*Supplementary Appendix for*

# The Slow Road from Serfdom: Labor Coercion and Long-Run Development in the Former Russian Empire

Johannes C. Buggle [*]       Steven Nafziger[†]

November 13, 2018

This appendix accompanies the paper *"The Slow Road from Serfdom: Labor Coercion and Long-Run Development in the Former Russian Empire"* by Johannes C. Buggle and Steven Nafziger.

## Contents

[*]Department of Economics, University of Lausanne. mail: johannes.buggle@unil.ch

[†]Department of Economics, Williams College. mail: snafzige@williams.edu

# A  Data Description and Sources

## A.1  Country-Level Data

**Date of Peasant Emancipation across Countries:** The emancipation dates for the 18 countries used in Figure 1 are taken from several sources. The dates refer to the legal emancipation of peasants. Information for Austria, Denmark, France, Switzerland, Estonia, Latvia, Germany (Prussia), Hungary, Russia and Ukraine, Belarus, Romania is taken from Blum (1978). Data for Netherlands from Jarrett (2013), for Bulgaria from Zagorov (1955), dates for Croatia from Magas (2008), and for Italy (Naples) and Spain from Atkin et al. (2011). For Poland we use the average date of the Duchy of Warsaw and the Kingdom of Poland which are taken from Davies (2005).

**Agricultural Employment 1900 and 2000:** The estimates of the share of labor in agriculture in 1900 and 2000 come from Wingender (2014).

## A.2  District-Level Data

**Serfdom Data:** The main explanatory variable, *Serfs % (1858)*, is constructed using the sum of total male and female serfs in 1858 per district (taken from Troinitskii (1982 (1861), who relies on the 10th and last tax census, or *revitsii*, conducted during the pre-1861 period), divided by district population in 1858 (taken from Bushen, ed (1863)). Other population sources from around 1861 give slightly different percentage serf values. We rely on total population as that appears to be more accurately recorded than any measure of the total number of peasants or the rural population alone. Using our two preferred sources (numerator and denominator), the weighted mean of the serfdom variable is approximately 36.5%. Markevich and Zhuravskaya (2018) and Castaneda Dower et al. (2018) utilize essentially the same variable in their district-level analyses, with the former relying on a slightly smaller sample. Our note on the share of *peasants* who were serfs in the Introduction – slightly less than 45% – reflects a rough estimate derived from the relevant provincial numbers in Table 1 of Markevich and Zhuravskaya (2018).

**Geographic Controls:**

- Longitude and latitude information based on own calculations at the district's centroid using ArcGIS.

- Distance to Moscow gives the distance in kilometers from the centroid of each historical district to Moscow.

- Cereal suitability (and its local standard deviation within historical districts), as well as suitability for growing wheat, rye, barley and oat, measure soil suitability for crop cultivation under low-input, rain-fed conditions and are taken from the FAO-GAEZ database. Cereal suitability is measured in eight suitability classes

(Very high; High; Good; Medium; Moderate; Marginal; Very marginal; Not suitable). For the set of flexible controls, we construct dummy variables for each class of (average) cereal suitability. Wheat, rye, barley, and oat suitability are defined as continuous suitability indices that vary between 0 and 10.

- Distance to Coast measures the nearest distance in kilometers of the district's centroid to the coastline.

- Forest Cover measures the average share of land covered by forests and is based on data from the FAO-GAEZ database.

- Ruggedness measures the average terrain slope and is based on data from the FAO-GAEZ database.

- River Density measures the total length of rivers divided by the area of the district. Information on the location of rivers is taken from O'Neill (2016).

- Mean Temperature, and Mean Precipitation during the growing season, are averages over the period 1901-2000, using gridded climate data from the CRU 3.10.1 database of the Climatic Research Unit at the University of East Anglia. We define the growing season as the period from April to September, based on information from the USDA publication *Major World Crop Areas and Climatic Profiles* (1994). Growing Season Variability in Temperature and Precipitation are defined as the average year-to-year variability in temperature and precipitation over the period 1901-2000.

- Podzol Soil measures the share of land in which podzol is the dominant soil and is taken from the FAO-GAEZ database.

- Distance Provincial Capital measures the distance in kilometers from the district's centroid to the centroid of the district where the Provincial capital is located.

- Distance City in 1600 measures the distance in kilometers from the district's centroid to the location of the nearest existing city with at least 5,000 inhabitants in 1600, as reported in the dataset of Bairoch et al. (1988).

- Length Growing Period measures the average number of days during which local environmental conditions allow crop growth. This information is take from the FAO-GAEZ database.

- Distance to St. Petersburg gives the distance in kilometers from the centroid of each historical district to St. Petersburg.

- Coal Territory is an indicator for the presence of coal territories, and is based on data from the *Coal Quality and Resources of the Former Soviet Union* project of the U.S. Geological Survey (Brownfield et al. 2001), accessed via http://pubs.usgs.gov/of/2001/ofr-01-104/fsucoal/html/data1.htm.

- (log) Population Density in 1858 is constructed by dividing the total population of a district in 1858, taken from Bushen, ed (1863), by the area of the district.

