, Wei and Dennis Tao Yang, "The Great Leap Forward: Anatomy of a Central Planning Disaster," *Journal of Political Economy*, 2005, *113* (4), 840–877.

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

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

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

Mace, James E., "Soviet Man-Made Famine in Ukraine," in William S. Parsons Totten Samuel and Israel W. Charny, eds., *Century of Genocide: Critical Essays and Eyewitness Accounts.*, Routledge, 2004.

Markevich, Andrei, "A Regional Perspective on the Economic Development of the late Russian Empire," *Working paper*, 2019.

_ and Ekaterina Zhuravskaya, "The Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire," *American Economic Review*, April 2018, *108* (4-5), 1074–1117.

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

Matsuura, Kenji and Cort Willmott, "Terrestrial Air Temperature and Precipitation: 1900-2014 Gridded Monthly Time Series (Version 4.01)," *National Oceanic and Atmospheric Administration*, 2014.

Meng, Xin, Nancy Qian, and Pierre Yared, "The Institutional Causes of China's Great Famine, 1959-1961," *The Review of Economic Studies*, 2015, *82* (4), 1568–1611.

Meslé, France, Jacques Vallin, and Evgeny Andreev, "Demographic Consequences of the Great Famine: Then and Now," in Andrea Graziosi, Lubomyr A. Hajda, and Halyna Hryn, eds., *After the Holodomor: The Enduring Impact of the Great Famine on Ukraine*, Harvard University Press, 2013.

Montalvo, Josz G. and Marta Reynal-Querol, "Ethnic Polarization, Potential Conflict, and Civil Wars," *American Economic Review*, June 2005, *95* (3), 796–816.

Murin, Yuri, ed., *Pisatel' i vozhd': perepiska M.A. Sholokhova s I.V. Stalinym 1931-1950 gody : sbornik dokumentov iz lichnogo arkhiva I.V. Stalina*, Raritet, 1997.

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

Ó Gráda, Cormac, *Famine: A Short History*, Princeton, NJ: Princeton University Press, 2009.

Petrova, Maria and David Yanagizawa-Drott, "Media Persuasion, Ethnic Hatred, and Mass Violence: a Brief Overview of Recent Research Advances," in Charles Anderton and Jurgen Brauer, eds., *Economics of Genocide, Mass Killing, and Their Prevention*, Oxford: Oxford University Press, 2016, chapter 12.

Polyakov, Yu.A. and V.B. Zhiromskaya, *Naselenie Rossii v XX veke: istoricheskie ocherky. Vol. 1 1900-1939*, Rosspen, 2000.

Polyan, Pavel, *Ne po svoej vole... Istoriya i geografiya prinuditelnikh migratsij v SSSR*, Moscow: O.G.I, 2001.

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

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

Rudnytskyi, Omelian, Natalia Levchuk, Oleh Wolowyna, Pavlo Shevchuk, and Alla Kovbasiuk, "Demography of a man-made human catastrophe: The case of massive famine in Ukraine 1932-1933," *Canadian Studies in Population*, 2015, *42* (1-2), 53–80.

Sen, Amartya, *Poverty and Famines. An Essay on Entitlement and Deprivation*, Oxford, UK: Oxford University Press, 1981.

Stalin, Iosif, "Concerning the National Question in Yugoslavia," *Bol'shevik*, 15 April 1925, *7*.

State Duma, "Russian: Postanovleniye GD FS RF ot 02.04.2008 N 262-5 GD O Zayavlenii Gosudarstvennoy Dumy Federal'nogo Sobraniya Rossiyskoy Federatsii "Pamyati zhertv goloda 30-kh godov na territorii SSSR"," 2008.

Tauger, Mark B., "The 1932 Harvest and the Famine of 1933," *Slavic Review*, 1991, *50* (1), 70–89.

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

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

Wheatcroft, S.G., "Soviet and Chinese Famines in historical perspective," in Mathias Middell and Felix Wemheuer, eds., *Hunger and Scarcity under Socialist Rule*, Leipzig: Leipzig University Press, 2012, pp. 307–342.

Wheatcroft, Stephen. G., "The Turn Away from Economic Explanations for Soviet Famines. Roundtable on Soviet Famines," *Contemporary European History*, 2018, *27* (3), 465–469.

Zelenin, I. E., N. A. Ivnitskii, V. V. Kondrashin, and E. N.. Oskolkov, *O golode 1932-1933 godov i ego otsekne na Ukraine*, Otechestvennaya istoriya 6, 1994.

Table 1: Food Availability

	1927	1928	1929	1930	1931	1932	1933	1937	1939
	I. Ukraine								
(1) Total population (mln)	29.0	29.6	30.3	30.8	31.3	31.7	31.9	28.4	29.6
(2) Rural population (mln)	23.6	24.6	24.9	25.1	25.0	24.8	25.0	18.8	18.7
(3) Production (mln tons)	.	14.9	18.7	23.2	16.8	9.1	17.7	.	.
(4) Procurement (mln tons)	4.0	2.0	5.3	7.7	7.3	4.2	6.1	.	.
(5) Rural retention (kg/person/day)	.	1.44	1.48	1.69	1.04	0.54	1.28	.	.
(6) Rural Retention, no procurement (kg/person/day)	.	1.66	2.06	2.53	1.84	1.01	1.95	.	.
(7) Rural and urban retention, no procurement (kg/person/day)	.	1.38	1.70	2.06	1.47	0.79	1.52	.	.
(8) Food needs for heavy labor (kg/person/day)	0.78	0.79	0.78	0.78	0.78	0.78	0.78	0.76	0.75
	II. USSR – no Ukraine								
(9) Total population (mln)	118.0	120.8	124.0	126.7	129.3	131.5	133.8	133.7	135.9
(10) Production (mln tons)	.	58.5	53.1	60.1	47.8	48.0	57.3	.	.
(11) Rural and urban retention, no procurement (kg/person/day)	.	1.33	1.17	1.30	1.01	1.00	1.17	.	.
(12) Food needs for heavy labor (kg/person/day)	0.78	0.78	0.78	0.78	0.78	0.77	0.77	0.76	0.76

Notes: Data for population and procurement are official statistics. Production is revised by the authors using archival sources. See the text and Appendix for details. Retention is the difference between production and procurement. Food needs are calculated by the authors and take into account the demographic composition (e.g., age, gender, rural/urban) as reported in the population censuses. They are based on official guidelines for maximum caloric needs for each group as reported by Lositskij (1926, 1928). Panel I includes Ukraine. Panel II includes all other republics in the USSR.

Table 2: Famine Mortality and Ethnic Ukrainian Population Share – Province-level Estimates

	Dependent Variable: Mortality in Year t+1									
	Control for Corrected Grain Production (1)	Baseline (2)	(3)	(4)	(5)	Omit Province FE (6)	Omit Ukraine (7)	Omit Ukraine, Lower Volga, North Caucasus, West Siberia (8)	Use 1926 Census Population Measures (9)	Other Minorities (10)
Dep. var. mean.	*0.022*	*0.022*	*0.022*	*0.022*	*0.022*	*0.022*	*0.022*	*0.022*	*0.023*	*0.022*
Dep. var. mean. in 1932	*0.031*	*0.031*	*0.031*	*0.031*	*0.031*	*0.031*	*0.029*	*0.026*	*0.033*	*0.031*
Ukrainians × Famine [1]	0.050***	0.051***	0.054***	0.053***	0.050***	0.051***	0.086***	0.059***	0.056***	0.049***
	(0.002)	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.007)	(0.006)	(0.005)	(0.007)
Standardized Coef.	*0.814*	*0.826*	*0.870*	*0.858*	*0.819*	*0.831*	*0.840*	*0.410*	*0.968*	*0.790*
Ukrainians [2]						-0.007***				
						(0.002)				
Other Minorities × Famine										-0.007*
										(0.004)
Grain	0.001	0.0002	0.0002	0.0002	0.0002	-0.0002	0.0003**	0.0007***	0.0001	0.0002
	(0.001)	(0.0001)	(0.0002)	(0.0002)	(0.0001)	(0.0004)	(0.0001)	(0.0002)	(0.0002)	(0.0001)
Grain × Famine	0.007**	0.0004	-0.003	-0.001	0.0004	-0.0001	-0.003	-0.002	-0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.001)	(0.002)	(0.003)
Standardized Coef.	*0.275*	*0.023*	*-0.137*	*-0.067*	*0.021*	*-0.005*	*-0.183*	*-0.113*	*-0.056*	*0.031*
Kulaks × Famine			1.805							
			(1.367)							
Livestock Change × Famine				0.016						
				(0.015)						
Mortality 1928 × Famine					-0.087					
					(0.485)					
Urbanization	0.016	-0.005	-0.006	-0.006	-0.005	-0.011**	-0.004	-0.004		-0.006
	(0.020)	(0.007)	(0.007)	(0.007)	(0.007)	(0.005)	(0.007)	(0.007)		(0.007)
Urbanization × Famine	0.011	0.003	-0.004	0.002	0.003	0.006	-0.003	-0.005	-0.005	0.000
	(0.009)	(0.007)	(0.006)	(0.006)	(0.007)	(0.008)	(0.007)	(0.005)	(0.005)	(0.007)
Observations	107	337	337	337	337	337	319	268	337	337
R-squared	0.838	0.776	0.791	0.784	0.777	0.420	0.758	0.808	0.765	0.780
[1] + [2]: Coef.						0.044				
p-val.						<0.001				

Notes: The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. Mortality is the number of deaths in each year and province divided by the total population in each year and province in columns (1)-(8) and (10); and the number of deaths in each year and province divided by the province's 1926 population in column (9). Ukrainians is the share of ethnic Ukrainians in the rural population. Famine is an indicator that equals one in 1932 and zero otherwise. Column (1) controls for grain production estimates revised by the authors using archival data (see text). Columns (2)-(10) control for per capita grain production predicted by exogenous factors. Urbanization is the urban population share. Column (3) controls for the number of kulak households exiled during 1930-31 per 1930 population interacted with the famine indicator. Column (4) controls for the drop in livestock (horses and cattle) per capita between 1929 and 1931 interacted with the famine indicator. Column (5) controls for the 1928 mortality rate interacted with the famine indicator. Column (9) controls for predicted grain per 1926 population and 1926 urban population share. In column (10), "other minorities" include all ethnic groups except Russians and Ukrainians, so the omitted category is only Russians. All regressions control for province and year FE. The standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table 3: Famine Severity and Ethnic Ukrainian Population Share— Robustness to Alternative Measures and Additional Controls

	Dep. var. mean.	Ukrainians × Famine			Observations	R-squared
		Coef.	Std. Err.	Std. Coef.		
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable: Natality in Year t+1						
(1) Baseline	0.039	-0.014***	(0.003)	*-0.432*	337	0.820
Dependent Variable: Birth Cohort Size in Year t+1/1939 Population						
(2) Baseline	0.024	-0.009***	(0.001)	*-0.747*	1,296	0.768
Dependent Variable: Mortality in Year t+1						
(3) Baseline	0.022	0.051***	(0.006)	*0.826*	337	0.776
A. Control for Weather instead of Predicted Grain						
(4) Spring, Summer, Winter Temp and Rain	0.022	0.051***	(0.004)	*0.830*	337	0.782
(5) Monthly Temp, Rain, and Quadratics (48 vars)	0.022	0.051***	(0.003)	*0.829*	337	0.828
(6) Monthly Temp, Rain, Temp × Rain (36 vars)	0.022	0.048***	(0.003)	*0.785*	337	0.814
(7) Monthly Weather Shock (12 vars)	0.022	0.050***	(0.004)	*0.818*	337	0.784
(8) vars)	0.022	0.048***	(0.003)	*0.781*	337	0.806
(9) Monthly Temp and Rain Deviations for Year t-1, t (48 vars)	0.022	0.047***	(0.002)	*0.772*	337	0.825
B. Alternative Measures of Mortality, Ukrainian Share						
(10) Urban Mortality	0.021	0.013	(0.008)	*0.252*	285	0.746
(11) Rural Mortality	0.022	0.063***	(0.005)	*0.869*	285	0.787
(12) Total Ukrainians	0.022	0.055***	(0.006)	*0.820*	337	0.776
(13) Urban Ukrainians	0.022	0.090***	(0.010)	*0.785*	337	0.771
(14) Mother Tongue Ukrainian 1926	0.022	0.056***	(0.008)	*0.747*	337	0.765
(15) Mother Tongue Ukrainian 1897	0.022	0.058***	(0.007)	*0.820*	337	0.781
C. Control for Demographic Structure, Suitability for Other Crops						
(16) Gender Ratio × Famine, Share of Children Under 10 × Famine	0.022	0.048***	(0.007)	*0.781*	337	0.783
(17) Latitude × Longitude × Famine	0.022	0.059***	(0.004)	*0.965*	337	0.806
(18) Famine × Suitability for: Potato, Sugarbeet, Sunflower, Wheat, Rye, Flax	0.022	0.056***	(0.004)	*0.913*	337	0.806
D. Control for Slow-Moving Economic, Cultural and Institutional Features						
(19) 1892 Famine	0.032	-0.0002	(0.003)	*-0.004*	1,297	0.864
(20) Nominal Income 1897 × Famine	0.022	0.052***	(0.006)	*0.845*	337	0.780
(21) Real Income 1897 × Famine	0.022	0.051***	(0.006)	*0.828*	337	0.776
(22) Labor Productivity 1897 × Famine	0.022	0.051***	(0.006)	*0.823*	337	0.776
(23) Rural Labor Productivity 1897 (Lower Bound) × Famine	0.022	0.050***	(0.005)	*0.811*	337	0.778
(24) Rural Labor Productivity 1897 (Upper Bound) × Famine	0.022	0.050***	(0.005)	*0.809*	337	0.779
(25) Value of Agricultural Equipment 1910 × Famine	0.022	0.046***	(0.004)	*0.747*	337	0.810
(26) Livestock 1916 × Famine	0.022	0.051***	(0.005)	*0.831*	337	0.777
(27) Share of Catholics 1897 × Famine, Share of Orthodox Christians 1897 × Famine	0.022	0.053***	(0.006)	*0.869*	337	0.782
(28) Share of Serfs 1858 × Famine	0.022	0.052***	(0.005)	*0.850*	337	0.779
(29) Peasant Revolts 1895–1914 × Famine	0.022	0.050***	(0.006)	*0.819*	337	0.797
(30) Baseline with Info on Land 1905	0.021	0.042***	(0.005)	*0.690*	286	0.792
(31) Share of Peasant Land in Repartition Commune 1905 × Famine	0.022	0.053***	(0.006)	*0.854*	286	0.799
(32) Share of Peasant Households in Repartition Commune 1905 × Famine	0.022	0.056***	(0.008)	*0.917*	286	0.800
(33) Peasant and Private Land Gini 1905 × Famine	0.022	0.040***	(0.006)	*0.645*	286	0.799
E. Control for Administrative Capacity and Political Zealousness						
(34) 1917 Bolshevik Vote Share × Famine	0.022	0.050***	(0.006)	*0.807*	337	0.781
(35) #Communists (averaged over 1922, 1927, and 1932) × Famine	0.022	0.051***	(0.005)	*0.825*	337	0.776
(36) #1930 Congress Delegates × Famine	0.022	0.050***	(0.006)	*0.815*	337	0.784
(37) LASSO	0.022	0.050***	(0.007)	*0.820*	337	

Notes: The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. Each row is one regression. All regressions control for urbanization, urbanization × famine, and province and year FE. Panels B-E also control for predicted grain and predicted grain × famine. In row (7), a shock is a dummy variable which equals 1 if temp or rain is one std. dev. or more different from historical province mean. The 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 Ethnic Ukrainian Population Share — District-level Estimates

	Dependent Variable: Mortality					
	Baseline with Year FE (1)	Baseline with Province-Year FE (2)	Monthly Temp and Rain Deviations for Year t-1, t (3)	Omit Ukraine (4)	Omit Ukraine, Lower Volga, North Caucasus, West Siberia (5)	Other Minorities (6)
Dep. var. mean.	0.029	0.029	0.029	0.026	0.024	0.029
Ukrainians × Famine	0.034***	0.024***	0.024***	0.024***	0.025***	0.026***
	(0.006)	(0.007)	(0.007)	(0.008)	(0.007)	(0.007)
	[0.006]	[0.007]	[0.007]	[0.008]	[0.011]	[0.007]
Standardized Coef.	0.438	0.311	0.311	0.209	0.214	0.334
Other Minorities × Famine						0.001
						(0.007)
Urbanization	-0.006	-0.003	-0.003	-0.009*	-0.006	-0.002
	(0.006)	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)
	[0.008]	[0.008]	[0.008]	[0.007]	[0.007]	[0.008]
Urbanization × Famine	0.0003	-0.001	-0.001	0.005**	0.007**	-0.001
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.005)
	[0.004]	[0.004]	[0.004]	[0.003]	[0.003]	[0.004]
Observations	3,274	3,274	3,274	2,498	1,996	3,274
R-squared	0.797	0.812	0.812	0.783	0.740	0.813
Ukrainians						
Mean	0.255	0.255	0.255	0.070	0.035	0.255
Std. Dev.	0.374	0.374	0.374	0.171	0.124	0.374

Notes: The sample includes Ukraine and Russia. Observations are at the district and year level. Mortality is the number of deaths divided by total population. Ukrainians is the share of ethnic Ukrainians in the rural population. Famine is an indicator that equals to one in 1933 and zero otherwise. All regressions control for grain suitability × famine. Except column (3), all regressions also control for monthly temperature and precipitation in years t and t-1 (48 additional controls); column (3), instead, controls for deviations in monthly temperature and precipitation from the long-term (1900-1950) means in years t and t-1 (48 additional controls). In column (1), estimates control for district and year FE. In columns (2)-(6) estimates control for district and province × year FE. In column (6), "other minorities" include all ethnic groups except Russians and Ukrainians, so the omitted category is Russians. Grain suitability is the FAO GAEZ wheat suitability index for low-input rain-fed agriculture. The standard errors in parentheses are adjusted for spatial correlation within 400 km. *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the district level are presented in square brackets.

Table 5: Procurement, Retention and Mortality

	Dependent Variable:			
	Procurement Share = Procurement/ Production	Retention = Production – Procurement	Mortality in Year t+1	
	(1)	(2)	(3)	(4)
Dep. var. mean.	*0.224*	*1.019*	*0.022*	*0.022*
Dep. var. mean. in 1932	*0.254*	*0.918*	*0.031*	*0.031*
Ukrainians × Famine	0.185***	-1.085***		
	(0.018)	(0.191)		
Retention			-0.030**	-0.029***
			(0.012)	(0.010)
Retention^2			0.009**	0.009**
			(0.004)	(0.004)
Retention × Famine				-0.015
				(0.027)
Retention^2 × Famine				0.004
				(0.009)
Observations	107	107	107	107
R-squared	0.936	0.782	0.634	0.643

Notes: The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. Production is revised by the authors based on archival documents (see text). Retention is measured in kilograms per person per day. Mortality is the number of deaths divided by the total population. All columns control for urbanization, urbanization × famine, and province and year fixed effects. The standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table 6: Grain Production, Procurement and Retention Targets from the First Five-Year Plan

	Dependent Variables: Centrally Planned Targets (kg/person/day)			
	Production	Procurement	Procurement Share = Procurement/ Production	Retention = Production – Procurement
	(1)	(2)	(3)	(4)
Dep. var. mean.	*1.673*	*0.268*	*0.089*	*1.404*
Ukrainians	0.776***	0.275***	0.165***	-0.275***
	(0.191)	(0.071)	(0.020)	(0.071)
Official Grain Production 1928	0.907***			
	(0.052)			
Production Target		0.540***	0.285***	0.460***
		(0.035)	(0.016)	(0.035)
Observations	90	90	90	90
R-squared	0.691	0.814	0.844	0.723

Notes: The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level for the years of the First Five-Year Plan, 1928-33. Ukrainians is the 1926 share of ethnic Ukrainians in the rural population. Official grain production 1928 is measured in kg/person/day. All estimates control for year fixed effects. Huber-White robust standard errors are in parentheses. *** p<0.01, ** p<0.05, * p<0.1

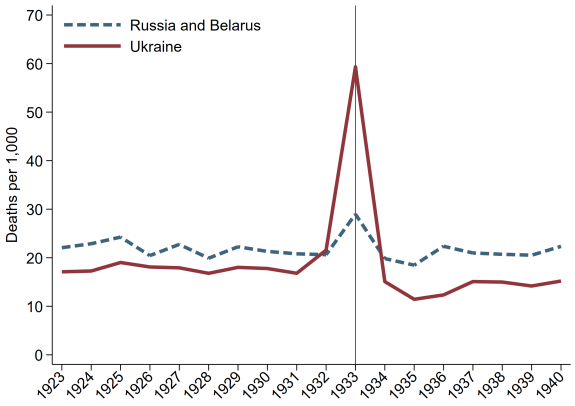
Table 7: Heterogeneous Effects of Grain Productivity on Famine Mortality in Ethnic Ukrainian Areas

	Dependent Variable: Mortality in Year t+1			
	Province- Level			District-Level
	(1)	(2)		(3)
Ukrainians × Official Grain Production 1928 × Famine	0.300***	0.263***	Ukrainians × Grain Suitability × Famine	0.071***
	(0.051)	(0.081)		(0.022)
Ukrainians × Famine	-0.174***	-0.065	Ukrainians × Famine	-0.011
	(0.035)	(0.156)		(0.017)
Official Grain Production 1928 × Famine	0.0002	0.001	Grain Suitability × Famine	0.004
	(0.002)	(0.002)		(0.005)
Ukrainians × Admin Capacity × Famine		0.022		
		(0.027)		
Admin. Capacity × Famine		-0.001		
		(0.001)		
Observations	337	337	Observations	3,376
R-squared	0.845	0.846	R-squared	0.783

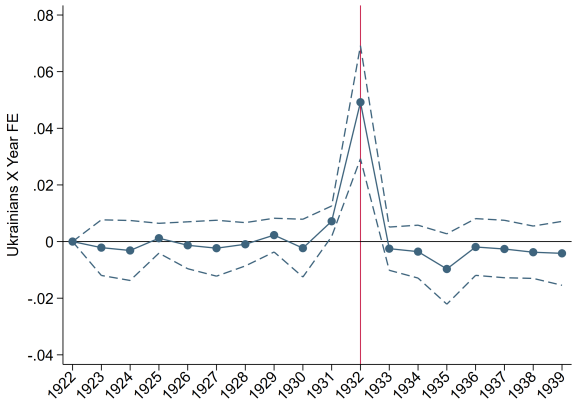
Notes: The province sample includes Ukraine, Russia and Belarus; the distict sample includes Ukraine and Russia. Cols (1) and (2) control for urbanization, urbanization × famine, Ukrainians × urbanization × famine, predicted grain, predicted grain × famine, Ukrainians × predicted grain × famine; and province and year fixed effects. In col (2), Admin. Capacity is the first principal component of the share of votes for the Bolshevik Party in the 1917 Constituency Assembly election, the number of Communist Party Members per 1,000 individuals in each province, and the number of delegates at the 1930 Party Congress. In cols (1) and (2), the standard errors in parentheses are adjusted for spatial correlation within 1,500 km. Col (3) controls for urbanization, urbanization × famine, Ukrainians × urbanization × famine, and district and province-year fixed effects. The standard errors in parentheses are adjusted for spatial correlation within 400 km. *** p<0.01, ** p<0.05, * p<0.1