4

**Road Density in the Countries of the former Soviet Union:** Road densities in 1996 are constructed using digitized maps provided by the *Coal Quality and Resources of the Former Soviet Union* project of the U.S. Geological Survey (Brownfield et al. 2001), accessed via http://pubs.usgs.gov/of/2001/ofr-01-104/fsucoal/html/data1.htm.

**Employment Shares 1897:** We construct employment shares in different sectors from the occupational census of 1897 by adding up the number of people that work in each sector per district, divided by the district labor force. The primary sector includes occupations in agriculture, fishing and hunting, forestry work, livestock cultivation and mining. The secondary sector contains occupations that process materials such as animals, metals, fiber elements, or wood, as well as occupations that produce textile, instruments, pottery, jewelry, metals, minerals, spirits and alcohol, carriage, chemicals, tobacco, printing and paper, foods and others. In a separate variable we only include the industrial occupations such as in textile production, production of instruments, metals, carriage etc. Service occupations contain those that belong to administration and police, public service, private service and medical service occupations. Commercial occupations are those that involve trading (e.g. grain trade, metal trade, textile trade etc.) as well as credit and insurance. Education and research occupations are teachers and educators and all occupations that involve science, literature and art.

**Other Historical Outcomes:** We drew on a variety of sources to construct other historical outcome measures to investigate the mechanisms behind the persistent impact of serfdom. Male literacy in 1897 is defined for rural residents in their 20s from data reported in Troinitskii, ed (1905). That source also provides the share of the adult male population with agriculture as the primary occupation. Rural enrollment rates for 1880 and 1894 are defined with both numerators and denominators taken from Tsentraf'nyi statisticheckikh komitet, Ministerstvo vnutrennykh del (1884) and Fal'bork and Charnoluskii, eds (1900-1905), respectively – see Nafziger (2012) for more information. Fal'bork and Charnoluskii, eds (1900-1905) also provide the number of formally recognized primary schools by district in 1856. The urbanization rate in 1913 is derived from Tsentraf'nyi statisticheckikh komitet, Ministerstvo vnutrennykh del (1914). The land Gini (both types), the percentage of land owned by the nobility or in communal tenure, and the amount of land possessed per peasant household, all defined in 1905, are from Tsentraf'nyi statisticheckikh komitet, Ministerstvo vnutrennykh del (1906), with additional details provided in Nafziger (2013). Finally, information on factory production and employment in 1868 is compiled from Tsentraf'nyi statisticheckikh komitet, Ministerstvo vnutrennykh del (1872).

**Ln Population Density in 2000** is based on the Gridded Population of the World (GPW), version 3, data.

**Night-Time Luminosity:** We use the log of average luminosity at night measured

as "Average Visible, Stable Lights, and Cloud Free Coverage" for several years from 1994 to 2012. The data is taken from the National Geophysical Data Center (http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html).

### A.3 Life in Transition Surveys (2006 - 2016)

**Log Equivalent Expenditure Per Capita** comes from several questions asking about expenditure during a 30 days and 12 month period. Since items and recall periods differ across the three waves, we harmonize and select common items across waves. Total expenditure is the sum of expenditure on food, beverages and tobacco; clothing and footwear; education; health; household durables. We annualized all monthly expenditures (food, beverages and tobacco in all waves; clothing and footwear in wave 2006. Dropping clothing and footwear - that does not have a consistent recall period across waves - from the calculation of expenditure in all waves does not change the results.). Expenditure is expressed in US Dollars (to convert expenditure in the waves 2010 and 2016 from local currencies to USD we use market exchange rates), and adjusted by household size. The equivalence scale assigns a value of 1 to the household head, of 0.5 to each additional adult member and of 0.3 to each child.

In addition, we use a second measure of household expenditure that is based on wave 2006 only for robustness in Table C6. This measure has the advantage to include two additional relevant household expenditure items (spending on transport and communication expenses, as well as spending on recreation, entertainment and meals outside the home.)

**Household controls:** Household size, share of male, share of persons aged 0-18, share of persons aged 60+, religious denomination of the respondent, survey round. Source is Life in Transition Survey, waves 2006, 2010, and 2016.

**Consumer goods:** Question asks "Do you or anyone in your household own any of the following?" We sum up the mentioning of ownership of a car, a mobile phone, or a computer. Source is Life in Transition Survey, waves 2006, 2010, and 2016.

**Access to Public Goods:** Question asks "Do you have access to [UTILITY] in this dwelling?" and UTILITY are different type of public goods, such as water, landline, heating, sewage, public sewage system. Source is Life in Transition Survey, wave 2006, 2010, and 2016.

**Sale Farm Products and Land Cultivation:** Sale Farm Products comes from the question "Which of these sources of livelihood apply to your household?", and takes on the value 1 if the household mentions "Sales or bartering of farm products", and 0 otherwise. Taken from waves 2006 and 2010. Land cultivation comes from the question "Does your household cultivate any land?" and takes on value 1 for households that

cultivate land, and 0 for those that do not. Taken from wave 2006.