Figure 1: Famine over Time

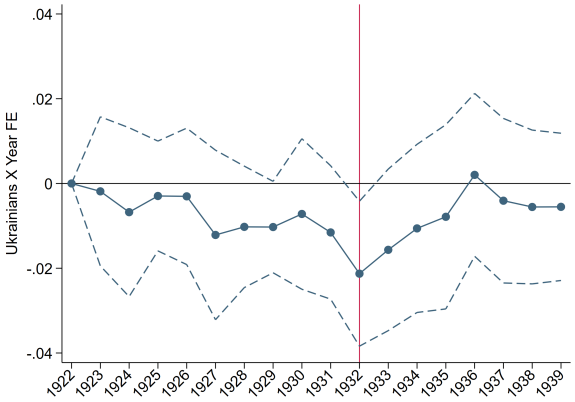
(a) Raw Mortality



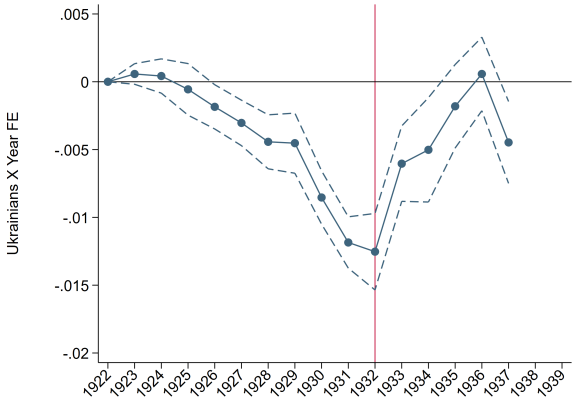
(b) Coefs. of *Ukrainians* × *Year FEs* on Mortality



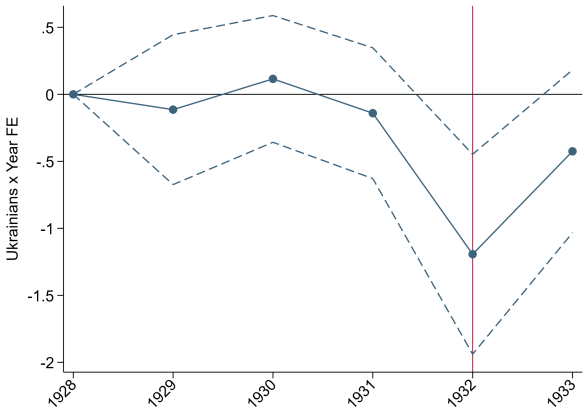
(c) Coefs. of *Ukrainians* × *Year FEs* on Natality



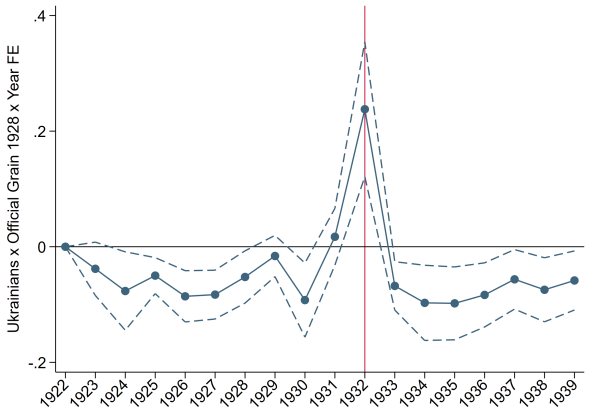
(d) Coefs. of *Ukrainians* × *Year FEs* on Birth Cohort Size



(e) Coefs. of *Ukrainians* × *Year FEs* on Retention



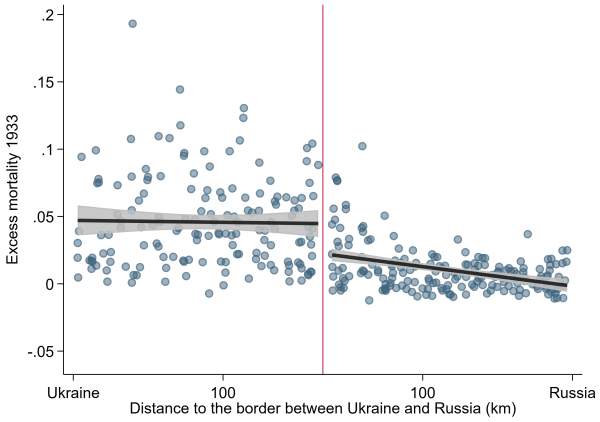
(f) Coefs. of *Ukrainians* × 1928 *Prod* × *Year FEs* on Mortality



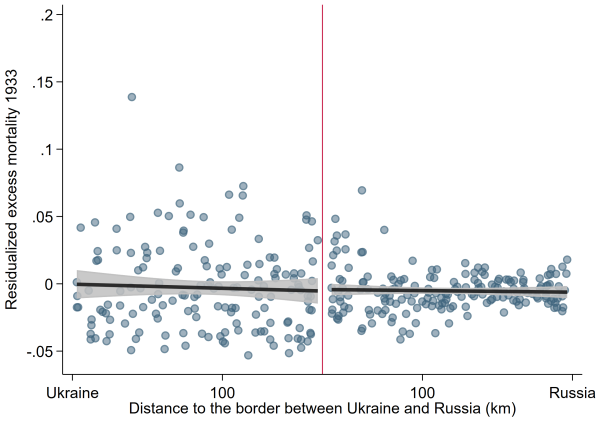
Notes: Mortality is the number of deaths per 1,000 individuals. Natality is the number of live births per 1,000 individuals. Birth cohort size is the number of individuals born each year divided by the province population as reported by the 1939 census. Retention is kgs grain retained per person per day. Figures 1b, 1c, 1d, 1e plot the interaction coefficients of Ukrainian population share and year dummy variables with their 95% confidence intervals. Figure 1f plots the triple interaction coefficient of Ukrainian population share, year dummy variables, and the 1928 grain production, and their 95% confidence intervals. The coefficients and their standard errors are presented in Appendix Tables A.5 and A.5.

Figure 2: Famine Mortality at the Ukrainian-Russian Border

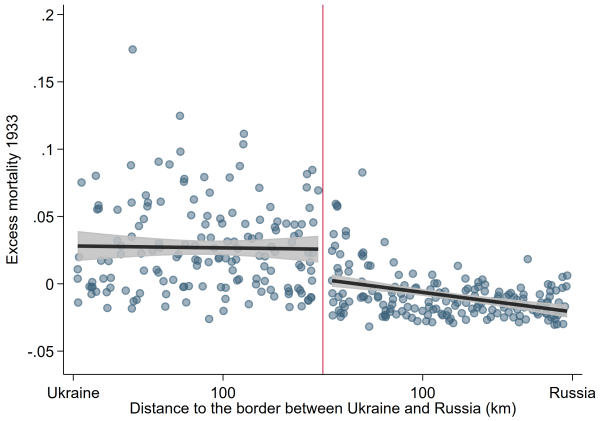
(a) Raw Mortality



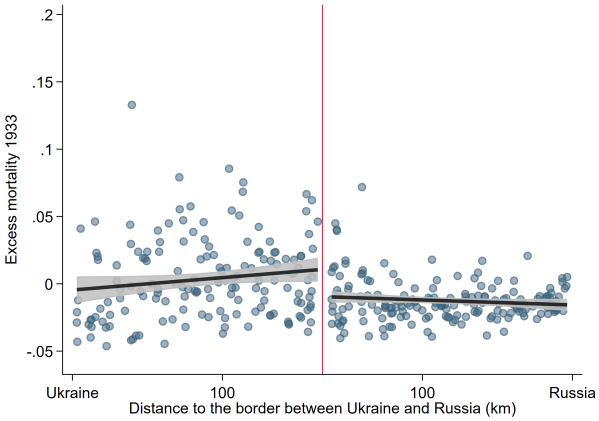
(b) Demeaned by 1926 Ukrainian Population Share



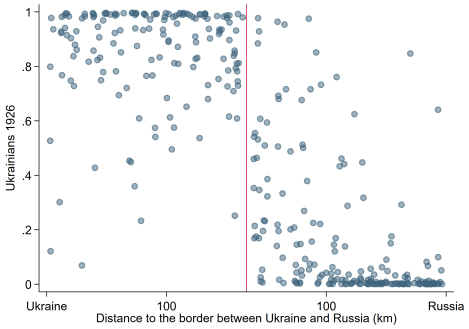
(c) Demeaned by 1926 Urbanization



(d) Demeaned by Agricultural Suitability, Weather



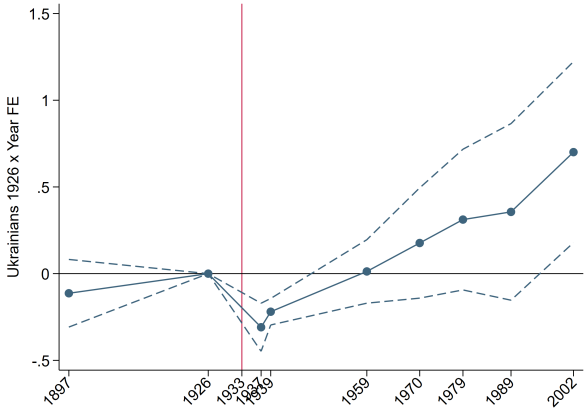
(e) Ukrainian Population Share



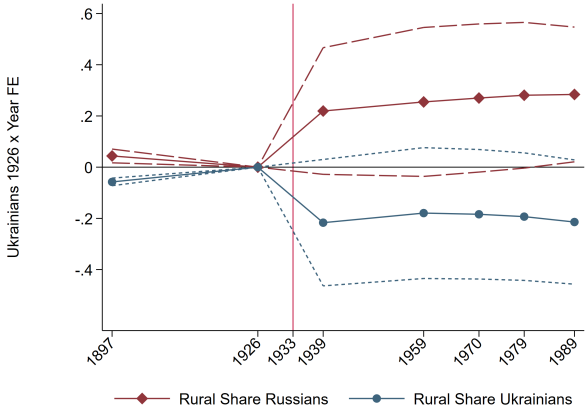
Notes: The figures plot excess mortality against distance to the Ukraine-Russia border and the fitted lines and their 95% confidence intervals for each district in Ukraine and Russia. Excess mortality is the difference between 1933 and 1928 mortality for each district. In Figure 2b, excess mortality is demeaned by the share of ethnic Ukrainians in each district. In Figure 2c, excess mortality is demeaned by the 1926 urban population share in each district. In Figure 2d, excess mortality is demeaned by grain suitability and monthly 1933 temperature and precipitation (24 controls). The estimated magnitude of the border effect is presented in Appendix Table A.8.

Figure 3: Rural Population, Grain Production, Mechanization After the Famine

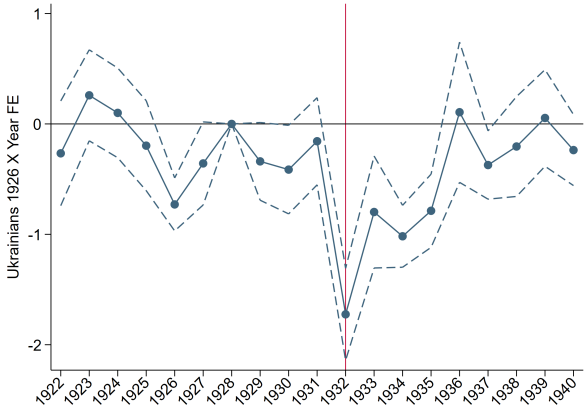
(a) Log Rural Population



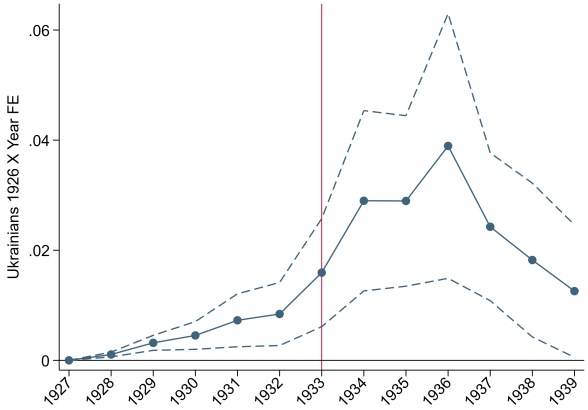
(b) Rural Ethnic Ukrainian and Russian Population Shares



(c) Log Grain Production



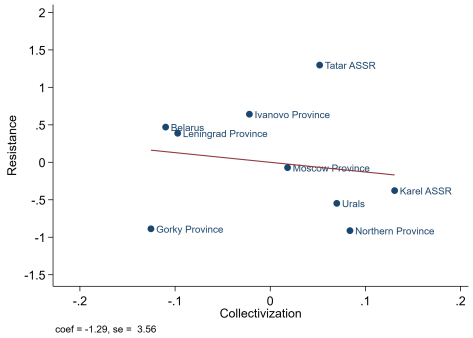
(d) Tractors (Tractor Horse Power per Hectare of 1928 Sown Area)



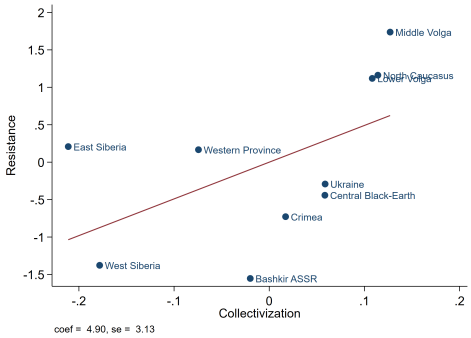
Notes: All figures plot the interaction coefficients of 1926 Ukrainian population share and year dummy variables in a regression that controls for province and year fixed effects. Figure 3c also controls for grain suitability interacted with year fixed effects and monthly temperature and precipitation for years t and $t-1$. Figure 3d also controls for grain suitability interacted with year fixed effects. Appendix Tables A.10 and A.11 report the coefficients and their standard errors. Figures 3a and 3b use population censuses. See the text for data sources of other figures.

Figure 4: Peasant Resistance to Collectivization

(a) Low Ukrainian Share



(b) High Ukrainian Share



Notes: Figures 4a and 4b plot the fitted lines and residuals for the correlation between collectivization and peasant resistance, controlling for official 1928 per capita grain production and urbanization. Figure 4a includes provinces with Ukrainian share below the sample median. Figure 4b includes provinces with Ukrainian share above or equal to the sample median. The p-value for the statistical difference of the two slopes is 0.093.

Online Appendix (Not for Publication)

A Revising Grain Production

Appendix Table A.2 presents the aggregate food counting exercise for Russia and the entire Soviet Union that are analogous to what is shown for Ukraine in Table 1. Panel I presents data for Russia. Panel II for all of the Soviet Union. We also compare our production corrections to those from Davies and Wheatcroft (2009), which are only available for the aggregate level and only as a range (henceforth, DW range).[2]

For brevity, we focus our discussion on the different corrections in Panel II. We approximate "true" production by using grain procurement ratios (as a share of production) reported by the State Planning Committee (Gosplan), and procurement levels, which are widely accepted as accurate since they were only ever used internally. Procurement stocks were directly observed and counted by the government.

Similarly, the accuracy of procurement ratios reported in the Gosplan has not been disputed. The Gosplan report was a candid evaluation of the First Five-Year Plan shown only to the highest ranking Soviet officials. The report was classified until after the fall of the Soviet Union. Procurement ratios are reported for each province and year from 1928 to 1932 in RSAE 4372/30/871 (p. 25). Dividing procurement levels by procurement ratios recovers production levels.

The main advantages of our correction are that it can be calculated at regional levels and require few, straightforward assumptions. Our correction assumes that a decline in procurement reflects a decline in production rather than a decline in the ability of the government to collect. If, for example, peasants hid more (less) grain to evade procurement in 1932 than in other years, then our estimates would understate (overstate) true production. Given historical accounts that intense campaigns to prevent peasants from hiding grain started with the introduction of collectivization several years before the famine, the *prima facie* assumption is that there was little evasion (and little change in evasion) between 1930 and 1932. In other analyses of the paper, we address the measurement error in production figures by not using reported production during this period.

The Gosplan report does not cover 1933. Thus, to estimate production for 1933, we follow the spirit of Davies and Wheatcroft (2009) and calculate 1933 production as collectivization rate $\times$ grain sown area $\times$ kolkhoz yield + (1 - collectivization rate) $\times$ grain sown area $\times$ official yield. 1933 kolkhoz yields are reported in the Gosplan report on the state of collective farms in 1932–33 (RSAE 1562/77/70, p. 31). Collectivization rate is the share of collectivized rural households. This calculation assumes that sown area was similar for collectivized and non-collectivized households. In reality, collectivized households were given more land to farm. It also assumes that individual farm yields were equal to official yields.

Panel II rows (6) and (8)-(9) show that, except for 1930, our corrections fall within the DW range for years for which both are available. Given the conventional wisdom that official aggregate numbers were inflated only in the years when production fell relative to expectations, it is reassuring that our corrected production is below officially reported production levels in row (7) for 1931 to 1933.

[2]The derivation of these estimates, which are presented in Davies and Wheatcroft (2009) Table 1, p. 448-449, is unclear. The authors explain that "Our estimates are based on a range of different data that were accepted internally by the best experts of the time, and our own assessments of the reliability of these different data" (p. 446). See their earlier work for more discussion.

Note that after 1933, the government switched from reporting realized production to the maximum possible yields given weather, geographic conditions, sown area, labor and other inputs. This change did not systematically differ based on the share of ethnic Ukrainians in each region and should not affect our study.[3]

B Predicting Grain Production

To estimate the grain production function, we use data from 1901 to 1915. We regress log grain production (total harvests) on log province area, log FAO GAEZ grain suitability index, their interaction, temperature and precipitation for each of the four seasons, their pairwise interactions and squared 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), and summer (July, August, September). The weather data are from Matsuura and Willmott (2014). Appendix Table A.3 presents 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, with two exceptions: Karelia and East Siberia provinces. The in-sample R-squared is 0.90. The out-of-sample R-squared is 0.77, see Appendix Figure A.3. The high out-of-sample predictive power is consistent with the lack of major technological changes in Soviet agriculture before the 1930s.

C *Dekulakization*

Appendix Table A.6 Panel A 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 in Davies and Wheatcroft (2009, Table 28), the number of exiled *kulak* households during 1930–31 according to a secret police report in Berelowitch and Danilov, eds (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 Natality and Birth Cohort Sizes

Appendix Figures A.1c and A.1d plot natality rates and birth cohort sizes over time. The temporal patterns of the raw data are less sharp, but consistent with mortality. Average natality rates and birth cohort sizes begin declining around 1928 and reach the lowest levels in 1933 and 1934. They decline before mortality rates began to spike, and took longer to recover back to trend than mortality rates. This is consistent with the fact that when food availability is low, food is generally allocated towards sustaining the existing population over giving birth, and that famine survivors need to recover their health before they can give birth. Note that the interpretation of natality and birth cohort sizes differ slightly. Lower natality changes reflect fewer births, while smaller birth cohorts reflect fewer births or lower survival rates.

[3] Davies and Wheatcroft (2009) discusses this change from "the barn harvest" to "the harvest on the root".

There is substantial cross-province variation in natality rates. Appendix Figure A.1b plots mean natality and the cross-province normalized standard deviation in natality for the full sample. It shows that the sharp rise in average famine mortality coincides with a sharp rise in the variance.

Table A.1: Ethnic Composition in the USSR

	1926 census						1939 census	
	Total		Urban		Rural		Total	
	mln.	%	mln.	%	mln.	%	mln.	%
	A. All 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
	B. USSR (Regression Sample, Subset of Panel A)							
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
	C. "Grain-producing" Provinces (Subset of Panel B)							
Ukrainians	28.5	43.8	3.0	29.2	25.5	46.6	25.3	37.1
Russians	27.3	41.9	5.2	50.0	22.0	40.3	32.8	48.1
Tatars	2.2	3.4	0.2	2.0	2.0	3.6	2.9	4.3

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

Table A.2: Caloric Accounting

		1927	1928	1929	1930	1931	1932	1933	1937	1939
		I. Russia								
(1)	Production (mln tons)	.	60.5	48.2	54.8	43.0	44.3	50.8	.	.
(2)	Procurement (mln tons)	6.2	9.6	10.2	13.9	14.8	13.7	15.3	.	.
(3)	Rural retention (kg/person/day)	.	1.63	1.19	1.27	0.87	0.96	1.12	.	.
(4)	Retention if no procurement	.	1.94	1.51	1.70	1.33	1.39	1.61	.	.
(5)	Food needs for heavy labor (kg/person/day)	0.78	0.78	0.78	0.78	0.78	0.78	0.79	0.75	0.75
		II. USSR								
(6)	Production (mln tons)	.	73.4	71.8	83.3	64.5	57.1	75.1	.	.
(7)	Official	74.1	73.3	71.7	83.5	69.5	69.9	89.8	120.3	100.9
(8)	DW Min	.	.	.	73.0	57.0	55.0	70.0	.	.
(9)	DW Max	.	.	.	77.0	65.0	60.0	77.0	.	.
(10)	Procurement (mln tons)	11.1	10.8	16.1	22.2	22.8	19.0	23.7	31.9	30.9
(11)	Rural retention (kg/person/day)	.	1.38	1.21	1.31	0.89	0.82	1.11	.	.
(12)	Retention if no procurement	.	1.62	1.55	1.78	1.38	1.23	1.62	.	.
(13)	Food needs for heavy labor (kg/person/day)	0.78	0.78	0.78	0.78	0.78	0.77	0.77	0.76	0.76

Notes: Data for population, production and procurement are official statistics. Rows (1) and (6) report production revised by the authors using archival data (see text). Row (7) reports official production figures. Rows (8)-(9) report production revised by Davis and Wheatcroft (2004) (see their Table 1). Retention is the difference between our revised production and procurement. Food needs are calculated by the authors and take into account the demographic composition (e.g., age, gender, rural/urban) as reported in the population censuses. They are based on official guidelines for maximum caloric needs for each group as reported by Lositskij (1926, 1928). Panel I includes Russia. Panel II includes all USSR.