**Education:** Highest education measures the highest educational achievement of the respondent in 5 classes (no degree; compulsory education; secondary education; professional training; tertiary education or postgraduate education). Post secondary education equals 1 if the highest educational level of the respondent above secondary education (i.e., professional training or tertiary education or postgraduate education), and 0 otherwise. Tertiary education equals 1 if the highest educational level of the respondent is tertiary education (i.e., tertiary education or postgraduate education), and 0 otherwise (Source is Life in Transition Survey, waves 2006,, and 2016.). Parents education is the average of the father's and mother's education measured in years (Source is Life in Transition Survey, wave 2010). Parents with tertiary education is the number of parents that have completed tertiary education (Source is Life in Transition Survey, waves 2006 and 2016). We also use the question "In your opinion, which of these fields should the first priority for extra (government) investment?" to construct an indicator variable equal to 1 if the respondent mentions education, asked in waves 2006, 2010, and 2016.

**Inequality:** From the household survey, we construct Gini inequality indices of expenditure and consumer good ownership across households that live within a common historical district using the "egen_inequal" command in Stata.

**Attitudes:**

- *Less or More Inequality?* Question asks "Now I'd like you to tell me your views on various issues. How would you place your views on this scale? 1 means you agree completely with the statement on the left; 10 means you agree completely with the statement on the right; and if your views fall somewhere in between, you can choose any number in between: Incomes should be made more equal (1) vs We need larger income differences as incentives for individual effort (10)". Source is Life in Transition Survey, waves 2010 and 2016.

- *Poor Unlucky?* Question asks "In your opinion, what is the main reason why there are some people in need in our country today? Because they have been unlucky." We create a dummy that takes on the value 1 if the respondent mentioned luck as the main reason. Source is Life in Transition Survey, waves 2006 and 2010.

- *Demonstrated/Striked/Signed A Petition:* Question asks "How likely are you to... attend a lawful demonstration - participate in a strike - sign a petition" and respondents' answers range from "have done (3) - might do (2) - would never do (1)". Source is Life in Transition Survey, waves 2006, 2010, and 2016.

- *Pref Market Economy:* Question asks "With which one of the following statements do you agree most? A market economy is preferable to any other form of eco-

nomic system - Under some circumstances, a planned economy may be preferable to a market economy - For people like me, it does not matter whether the economic system is organized as a market economy or as a planned economy" The variable takes on the value 1 if the respondent states that "A market economy is preferable to any other form of economic system " and 0 otherwise. Source is Life in Transition Survey, waves 2006, 2010, and 2016.

- *Pref Democracy:* Question asks "With which one of the following statements do you agree most? Democracy is preferable to any other form of political system - Under some circumstances, an authoritarian government may be preferable to a democratic one - For people like me, it does not matter whether a government is democratic or authoritarian" The variable takes on the value 1 if the respondent states that "Democracy is preferable to any other form of political system " and 0 otherwise. Source is Life in Transition Survey, waves 2006, 2010, and 2016.

- *Trust:* The variable takes on the value 1 if the respondent states that "Most people can be trusted" and 0 otherwise. Source is Life in Transition Survey, waves 2006, 2010, and 2016.

- *Xenophobia:* The variable are derived from questions asking respondents about "People you do not want to have as neighbours ...". Each variable takes on the value 1 if the respondent mentions the category ("different race"/ "immigrant" / "different religion" / "Jewish people" / "different language"), and 0 otherwise. Source is Life in Transition Survey, wave 2016.

- *Communism:* The variable takes on the value 1 if at least one family member of the respondent was in the communist party, and 0 otherwise. Source is Life in Transition Survey, wave 2016.

### A.4 City-Level Data

**City Population 1897-2002:** The population of Russian cities for the years 1897, 1926, 1939, 1959, 1970, 1979, 1989, 2002 has been compiled by Mikhailova (2012) and is based on the population census of the Russian Empire (1897), the Soviet Union (1926-1989) and the Russian Federation (2002). For the cross-sectional sample, we construct a balanced sample of 366 population centers with city status. For the sample from 1800 - 2002 used in the panel analysis, we merge the data of Mikhailova (2012) to the population data constructed by Bairoch et al. (1988) that add the years 1800 and 1850, for a total of 99 cities.

**Military Industry 1939 - 1989**: Data on the location of defense-related factories during the Soviet period (1939, 1945, 1959, 1970, 1979, 1989) for a sample of 278 Russian and Ukrainian cities is taken from Acemoglu et al. (2011), who matched 17,914 establishments to their current location. The source data is the "Factories, Research and

Design establishments of the Soviet Defence Industry" database Version 11 (2010) by Keith Dexter and Ivan Rodionov. We use the total number of factories per city, as well as the growth rates of factories during various time periods as outcome variables.

## A.5 Firm-Level Data 1989

We use information on the characteristics of firms that existed in 1989 from the *1989 Soviet Census of Manufacturers*. For each firm we construct the (log) of total employment, the (log) of total turnover, and the (log) of turnover per worker. Additionally, using the Standard Industrial Classification codes we identify agricultural and primary sector firms with an indicator variable. In addition, we count the number of firms per town and district in 1989.