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

	Dependent Variable: Log Grain Production		
	(1)		(2)
Log area	0.352***	Fall temperature × Fall precipitation	0.0005*
	(0.067)		(0.0002)
Log grain suitability	-4.643***	Winter temperature × Winter precipitation	0.001*
	(0.640)		(0.0003)
Log area × Log grain suitability	0.278***	Spring temperature × Spring precipitation	0.0004
	(0.023)		(0.0003)
Fall temperature	0.015	Summer temperature × Summer precipitation	0.001***
	(0.037)		(0.0003)
Winter temperature	0.027	Fall temperature^2	0.004**
	(0.043)		(0.002)
Spring temperature	-0.169**	Winter temperature^2	0.0003
	(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	0.00002
	(0.007)		(0.00002)
Spring precipitation	0.01	Winter precipitation^2	0.00004
	(0.007)		(0.00003)
Summer precipitation	-0.025***	Spring precipitation^2	-0.00003**
	(0.009)		(0.00002)
		Summer precipitation^2	0.00003**
			(0.00001)
Observations		220	
R-squared		0.998	

Notes: The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. Log grain is the logarithm of the grain harvest. Log area is the logarithm of province area. Log grain suitability is the logarithm of the province's FAO GAEZ wheat 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. *** p<0.01, ** p<0.05, * p<0.1

Table A.4: Standard Errors of the Baseline Estimate of Ukrainians × Famine

Time correlation, years:	Spatial correlation, km: 900 (1)	1,000 (2)	1,100 (3)	1,200 (4)	1,300 (5)	1,400 (6)	Baseline 1,500 (7)	1,600 (8)	1,700 (9)	1,800 (10)	1,900 (11)	2,000 (12)
	A. Correct for heteroskedasticity and autocorrelation + imposes a distance linear decay between observations within the cutoff											
Baseline: 0	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
1	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
2	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
3	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
4	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
5	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
6	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
7	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
8	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
9	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
10	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
11	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
12	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
13	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
14	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
15	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
16	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
17	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
18	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
19	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
20	0.007	0.007	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006	0.006
	B. Correct for heteroskedasticity and autocorrelation											
0	0.002	0.004	0.004	0.004	0.005	0.005	0.005	0.005	0.005	0.005	0.005	0.005
1	0.003	0.005	0.004	0.004	0.005	0.005	0.006	0.006	0.005	0.006	0.005	0.005
2	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
3	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
4	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
5	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
6	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
7	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
8	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
9	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
10	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
11	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
12	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
13	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.005	0.006	0.005	0.005
14	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.005
15	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.005
16	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.005
17	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.005
18	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.005
19	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.005
20	0.003	0.005	0.005	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.005	0.005

Notes: The table reports standard errors of the Ukrainians × Famine coefficient from the baseline estimate (Table 2 column 2) for different levels of spatial and over time correlation. See Colella et al. (2019) for details.

Table A.5: The Dynamic Relationship between Ukrainian Population Share, Famine Intensity and Grain Retention

	Dependent Variable:							
	Mortality in Year t+1	Rural Mortality in Year t+1	Natality in Year t+1	Rural Natality in Year t+1	Total Birth Cohort Size in Year t+1/Total 1939 Population	Rural Birth Cohort Size in Year t+1/Rural 1939 Population		Mortality in Year t+1
	(1)	(2)	(3)	(4)	(5)	(6)		(7)
Ukrainians × 1923	-0.002		-0.002		0.001	0.001*	Ukrainians × Official Grain 1928 × 1923	-0.038
	(0.003)		(0.007)		(0.0004)	(0.0005)		(0.024)
Ukrainians × 1924	-0.003		-0.007		0.0004	0.001	Ukrainians × Official Grain 1928 × 1924	-0.077**
	(0.003)		(0.007)		(0.001)	(0.001)		(0.035)
Ukrainians × 1925	0.001		-0.003		-0.001	-0.0002	Ukrainians × Official Grain 1928 × 1925	-0.050***
	(0.002)		(0.005)		(0.001)	(0.001)		(0.016)
Ukrainians × 1926	-0.001	-0.003	-0.003	-0.0002	-0.002**	-0.002*	Ukrainians × Official Grain 1928 × 1926	-0.086***
	(0.003)	(0.002)	(0.006)	(0.002)	(0.001)	(0.001)		(0.023)
Ukrainians × 1927	-0.002	-0.004	-0.012	-0.009**	-0.003***	-0.003***	Ukrainians × Official Grain 1928 × 1927	-0.083***
	(0.004)	(0.003)	(0.008)	(0.004)	(0.001)	(0.001)		(0.022)
Ukrainians × 1928	-0.001	-0.003**	-0.010*	-0.007***	-0.004***	-0.005***	Ukrainians × Official Grain 1928 × 1928	-0.052**
	(0.003)	(0.001)	(0.005)	(0.002)	(0.001)	(0.001)		(0.023)
Ukrainians × 1929	0.002	0.001	-0.010***	-0.007***	-0.005***	-0.005***	Ukrainians × Official Grain 1928 × 1929	-0.016
	(0.002)	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)		(0.018)
Ukrainians × 1930	-0.002	-0.003	-0.007	-0.003	-0.009***	-0.011***	Ukrainians × Official Grain 1928 × 1930	-0.092***
	(0.004)	(0.003)	(0.007)	(0.003)	(0.001)	(0.001)		(0.033)
Ukrainians × 1931	0.007***	0.008***	-0.012*	-0.007***	-0.012***	-0.015***	Ukrainians × Official Grain 1928 × 1931	0.017
	(0.002)	(0.001)	(0.006)	(0.002)	(0.001)	(0.001)		(0.025)
Ukrainians × 1932	0.049***	0.059***	-0.021***	-0.017***	-0.013***	-0.016***	Ukrainians × Official Grain 1928 × 1932	0.238***
	(0.007)	(0.006)	(0.006)	(0.003)	(0.001)	(0.002)		(0.059)
Ukrainians × 1933	-0.003	-0.005***	-0.016**	-0.012***	-0.006***	-0.008***	Ukrainians × Official Grain 1928 × 1933	-0.068***
	(0.003)	(0.002)	(0.007)	(0.004)	(0.001)	(0.002)		(0.021)
Ukrainians × 1934	-0.004	-0.007***	-0.011	-0.007**	-0.005**	-0.007***	Ukrainians × Official Grain 1928 × 1934	-0.097***
	(0.003)	(0.002)	(0.008)	(0.003)	(0.002)	(0.002)		(0.033)
Ukrainians × 1935	-0.010**	-0.013***	-0.008	-0.005	-0.002	-0.003	Ukrainians × Official Grain 1928 × 1935	-0.098***
	(0.005)	(0.004)	(0.009)	(0.004)	(0.002)	(0.002)		(0.032)
Ukrainians × 1936	-0.002	-0.006**	0.002	0.003	0.001	0.001	Ukrainians × Official Grain 1928 × 1936	-0.083***
	(0.003)	(0.002)	(0.007)	(0.003)	(0.001)	(0.001)		(0.028)
Ukrainians × 1937	-0.003	-0.006***	-0.004	-0.002	-0.004***	-0.005***	Ukrainians × Official Grain 1928 × 1937	-0.056**
	(0.003)	(0.002)	(0.008)	(0.003)	(0.002)	(0.002)		(0.026)
Ukrainians × 1938	-0.004	-0.006***	-0.006	-0.003			Ukrainians × Official Grain 1928 × 1938	-0.074***
	(0.003)	(0.002)	(0.007)	(0.003)				(0.028)
Ukrainians × 1939	-0.004	-0.007**	-0.006	-0.002			Ukrainians × Official Grain 1928 × 1939	-0.058**
	(0.004)	(0.003)	(0.007)	(0.003)				(0.026)
Observations	337	285	337	285	1,296	1,296		337
R-squared	0.811	0.819	0.870	0.926	0.822	0.767		0.914

Notes: The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. Mortality is the number of deaths divided by the total population. Natality is the number of live births divided by the total population. Birth cohort size is the total/rural birth cohort size reported by the 1939 Census divided by the total/rural 1939 population. Columns (1)-(4) control for urbanization interacted with year indicators, predicted grain interacted with year indicators, and province and year fixed effects. Columns (5)-(6) control for 1926 urbanization interacted with year indicators, predicted grain interacted with year indicators, and province and year fixed effects. Column (7) controls for urbanization × Ukrainians × year indicators, predicted grain × Ukrainians × year indicators and all lower-order interaction terms; and province and year fixed effects. The 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: Famine Mortality and Ukrainian Population Share – Robustness to *Dekulakization* and Demographic Structure

	Dependent Variable: Mortality in Year t+1					
	Dep. var. mean.	Ukrainians × Famine			Observations	R-squared
		Coef.	Std. Err.	Std. Coef.		
	(1)	(2)	(3)	(4)	(5)	(6)
(1) Baseline	0.022	0.051***	(0.006)	0.826	337	0.776
A. Alternative Controls for *Dekulakization*						
(2) Exiled Kulaks (DW) × Famine	0.022	0.054***	(0.005)	0.876	337	0.792
(3) Exiled Kulaks (OGPU) × Famine	0.022	0.053***	(0.006)	0.862	337	0.789
(4) Planned Kulaks 1930 × Famine	0.022	0.044***	(0.007)	0.712	337	0.807
(5) Total Kulaks (OGPU Estimate) × Famine	0.022	0.058***	(0.007)	0.944	267	0.780
(6) Arrested Kulaks 1930 × Famine	0.022	0.051***	(0.006)	0.836	302	0.787
(7) First Principal Component of Kulak Variables × Famine	0.022	0.053***	(0.005)	0.868	249	0.800
B. Alternative Controls for Demographic Structure						
(8) Share of Infants × Famine, Gender Ratio × Famine	0.022	0.055***	(0.006)	0.901	337	0.780
(9) Share of Children 5 and Younger × Famine, Gender Ratio × Famine	0.022	0.046***	(0.007)	0.756	337	0.785
(10) Share of Children 10 and Younger × Famine, Gender Ratio × Famine	0.022	0.048***	(0.007)	0.781	337	0.783
(11) Share of Adults 50 and Older × Famine, Gender Ratio × Famine	0.022	0.051***	(0.005)	0.833	337	0.784
(12) Share of Adults 70 and Older × Famine, Gender Ratio × Famine	0.022	0.051***	(0.005)	0.837	337	0.784

Notes: The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. Ukrainian population share is the share of self-reported ethnic Ukrainians in the rural population. Famine is an indicator that equals one in 1932 and zero otherwise. Each row reports the interaction of Ukrainians and Famine coefficient from a separate regression. All regressions control for urbanization, urbanization × famine, predicted grain, predicted grain × famine, and province and year fixed effects. Additional controls are stated in the rows. In Panel A row (2), exiled kulaks (DW) are the number of dekulakized and exiled households in 1930–31 per 1930 population according to Davies and Wheatcroft (2004), Table 28. In row (3), exiled kulaks (OGPU) are 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. In row (4), planned kulaks are the planned number of dekulakizations per capita in February 1930 (the average between lower and upper bounds). In row (5), total kulaks (OGPU estimate) are 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). In row (6), arrested kulaks are 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). The 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: Famine Mortality and Ethnic Ukrainian Population Share — Robustness Checks for District-level Estimates

	Dependent Variable: Mortality in 1933								
		Alternative Measures of Mortality		Alternative Measures of Ukrainian Population Share			Additional Controls		
	Baseline with Province-Year FE	Urban Mortality	Rural Mortality	Total Ukrainians	Urban Ukrainians	Mother Tongue Ukrainian	Gender Ratio	Latitude × Longitude × Famine	Famine × Suitability for: Potato, Sugarbeet, Sunflower, Rye, Flax
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dep. var. mean.	*0.029*	*0.024*	*0.031*	*0.029*	*0.029*	*0.029*	*0.029*	*0.029*	*0.029*
Ukrainians × Famine	0.024***	0.0004	0.036***	0.031***	0.027***	0.035***	0.024***	0.023***	0.025***
	(0.007)	(0.007)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)
	[0.007]	[0.014]	[0.011]	[0.007]	[0.012]	[0.007]	[0.007]	[0.007]	[0.007]
Standardized Coef.	*0.311*	*0.007*	*0.441*	*0.384*	*0.276*	*0.444*	*0.311*	*0.294*	*0.324*
Urbanization	-0.003	-0.008	-0.011	-0.006	-0.004	-0.003	-0.002	-0.001	-0.002
	(0.006)	(0.007)	(0.008)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
	[0.008]	[0.010]	[0.011]	[0.008]	[0.009]	[0.008]	[0.008]	[0.008]	[0.008]
Urbanization × Famine	-0.001	0.012***	0.007	0.002	0.001	-0.001	-0.003	-0.001	-0.001
	(0.005)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
	[0.004]	[0.006]	[0.007]	[0.004]	[0.005]	[0.004]	[0.004]	[0.004]	[0.004]
Observations	3,274	1,648	2,832	3,274	1,960	3,264	3,274	3,274	3,274
R-squared	0.812	0.913	0.920	0.814	0.840	0.816	0.813	0.814	0.818
Ukrainians									
Mean	0.255	0.255	0.255	0.243	0.228	0.236	0.255	0.255	0.255
Std. Dev.	0.374	0.374	0.374	0.358	0.294	0.368	0.374	0.374	0.374

Notes: The sample includes Ukraine and Russia. Observations are at the district and year level. All regressions control for grain suitability × famine, monthly temperature and precipitation in years t and t-1 (48 additional controls), district and province × year fixed effects. Column (7) also controls for the male/female ratio × famine. Column (8) also controls for latitude × longitude × famine and all lower-order interactions. Column (9) controls for agricultural suitability of each of the 5 crops listed in the column heading interacted with the famine dummy. Grain suitability is the FAO GAEZ wheat suitability index for low-input rain-fed agriculture. The standard errors in parentheses are adjusted for spatial correlation within 400 km. *** p<0.01, ** p<0.05, * p<0.1. Standard errors clustered at the district level are presented in square brackets.

Table A.8: Famine Mortality at the Ukrainian-Russian Border

	Dependent Variable: Excess Mortality 1933			
Window, km	[-100,100]	[-150,150]	[-200,200]	[-250,250]
	(1)	(2)	(3)	(4)
	A. Raw Mortality			
Russia dummy	-0.025***	-0.032***	-0.036***	-0.035***
	(0.004)	(0.004)	(0.005)	(0.006)
Observations	144	211	274	336
R-squared	0.174	0.258	0.292	0.304
	B. Demeaned by 1926 Ukrainian Share			
Russia dummy	0.001	-0.003	-0.003	-0.002
	(0.006)	(0.005)	(0.005)	(0.004)
Observations	144	211	274	336
R-squared	0.000	0.004	0.005	0.002
	C. Demeaned by 1926 Urbanization			
Russia dummy	-0.025***	-0.032***	-0.036***	-0.035***
	(0.004)	(0.004)	(0.005)	(0.006)
Observations	144	211	274	336
R-squared	0.176	0.261	0.294	0.307
	D. Demeaned by Agricultural Suitability and Weather			
Russia dummy	-0.017***	-0.019***	-0.019***	-0.016***
	(0.004)	(0.003)	(0.003)	(0.004)
Observations	144	211	274	336
R-squared	0.098	0.144	0.139	0.104

Notes: The sample includes Russia and Ukraine. Observations are districts within a the distance from Ukrainian-Russian border stated in the column headings. Excess mortality 1933 is the difference between 1933 and 1928 mortality for each district. In Panel B, excess mortality 1933 is demeaned by the share of ethnic Ukrainians in each district. In Panel C, excess mortality 1933 is demeaned by the 1926 urban population share in each district. In Panel D, excess mortality 1933 is demeaned by grain suitability and 1933 monthly temperature and precipitation (24 controls) in each district. The Russia dummy equals to one on the Russian side of the border. The standard errors in parentheses are adjusted for spatial correlation within 400 km. *** p<0.01, ** p<0.05, * p<0.1

Table A.9: Mortality, Natality and Grain Retention

	Dependent Variable:								
	Mortality in Year t+1			Natality in Year t+1				Retention	
	(1)	(2)	(3)	(4)	(5)	(6)		(7)	(8)	(9)
Dep. var. mean.	*0.022*	*0.022*	*0.022*	*0.038*	*0.038*	*0.038*	*Dep. var. mean.*	*1.019*	*1.019*	*1.019*
Retention: < 0.5	0.037***	0.036***	0.036***	0.045***	0.046***	0.046***	Ukrainians × 1929	-0.114	-0.111	-0.115
	(0.009)	(0.009)	(0.009)	(0.003)	(0.003)	(0.003)		(0.285)	(0.284)	(0.286)
Retention: (0.5, 1.0]	0.026***	0.025***	0.025***	0.045***	0.045***	0.046***	Ukrainians × 1930	0.115	0.119	0.116
	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)		(0.241)	(0.240)	(0.242)
Retention: (1.0, 1.5]	0.020***	0.019***	0.020***	0.049***	0.051***	0.051***	Ukrainians × 1931	-0.141	-0.135	-0.137
	(0.004)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)		(0.249)	(0.247)	(0.249)
Retention: (1.5, 2.0]	0.020***	0.020***	0.021***	0.052***	0.053***	0.053***	Ukrainians × 1932	-1.193***	-1.192***	-1.160***
	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)		(0.380)	(0.377)	(0.352)
Retention: > 2.0	0.016***	0.015***	0.015***	0.056***	0.058***	0.058***	Ukrainians × 1933	-0.425	-0.416	-0.412
	(0.005)	(0.005)	(0.005)	(0.003)	(0.003)	(0.003)		(0.310)	(0.306)	(0.306)
Controls:							Controls:			
Urbanization × Famine		Y	Y		Y	Y	Urbanization × Famine		Y	Y
Official Grain 1928 × Famine			Y			Y	Official Grain 1928 × Famine			Y
Observations	107	107	107	107	107	107	Observations	107	107	107
R-squared	0.964	0.964	0.968	0.995	0.995	0.995	R-squared	0.788	0.790	0.793

Notes: The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. Mortality is the number of deaths divided by the total population. Natality is the number of live births divided by the total population. Retention is the author's revised grain production (see text) minus grain procurement, measured in kilograms per person per day. Ukrainians is the 1926 Ukrainian population share. All regressions control for urbanization, province and year fixed effects. In columns (1)-(6), Huber-White robust standard errors are in parentheses; in columns (7)-(9), the standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Table A.10: Agricultural Output and Mechanization after the Famine

	Dependent Variable:	
	Log(Grain)	Tractors (Tractor Horse Power/1928 Sown Area)
	(1)	(2)
Dep. var. mean.	*9.943*	*0.027*
Ukrainians × 1922	-0.266	
	(0.242)	
Ukrainians × 1923	0.259	
	(0.210)	
Ukrainians × 1924	0.100	
	(0.207)	
Ukrainians × 1925	-0.197	
	(0.208)	
Ukrainians × 1926	-0.729***	
	(0.122)	
Ukrainians × 1927	-0.358*	
	(0.192)	
Ukrainians × 1928		0.001***
		(0.000)
Ukrainians × 1929	-0.339*	0.003***
	(0.179)	(0.001)
Ukrainians × 1930	-0.414**	0.005***
	(0.205)	(0.001)
Ukrainians × 1931	-0.157	0.007***
	(0.201)	(0.002)
Ukrainians × 1932	-1.725***	0.008***
	(0.213)	(0.003)
Ukrainians × 1933	-0.798***	0.016***
	(0.259)	(0.005)
Ukrainians × 1934	-1.018***	0.029***
	(0.143)	(0.008)
Ukrainians × 1935	-0.787***	0.029***
	(0.169)	(0.008)
Ukrainians × 1936	0.106	0.039***
	(0.324)	(0.012)
Ukrainians × 1937	-0.372**	0.024***
	(0.158)	(0.007)
Ukrainians × 1938	-0.204	0.018**
	(0.231)	(0.007)
Ukrainians × 1939	0.054	0.013**
	(0.223)	(0.006)
Ukrainians × 1940	-0.237	
	(0.164)	
Observations	354	247
R-squared	0.986	0.899

Notes: Ukrainians is the share of ethnic Ukrainians in the rural population in 1926. The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. All regressions control for grain suitability × year FE, province and year fixed effects. Column (1) also controls for monthly temperature and precipitation in year t and t-1 (48 additional controls). The standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

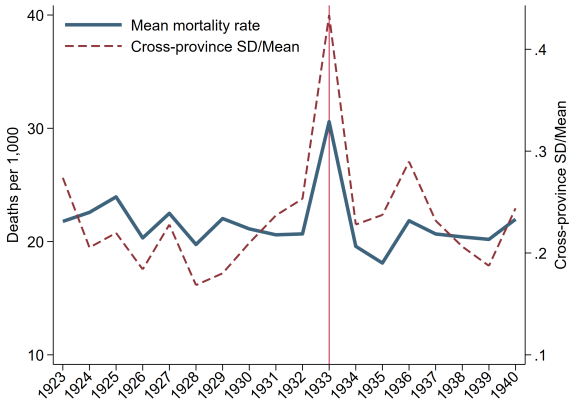
Table A.11: Population and Ethnic Composition after the Famine

	Dependent Variable:		
	Log Rural Population	Rural Share Ukrainians	Rural Share Russians
	(1)	(2)	(3)
Dep. var. mean.	*14.754*	*0.086*	*0.663*
Ukrainians × 1897	-0.113	-0.058***	0.044***
	(0.099)	(0.009)	(0.016)
Ukrainians × 1937	-0.309***		
	(0.071)		
Ukrainians × 1939	-0.220***	-0.217	0.219
	(0.039)	(0.150)	(0.150)
Ukrainians × 1959	0.013	-0.179	0.255
	(0.093)	(0.155)	(0.177)
Ukrainians × 1970	0.177	-0.184	0.270
	(0.162)	(0.154)	(0.176)
Ukrainians × 1979	0.312	-0.193	0.281
	(0.207)	(0.151)	(0.173)
Ukrainians × 1989	0.356	-0.214	0.284*
	(0.260)	(0.148)	(0.160)
Ukrainians × 2002	0.701***		
	(0.266)		
Observations	171	133	133
R-squared	0.952	0.941	0.957

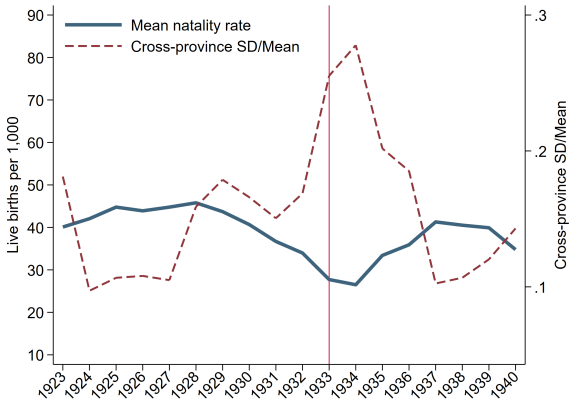
Notes: The dependent variables are reported by population censuses. Ukrainians is the share of ethnic Ukrainians in the rural population in 1926. The sample includes Ukraine, Russia and Belarus. Observations are at the province and year level. All regressions control for province and year fixed effects. The standard errors in parentheses are adjusted for spatial correlation within 1,500 km. *** p<0.01, ** p<0.05, * p<0.1