## A.6 Descriptives

TABLE A1: SUMMARY STATISTICS

| District-Level | Mean | S.D. | N |
|---|---|---|---|
| Serfs % (1858) | 0.39 | 0.25 | 490 |
| Quit-Rent % (1858) | 0.11 | 0.17 | 468 |
| Corvée % (1858) | 0.26 | 0.21 | 468 |
| Household Serfs % (1858) | 0.03 | 0.02 | 486 |
| Latitude | 54.09 | 3.80 | 490 |
| Longitude | 36.97 | 8.44 | 490 |
| Distance to Moscow | 6.29 | 3.16 | 490 |
| Distance to Coast | 5.97 | 3.49 | 490 |
| Forest Cover | 36.07 | 23.57 | 490 |
| Ruggedness | 91.36 | 4.82 | 490 |
| River Density | 0.02 | 0.01 | 490 |
| Cereal Suitability | 6.24 | 1.30 | 490 |
| Mean Temperature Apr-Sep | 1.42 | 0.19 | 490 |
| Mean Precipitation Apr-Sep | 56.66 | 8.09 | 490 |
| Podzol Soil | 0.35 | 0.37 | 490 |
| Wheat Suitability | 6.63 | 2.17 | 490 |
| Rye Suitability | 6.01 | 2.25 | 490 |
| Barley Suitability | 6.08 | 2.33 | 490 |
| Oat Suitability | 6.08 | 2.29 | 490 |
| Distance City in 1600 | 0.23 | 0.18 | 490 |
| Distance Provincial Capital | 1.24 | 0.97 | 490 |
| SD Cereal Suitability | 0.87 | 0.36 | 490 |
| Length Growing Period | 159.51 | 33.82 | 490 |
| Growing Season Variability Precipitation | 27.18 | 2.39 | 490 |
| Growing Season Variability Temperature | 0.18 | 0.02 | 490 |
| Coal Territory 0/1 | 0.39 | 0.49 | 490 |
| Distance to St. Petersburg | 9.25 | 3.81 | 490 |
| (ln) Pop Density 1858 | 3.92 | 0.91 | 478 |
| Perc. Christian 1870 | 0.13 | 0.39 | 490 |
| Perc. Muslim 1870 | 0.06 | 0.20 | 490 |
| Perc. Orthodox 1870 | 0.95 | 0.44 | 490 |
| Perc. Armenian 1870 | 0.04 | 0.19 | 490 |
| Perc. Jewish 1870 | 0.06 | 0.18 | 490 |
| Urbanization Rate 1863 | 8.87 | 10.11 | 483 |
| Urbanization Rate 1913 | 10.10 | 12.19 | 490 |

| | | | |
|---|---|---|---|
| Factories Per 1,000 ppl 1868 | 27.13 | 60.22 | 483 |
| Log Production per worker 1868 | 13.00 | 0.94 | 434 |
| Road Density | 0.01 | 0.01 | 490 |
| Gulag 0/1 | 0.20 | 0.40 | 490 |
| Log Population Density 2000 | 3.42 | 1.09 | 490 |
| Primary Employment 1897 | 0.62 | 0.15 | 490 |
| Secondary Employment 1897 | 0.13 | 0.09 | 490 |
| Industry Employment 1897 | 0.09 | 0.08 | 490 |
| Service Employment 1897 | 0.01 | 0.01 | 490 |
| Education & Research Employment 1897 | 0.01 | 0.00 | 490 |
| Commerce Employment 1897 | 0.03 | 0.02 | 490 |
| Log Light Density in 1994 | 0.85 | 0.95 | 490 |
| Log Light Density in 1999 | 0.60 | 1.08 | 490 |
| Log Light Density in 2005 | 0.31 | 1.03 | 490 |
| Log Light Density in 2008 | 0.57 | 1.01 | 490 |
| Log Light Density in 2012 | 1.03 | 0.98 | 490 |
| Schools before 1856 per thousand | 0.05 | 0.13 | 486 |
| Enrollment 1880 | 8.57 | 7.47 | 489 |
| Total Schools per thousand 1911 | 1.50 | 0.64 | 486 |
| Land Owned by Nobles (perc.), 1905 | 21.00 | 14.02 | 490 |
| Land Gini 1905 | 0.49 | 0.16 | 467 |
| *LiTS Household Survey 2006 – 2016* | Mean | S.D. | N |
| (ln) Equivalent Expenditures Per Capita | 4.96 | 0.85 | 5823 |
| Household Size | 2.34 | 1.24 | 22061 |
| Share Young | 0.12 | 0.20 | 22061 |
| Share Old | 0.33 | 0.44 | 22061 |
| Share Male | 0.42 | 0.30 | 22061 |
| No Religion | 0.14 | 0.34 | 21738 |
| Buddhist Religion | 0.00 | 0.04 | 21738 |
| Christian Religion | 0.83 | 0.38 | 21738 |
| Jewish Religion | 0.00 | 0.04 | 21738 |
| Muslim Religion | 0.01 | 0.09 | 21738 |
| Rural PSU | 0.35 | 0.48 | 22061 |
| Age of Respondent Individual Questions | 49.87 | 18.40 | 22052 |
| Age Sq. of Respondent Individual Questions | 2825.48 | 1893.00 | 22052 |
| Gender of Respondent Individual Questions | 0.37 | 0.48 | 22058 |
| Consumer Goods | 1.75 | 1.03 | 22056 |
| Sale Farm Products | 0.08 | 0.27 | 13240 |

| Table A1 (continued) | | | |
|---|---|---|---|
| Land Cultivation | 0.53 | 0.50 | 6177 |
| Highest Education | 3.75 | 1.02 | 22060 |
| Post Secondary | 0.59 | 0.49 | 22060 |
| Tertiary | 0.29 | 0.45 | 22060 |
| Parents Education Years | 9.17 | 4.26 | 4410 |
| Parents w/ Tertiary | 0.31 | 0.64 | 13405 |
| Education Gov. Priority? | 0.39 | 0.49 | 21722 |
| Poor: unlucky | 0.08 | 0.27 | 13239 |
| Equal incomes vs inequality | 4.82 | 2.94 | 15255 |
| Demonstrated 0/1 | 0.33 | 0.47 | 21061 |
| Striked 0/1 | 0.29 | 0.45 | 21061 |
| Joined Party 0/1 | 0.15 | 0.36 | 12240 |
| Average Protest | 0.35 | 0.41 | 21061 |
| Trust | 0.37 | 0.48 | 21520 |
| Prefer Market Economy | 0.36 | 0.48 | 19864 |
| Prefer Democracy | 0.47 | 0.50 | 20072 |
| Neighbor: Race | 0.18 | 0.38 | 8821 |
| Neighbor: Immigrant | 0.31 | 0.46 | 8821 |
| Neighbor: Religion | 0.12 | 0.32 | 8821 |
| Neighbor: Jew | 0.09 | 0.29 | 8821 |
| Neighbor: Language | 0.09 | 0.29 | 8821 |
| Average Xenophobia | 0.03 | 1.54 | 8821 |
| At Least One Communist Family Member | 0.34 | 0.47 | 8821 |
| Access to Water | 0.84 | 0.37 | 22033 |
| Access to Landline | 0.60 | 0.49 | 21927 |
| Access to Heating | 0.57 | 0.49 | 21812 |
| Access to Sewage | 0.62 | 0.48 | 6125 |
| PC Public Goods | 0.03 | 1.55 | 6119 |

| *City Population Cross-Section* | Mean | S.D. | N |
|---|---|---|---|
| Log City Population 1897 | 9.00 | 1.00 | 366 |
| Log City Population 1926 | 9.33 | 1.06 | 366 |
| Log City Population 1939 | 9.82 | 1.18 | 366 |
| Log City Population 1959 | 10.10 | 1.29 | 366 |
| Log City Population 1970 | 10.33 | 1.34 | 366 |
| Log City Population 1989 | 10.53 | 1.39 | 366 |
| Log City Population 2002 | 10.47 | 1.41 | 366 |

| *City Population Panel* | Mean | S.D. | N |
|---|---|---|---|
| (Log) City Population | 9.95 | 1.22 | 5031 |

| *Factory Location* | Mean | S.D. | N |
|---|---|---|---|
| Number of Establishments in 1939 | 5.86 | 13.61 | 278 |
| Number of Establishments in 1945 | 7.54 | 17.70 | 278 |
| Number of Establishments in 1959 | 10.57 | 25.02 | 278 |
| Number of Establishments in 1970 | 13.86 | 31.25 | 278 |
| Number of Establishments in 1979 | 14.66 | 32.40 | 278 |
| Number of Establishments in 1989 | 15.55 | 34.01 | 278 |

| *Firms 1989* | Mean | S.D. | N |
|---|---|---|---|
| Number of Firms 1989 (per City) | 5.33 | 19.58 | 2656 |
| Agriculture | 0.07 | 0.25 | 14154 |
| Primary Sector | 0.09 | 0.29 | 14154 |
| (log) Employment | -1.46 | 1.41 | 14055 |
| (log) Turnover | 1.43 | 1.79 | 13933 |
| (log) Turnover per Worker | 2.89 | 0.93 | 13923 |