Figure A.1: Famine Severity over Time

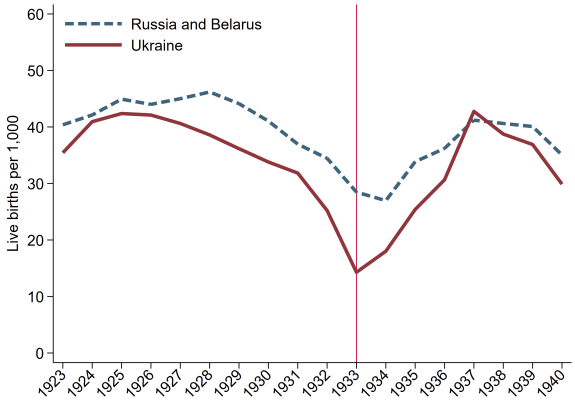
(a) Raw Mortality Mean and Variance



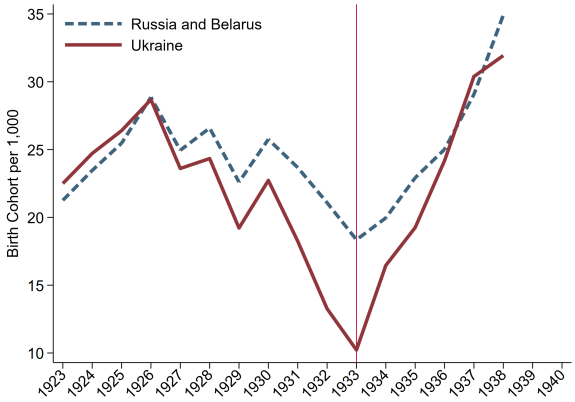
(b) Raw Natality Mean and Variance



(c) Raw Natality



(d) Raw Birth Cohort Size



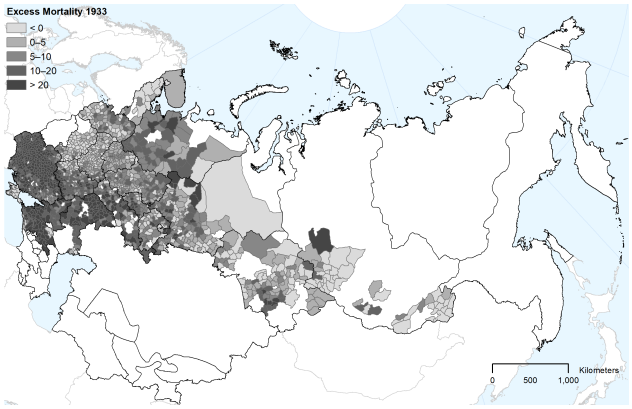
Notes: Mean mortality (natality) rate is the average mortality (natality) rate across provinces in each year. Cross-province SD/Mean is the standard deviation in mortality (natality) rates across provinces in year t divided by the mean mortality (natality) rate in year t. Natality is the number of live births per 1,000 individuals. Birth cohort size is the number of individuals born each year divided by the total population as reported by the 1939 census.

Figure A.2: Maps

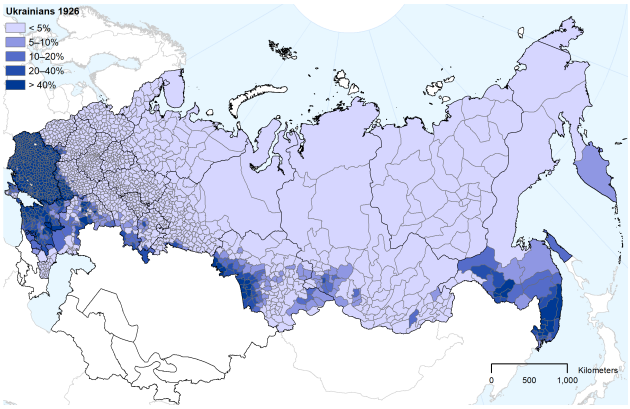
(a) Province Excess Mortality 1933 and Grain-Producing Regions

(b) Province Ethnic Ukrainians (1926) and Grain-Producing Regions

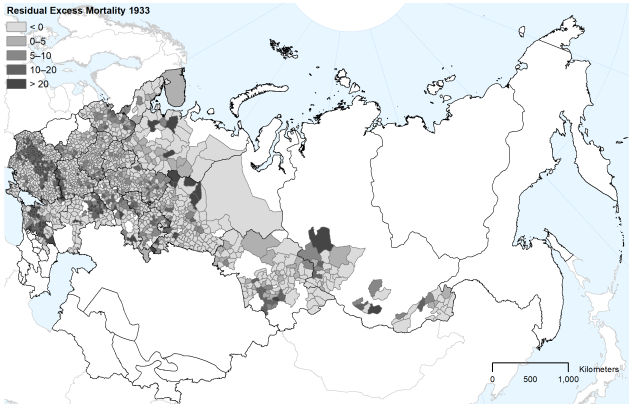
(c) District Excess Mortality 1933



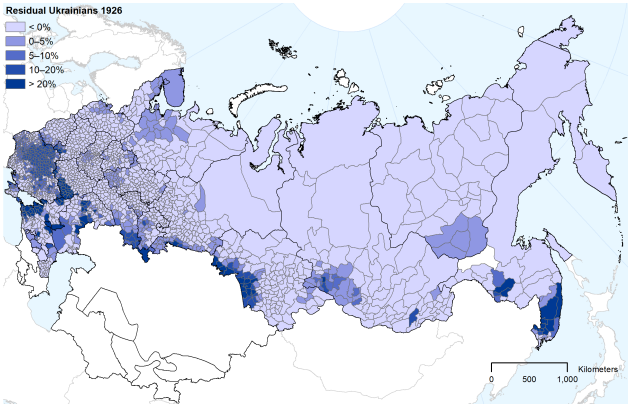
(d) District Ethnic Ukrainians 1926



(e) District Excess Mortality 1933 Demeaned by Province Fixed Effects



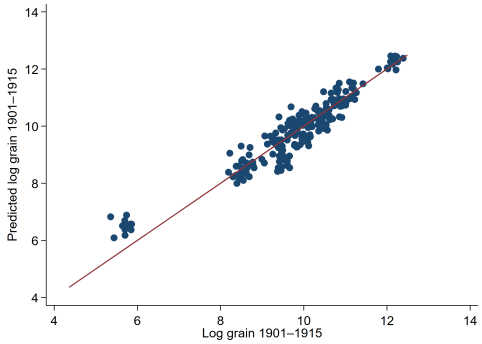
(f) District Ethnic Ukrainians 1926 Demeaned by Province Fixed Effects



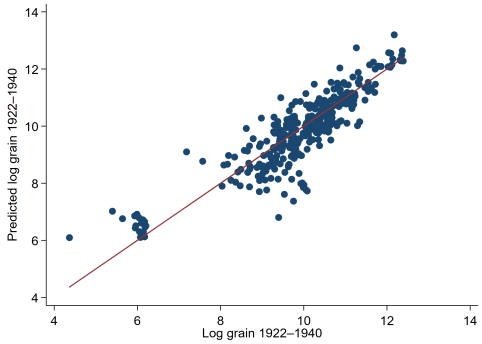
Notes: Excess mortality 1933 is mortality in 1933 minus mortality in 1928. Ethnic Ukrainians 1926 is the share of ethnic Ukrainians in the rural population according to the 1926 Population Census.

Figure A.3: Predicted Grain

(a) In-Sample Fit, $R^2 = 0.90$



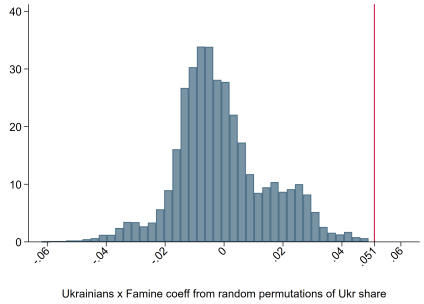
(b) Out-of-Sample Fit, $R^2 = 0.77$



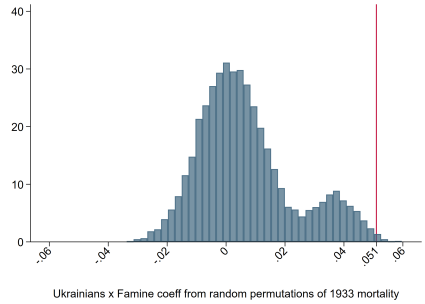
Notes: Log reported grain is plotted against log predicted grain with a 45-degree line for 1901–1915, a sample on which grain production function is estimated (in-sample fit) in Figure A.3a, and for 1922–1940 (out-of-sample fit) in Figure A.3b. See Appendix Section B for details.

Figure A.4: Random Permutation Test

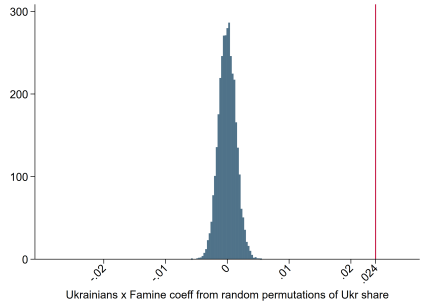
(a) Permute Ukrainian Population Share, Province-Level Panel



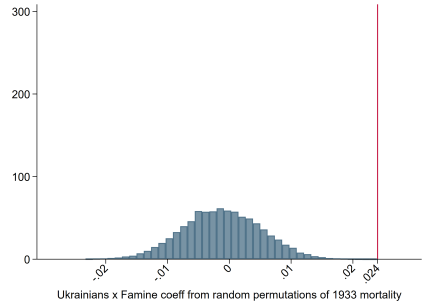
(b) Permute 1933 Mortality, Province-Level Panel



(c) Permute Ukrainian Population Share, District-Level Panel



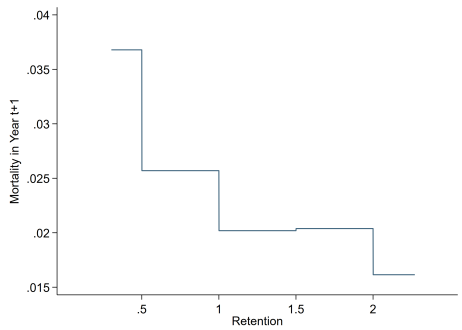
(d) Permute 1933 Mortality, District-Level Panel



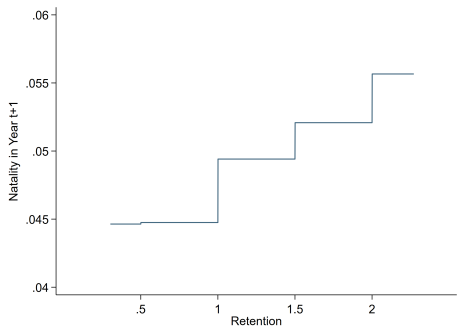
Notes: The figures are histograms of the interaction coefficients ($Ukrainians_i \times Famine_t$). The red vertical lines indicate the baseline estimates reported in the tables.