## A.7   Distribution of Serfdom

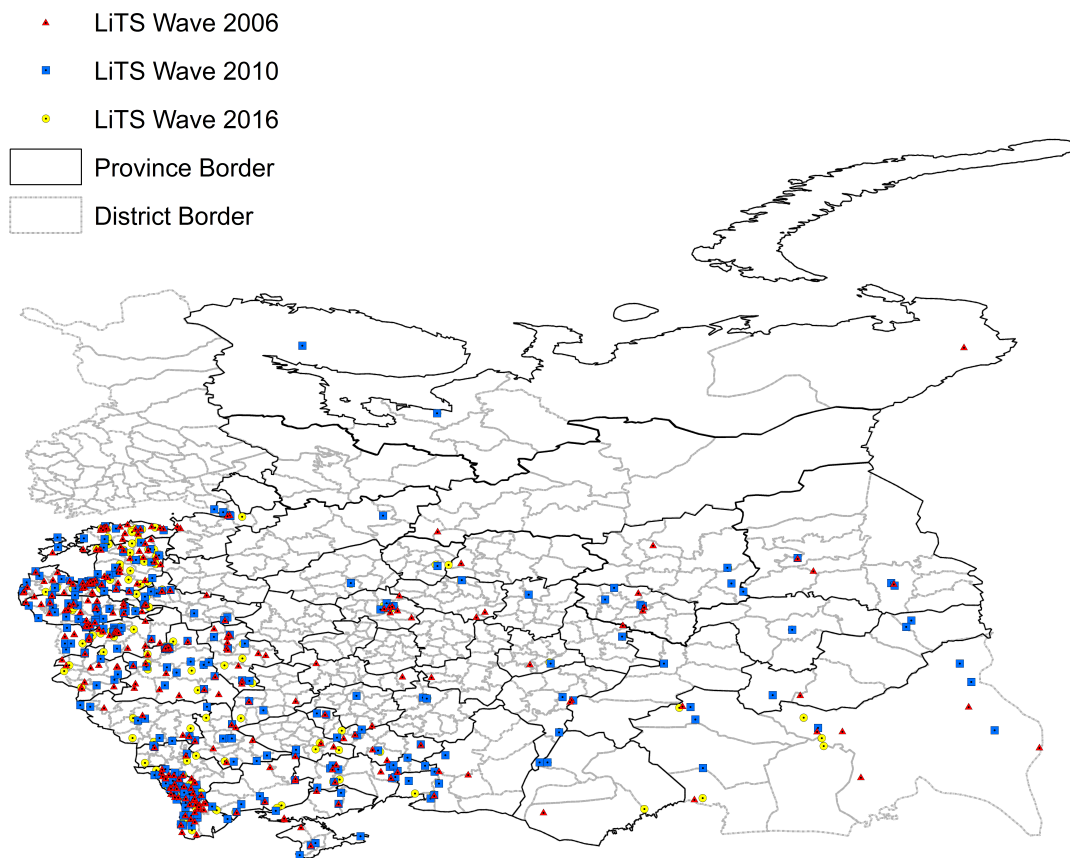FIGURE A1: DISTRIBUTION OF SERFS AS SHARE OF POPULATION, C. 1858. N = 495.
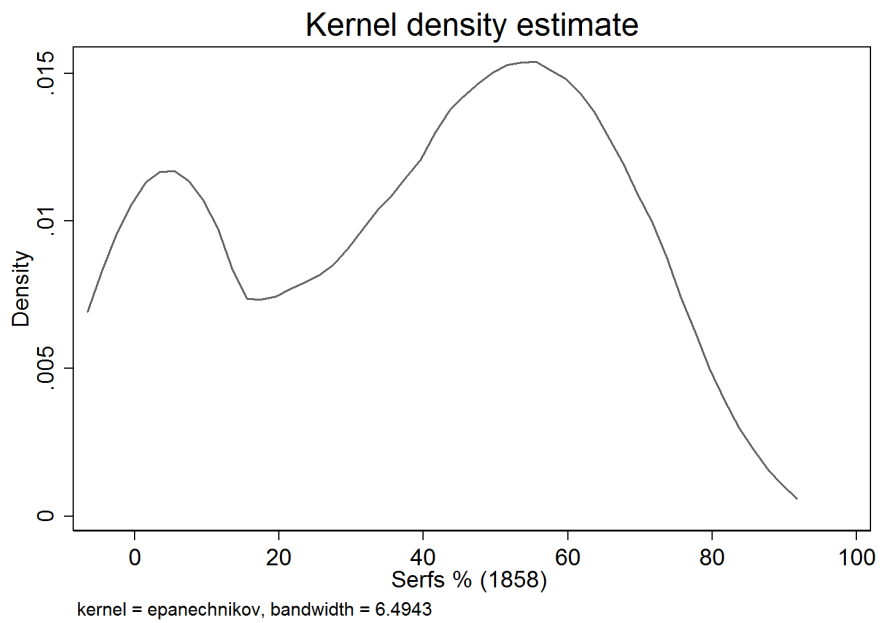


NOTES: This figure plots the distribution of average serfdom across districts.

## A.8 Spatial Distribution of Primary Sampling Units

FIGURE A2: LOCATION OF LITS PRIMARY SAMPLING UNITS (PSU)



NOTES: This figure plots the location of the LiTS Primary Sampling Units, as well as historical district and province borders.

## B   Assessing Bias with the Omnibus Test

To test for a bias deriving from the potential association of the determinants of historical serfdom with economic development today, we perform an omnibus test (similar to Satyanath et al. (2017)). To do this, we estimate a first-stage regression that regresses economic development today on all controls and province fixed effects, thereby generating predicted measures of economic development. In the second stage, we estimate the relationship between the predicted measure of economic development and historical serfdom. Thus, the omnibus test examines whether in the second stage, the variation in development that stems from the geographic controls is correlated with historical serfdom. To guarantee the widest geographic coverage, we measure economic development with light density at the district-level in 2008.

The results are reported in Table B1. The dependent variable is predicted light density obtained from regressing light density on the full set of geographic controls (either linear or flexible), the distance to the provincial capital and the nearest cities, as well as province fixed effects. Overall, we do not find any association between historical serfdom and predicted light-density, both without and with additionally controlling for province fixed effects. The magnitudes are positive without province fixed effects, and negative, but very small compared to the original effect (-0.8), conditional on province fixed effects. The 95% confidence intervals imply that maximum value of the coefficients are between -0.45 and - 0.5. Thus, the omnibus test dismisses a bias larger than two-thirds in our original result.

TABLE B1: OMNIBUS TEST

| | Predicted Log-Light Density (2008) | | | |
|---|---|---|---|---|
| Controls used 1st Stage: | Linear Controls | | Flexible Controls | |
| | (1) | (2) | (3) | (4) |
| Serfs % (1858) | 0.200 | -0.066 | 0.196 | -0.077 |
| | (0.314) | (0.190) | (0.326) | (0.207) |
| | [-0.432,0.832] | [-0.449,0.316] | [-0.459,0.852] | [-0.492,0.339] |
| Fixed Effects | | Province | | Province |
| R-squared | 0.00 | 0.82 | 0.00 | 0.76 |
| Observations | 490 | 490 | 490 | 490 |

NOTE: The unit of observation is the district. The dependent variable is predicted log light density in 2008. The omnibus test regresses in a first stage log light density in 2008 on either the full set of geographic controls, the distance to the provincial capital and the nearest cities, and province fixed effects. The predicted light density is then regressed on serfdom, either unconditional or conditional on province fixed effects. The omnibus test examines whether the variation in light density, that is determined by the set of controls used, is associated with serfdom. Heteroscedastic-robust standard errors in parentheses, clustered at the province.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# C   Additional Results: Households and Individuals

## C.1   Additional Robustness Checks

This section performs additional robustness checks of the long-term effect of serfdom on household well-being today.

Figure C1 illustrates the conditional relationship between household expenditure and serfdom.

Table C1 shows that adjusting standard errors for spatial dependence using a cutoff distance of 250km is the most conservative approach, i.e. produces the largest standard errors.

Table C2 reports the coefficients of all geographic controls.

Table C3 tests whether the main effect of serfdom on household consumption today is sensitive to the inclusion of the set of household controls, some of which could be endogenous to the historical process.

Table C4 tests sensitivity to controlling for wheat, rye, barley, and oat suitability instead of the combined cereal suitability index.

Table C5 replicates the main results with consumer good ownership as dependent variable.

Table C6 replicates the main results with household expenditure calculated from a wider range of items in 2006.

Table C7 tests whether the main results are robust to controlling for whether the PSU is classified as rural or urban settlement. Since this rural/urban status an endogenous control, and likely part of the agglomeration mechanism that we uncover, it is not included in the base set of household controls. Table C7 shows that the main results are robust to its inclusion.

Figure C2 shows strong positive correlations between average district-level night-time lights in 2008 and average expenditure, as well as average consumer good ownership, conditional on either country or province dummies.

FIGURE C1: EFFECT OF SERFDOM ON HOUSEHOLD EXPENDITURE
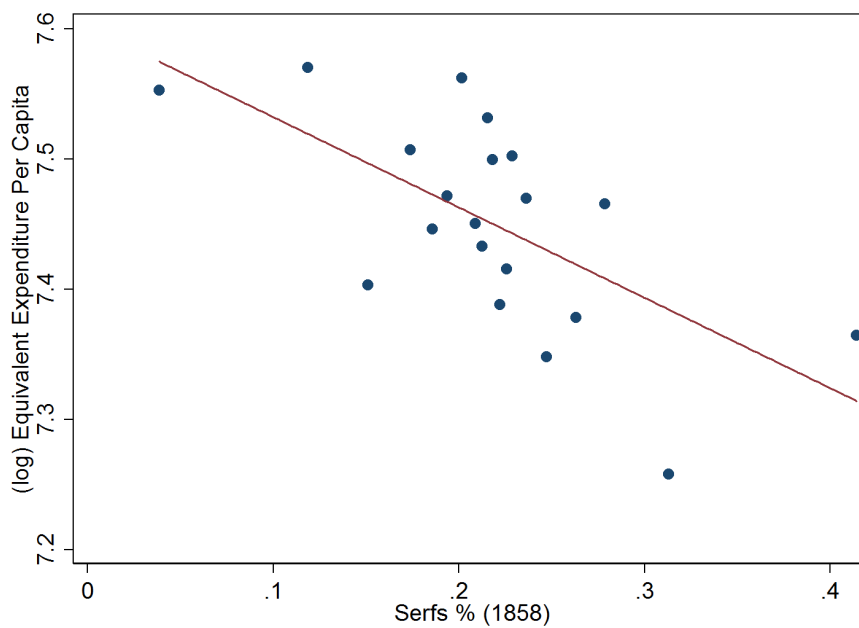
NOTES: This figure plots the relationship between household expenditure and serf-dom, conditional on the full set of household household, flexible geographic controls, and province fixed effects. Households are grouped into 20 bins.

TABLE C1: ADJUSTING S.E. FOR SPATIAL DEPENDENCE: ALTERNATIVE CUTOFF DISTANCES

| Cutoff in km | (ln) Equivalent Expenditures Per Capita | | | | | |
|---|---|---|---|---|---|---|
| | 150 | 200 | 250 | 300 | 350 | 400 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Serfs % (1858) | -0.644*** | -0.644*** | -0.644*** | -0.644*** | -0.644*** | -0.644*** |
| | (0.093) | (0.080) | (0.128) | (0.125) | (0.085) | (0.097) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Flexible Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province |
| Observations | 17155 | 17155 | 17155 | 17155 | 17155 | 17155 |
| R-squared | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 | 0.41 |