Figure A.5: Piece-wise Linear Estimation of Famine Severity and Grain Retention

(a) Mortality



(b) Natality



Notes: The figures plot the coefficients for different levels of grain retention (production - procurement), controlling for province and year fixed effects, from the piece-wise linear function in equation (2). The coefficients and standard errors are shown in Appendix Table A.9.

Data Appendix

Province-level Panel

The province-level panel includes 1922 to 1940 and 19 provinces within the republics of Belarus, Russia and Ukraine. These provinces correspond to the 1932 administrative division. Belarus and Ukraine were each a single province. The omitted territories are those with no reliable mortality data: Far Eastern Province, Yakut Autonomous SSR, and the North Caucasus ethnic territories: Chechen Autonomous Province, Cherkess Autonomous Province, Dagestan Autonomous SSR., Ingush Autonomous Province, Kabardino-Balkarian Autonomous Province, Karachay Autonomous Province, North Ossetian Autonomous Province. Figure A.2a maps the provinces in our sample. Omitted territories are in white.

Total, Rural and Urban Population For the years 1927, 1937 and 1939, we use data from population censuses.[4] For earlier years, 1920, 1923, 1924 and 1925 we use population estimates constructed by Soviet statisticians (see the exact sources below). For 1933, we use the same source as the 1933 district mortality: in addition to reporting the number of live births and deaths it also reports the population for the beginning of the year that was served by the civil acts registration bureaus (ZAGSy). For the remaining years, we interpolate population. In the paper, we check that our main finding is not an artifact of the interpolation by repeating our estimates with only 1926 population census data.

Administrative boundaries differ across years. However, since the data are reported at disaggregated administrative units, we are able to harmonize the data over time by using ArcGIS and manually aggregate the population data to the 1932 province borders. One issue is that small changes in borders that occur over time lead to large changes in population if we assume that the population is uniformly distributed across space in very large and sparsely populated provinces such as Ural, and West and East Siberia. To address this, we use the 1897 Population Census (the most recent available census prior to the start of our sample), which can be disaggregated to the *Uezd* level (of which there are 817 for the Russian Empire). These data allow us to calculate population density, which we use to attribute population to the 1932 province borders.

The data sources are as follows. 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: De-

[4]The 1926 Census took place on December 17, 1926, all other Soviet censuses took place in January; we use 1926 Census counts for the 1927 population. Population in the 1939 Census is widely believed to be inflated. This was recently corrected by Russian demographers using archival data (Bogoyavlensky, 2013). We use the corrected count.

cember 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.

Natality and Mortality Mortality is the total number of deaths divided by population (crude death rate). Natality is the number of live births divided by population (crude birth rate). We assign these data into 1932 province boundaries following the same procedure as for population. The rural-urban decomposition of deaths and births is available since 1926.

The data sources are as follows. 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 USSR] (1928) *"Yestestvennoye dvizheniye naseleniya Soyuza S.S.R. 1923–1925 [Natural movement of the population of the USSR]"*, 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.

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 (*volost*)-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 one-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 We map the grain data into 1932 provinces borders following the same procedure as for population. The years 1922, and 1924 to 1927 are reported for larger units than our provinces. Thus, we map them into the province borders proportional to the 1913 *Uezd* sown area data.

The data sources are as follows. 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 USSR 1918–1923. Five years of work of the Central Statistical Office.]"*, Volume XVIII of Trudy Tsentralnogo Statisticheskogo Upravleniya [Proceedings of the Central Statistical Office], Part VI, Tables 7 and 8. 1923: Tsentralnoye Statisticheskoye Upravleniye [Central Statistical Office] (1924) *"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 USSR] (1929) *"Selskoye khozyaystvo SS.S.R. 1925–1928. Sbornik statisticheskikh svedeniy k XVI Vsesoyuznoy partkonferentsii [Agriculture of the USSR 1925–1928. A collection of statistical information for the XVI All-Union Party Congress]"*, Part III. 1928: RGAE 1562/329/1409. 1929–1930: Gosudarstvennoye sotsialno-ekonomicheskoye izdatelstvo [State Socio-Economic Publishing House] (1932) *"Narodnoye khozyaystvo SS.S.R.. Statisticheskiy spravochnik 1932 [The national economy of the USSR. Statistical Handbook 1932]"*, Part II.3.A, Tables 30 and 33. 1931: Gosudarstvennoye izdatelstvo kolkhoznoy i sovkhoznoy literatury "Selkhozgiz" [State publishing house of collective and state farm literature "Selkhozgiz"] (1936) *"Selskoye khozyaystvo SS.S.R.. Yezhegodnik 1935 [Agriculture of the USSR. Yearbook 1935]"*, p. 269, Tables 106 and 107. 1932–1940: RGAE 1562/329/1409.

Procurement We calculated 1925 to 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 to 1933 data is used as reported.

The data sources are as follows. 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 USSR] (1929) *"Selskoye khozyaystvo SS.S.R. 1925–1928. Sbornik statisticheskikh svedeniy k XVI Vsesoyuznoy partkonferentsii [Agriculture of the USSR 1925–1928. A collection of statistical information for the XVI All-Union Party Congress]", Part V.* 1928: calculated from the 1928 grain harvest and procurement as a share of harvest from RGAE 4372/30/871 p. 30. 1929: Narodnyy Komissariat Snabzheniya SS.S.R. [People's Commissariat of Supply of the USSR] (1932) *"Yezhegodnik khlebooborota N4 [Yearbook of grain turnover N 4]"*, Tables 3 and 10. 1930: Narodnyy Komissariat Snabzheniya SS.S.R. [People's Commissariat of Supply of the USSR] (1932) *"Yezhegodnik khlebooborota N4 [Yearbook of grain turnover N 4]"*, Table 29 and Table 36. 1931: Komitet po zagotovkam S.-Kh produktov pri SNK SS.S.R. [Committee for Procurement of Agricultural Products under the Council of People's Commissars of the USSR] (1934) *"Yezhegodnik khlebooborota za 1931-32, 1932-33 i predvaritelnyye itogi zagotovok 1933 g. [Yearbook of grain turnover for 1931-32, 1932-33 and preliminary results of procurement in 1933]"*, Table 21. 1932: Komitet po zagotovkam S.-Kh produktov pri SNK SS.S.R. [Committee for Procurement of Agricultural Products under the Council of People's Commissars of the USSR] (1934) *"Yezhegodnik khlebooborota za 1931-32, 1932-33 i predvaritelnyye itogi zagotovok 1933 g. [Yearbook of grain turnover for 1931-32, 1932-33 and preliminary results of procurement in 1933]",* Table 33. 1933: Komitet po zago-

tovkam S.-Kh produktov pri SNK SS.S.R. [Committee for Procurement of Agricultural Products under the Council of People's Commissars of the USSR] (1934) *"Yezhegodnik khlebooborota za 1931-32, 1932-33 i predvaritelnyye itogi zagotovok 1933 g. [Yearbook of grain turnover for 1931-32, 1932-33 and preliminary results of procurement in 1933]",* Table 53.

Grain Targets The First Five-Year plan reports plans for population, grain production, and procurement ("marketable grain"). Aggregate Soviet-level figures are available for every year from 1928 to 1933; from these yearly figures it is clear that the planners just linearly interpolated targets between 1928 and 1933. Disaggregated province-level figures are available for 1928 and 1933; we linearly interpolate planned grain production, procurement and population targets between these years. Summing up the interpolated province targets mechanically results in the aggregate target for that year. Conceptually, the interpolation assumes that the central planners used the same formula to set targets at the aggregate and province levels.

The provinces in the Plan are slightly larger than the provinces in our sample (e.g., West and East Siberia are reported together, Ivanovo and Moscow provinces are united into a Central Industrial Region, Leningrad province and Karelia are united, Tatar and Chuvas regions are united with the Middle Volga region). We calculate grain production and procurement targets in kilograms per person per day (planned grain divided by planned population). Planovoye khozyaystvo [Planned economy] (1930) *"Pyatiletniy plan narodno-khozyaystvennogo stroitel'stva SSSR. Tom 3: rayonnyy razrez plana [Five-year plan for the national economic construction of the USSR. Volume 3: regional aspect of the plan.]"*

Collectivization 1927: Statizdat TSSU SS.S.R. [Statistical publishing house of the Central Statistical Office of the USSR] (1929) *"Kollektivizatsiya Sovetskoy derevni. Predvaritelnyye itogi sploshnykh obsledovaniy 1928 i 1929 gg. [Collectivization of the Soviet countryside. Preliminary results of comprehensive surveys in 1928 and 1929]"*, Table 10. 1928: RGAE 1562/82/271. 1929: Gosplan S.S.S.R. i RSFSR. Ekonomiko-statisticheskiy sektor [State Planning Committee of the USSR and the RSFSR. Economic and statistical sector] (1931) *"Kolkhozy v 1929 g. Itogi sploshnogo obsledovaniya kolkhozov [Collective farms in 1929. Results of a comprehensivy survey of collective farms]"*. 1930: Gosplan S.S.S.R.. Upravleniye Narodnokhozyaystvennogo Ucheta [State Planning Committee of the USSR. Department of National Economic Accounting] (1931) *"Kolkhozy v 1930 g. Itogi raportov kolkhozov k XVI s'yezdu VKP(b) [Collective farms in 1930. Resume of the collective farms' reports to the XVI Congress of the CPSU(b)]"*. 1931: Izd. Kolkhoztsentra SS.S.R. i RSFSR [Publishing House of the Collective Farm Center of the USSR and the RSFSR] (1931) *"Kolkhoznoye stroitelstvo v SS.S.R. [Collective farms building in the USSR]"*, p. 15 and Davies and Wheatcroft (2009), 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 USSR]"*, 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 USSR 1939 (Statistical handbook)]"*, Part IV.

Dekulakization The baseline measure of *kulak* households exiled during 1930 to 1931 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 to 1931 according to Davies and Wheatcroft (2009) (Table 28) per 1930 population. Exiled *kulaks* (OGPU) is the number of *dekulakized* and exiled households of all categories between January 1, 1930 and July 1, 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 January 1, 1930 and April 1, 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 from 1895 to 1914 are from Gokmen and Kofanov (2020).

Bolshevik Votes 1917 Bolshevik vote share is from Protasov et al. (2014). Data is calculated in our province borders using district (*uezd*)-level 1917 map from Castañeda Dower and Markevich (2021).

Communists 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). The reported data are more disaggregated than our provinces. For 1931, we use the reported data.

The data sources are as follows. 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]"*

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.

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 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 Panel

District-level dataset includes district for which we were able to collect mortality data for 1933: 1928 and 1933, approximately 1,600 districts within the republics of Russia and Ukraine. Appendix Figure A.2c maps the districts in our sample (omitted territories are in white).

Mortality 1928: Russia: State archive of the Russian federation (GARF) 374/23/7, 13, 31–32, 67, 72–91, 132, 158; Ukraine: Tsentralna Statistichna Uprava USRR [Central Statistical Office of Ukraine] (1929) "Ukraina: Statisticheskiy Schorichnik 1929 [Ukraine: Statistical Yearbook 1929]." 1933: RGAE 1562/329/18–19.

Ethnic Composition Ethnic composition comes from the 1926 Population Census. 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.

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.

Collectivization Gosplan S.S.S.R.. Upravleniye Narodnokhozyaystvennogo Ucheta [State Planning Committee of the USSR. Department of National Economic Accounting] (1931) "Kolkhozy v 1930 g. Itogi raportov kolkhozov k XVI s'yezdu VKP(b) [Collective farms in 1930. Resume of the collective farms' reports to the XVI Congress of the CPSU(b)]." 1930 districts matched to 1933 districts by name.