NOTE: The unit of observation is the household. Household controls include the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Standard errors adjusted for spatial dependence reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## TABLE C2: MAIN RESULTS W/ ALL GEOGRAPHIC CONTROLS REPORTED

| | (ln) Equivalent Expenditures Per Capita | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Serfs % (1858) | -0.373*** | -0.431*** | -0.379*** | -0.677*** | -0.694*** | -0.644*** |
| | (0.117) | (0.111) | (0.104) | (0.185) | (0.190) | (0.185) |
| Latitude | 0.042 | 0.002 | -0.010 | 0.185*** | 0.145*** | 0.170*** |
| | (0.026) | (0.016) | (0.019) | (0.064) | (0.053) | (0.055) |
| Longitude | 0.002 | 0.007 | 0.007 | 0.016 | 0.013 | 0.034 |
| | (0.009) | (0.009) | (0.009) | (0.019) | (0.025) | (0.023) |
| Distance to Moscow | 0.002 | -0.008 | -0.009 | 0.048 | 0.052 | 0.079** |
| | (0.015) | (0.009) | (0.011) | (0.055) | (0.035) | (0.037) |
| Distance to Coast | -0.026* | -0.028 | -0.023 | -0.074 | -0.025 | -0.022 |
| | (0.015) | (0.017) | (0.020) | (0.047) | (0.049) | (0.048) |
| Forest Cover | -0.001 | -0.003 | -0.004* | 0.002 | -0.002 | -0.001 |
| | (0.002) | (0.002) | (0.002) | (0.003) | (0.003) | (0.003) |
| Ruggedness | -0.020*** | -0.026*** | -0.022*** | -0.013* | -0.023*** | -0.016* |
| | (0.006) | (0.006) | (0.007) | (0.007) | (0.007) | (0.009) |
| River Density | 0.428 | | | -2.362 | | |
| | (3.661) | | | (3.480) | | |
| Cereal Suitability | -0.016 | | | -0.038 | | |
| | (0.027) | | | (0.040) | | |
| Mean Temperature Apr-Sep | 1.155* | | | 1.766* | | |
| | (0.591) | | | (0.900) | | |
| Mean Precipitation Apr-Sep | 0.011 | | | 0.011 | | |
| | (0.008) | | | (0.015) | | |
| Podzol Soil | 0.332** | | | -0.066 | | |
| | (0.142) | | | (0.178) | | |
| River Density Quartile 2 | | 0.005 | -0.006 | | -0.025 | -0.042 |
| | | (0.044) | (0.037) | | (0.047) | (0.051) |
| River Density Quartile 3 | | 0.162*** | 0.143*** | | 0.056 | 0.051 |
| | | (0.056) | (0.049) | | (0.076) | (0.074) |
| River Density Quartile 4 | | 0.066 | 0.047 | | -0.046 | -0.023 |
| | | (0.073) | (0.081) | | (0.097) | (0.106) |
| Mean Precipitation apr-sep Quartile 2 | | -0.139 | -0.110 | | -0.183 | -0.099 |
| | | (0.134) | (0.118) | | (0.141) | (0.142) |
| Mean Precipitation apr-sep Quartile 3 | | 0.257** | 0.257** | | 0.296 | 0.245 |
| | | (0.105) | (0.104) | | (0.201) | (0.197) |
| Mean Precipitation apr-sep Quartile 4 | | 0.235 | 0.227* | | 0.329 | 0.285 |
| | | (0.140) | (0.134) | | (0.233) | (0.223) |
| Mean Temperature apr-sep Quartile 2 | | 0.084 | 0.049 | | 0.132* | 0.067 |
| | | (0.051) | (0.067) | | (0.076) | (0.095) |
| Mean Temperature apr-sep Quartile 3 | | 0.014 | -0.038 | | 0.042 | 0.034 |
| | | (0.114) | (0.140) | | (0.220) | (0.225) |
| Mean Temperature apr-sep Quartile 4 | | -0.020 | -0.100 | | 0.228 | 0.163 |
| | | (0.146) | (0.154) | | (0.215) | (0.209) |
| Very marginal suitabality | | 0.563*** | 0.185 | | 1.273*** | 0.877* |
| | | (0.175) | (0.335) | | (0.369) | (0.459) |
| Marginal suitability | | 0.658** | 0.271 | | 3.506*** | 3.127*** |
| | | (0.267) | (0.456) | | (0.923) | (0.997) |
| Moderate suitability | | 0.549** | 0.242 | | 3.365*** | 2.971*** |
| | | (0.271) | (0.421) | | (0.933) | (1.010) |
| Medium suitability | | 0.558* | 0.235 | | 3.369*** | 2.917*** |
| | | (0.279) | (0.434) | | (0.933) | (1.014) |
| Good suitability | | 0.521* | 0.152 | | 3.268*** | 2.772** |
| | | (0.263) | (0.444) | | (0.937) | (1.032) |
| High suitability | | 0.635** | 0.263 | | 3.337*** | 2.840*** |
| | | (0.284) | (0.468) | | (0.974) | (1.051) |
| Very high suitability | | 0.581** | 0.195 | | 3.235*** | 2.789** |
| | | (0.274) | (0.440) | | (0.972) | (1.039) |
| Podzol Soil Quartile 2 | | -0.055 | -0.026 | | -0.076 | -0.028 |
| | | (0.078) | (0.073) | | (0.068) | (0.067) |
| Podzol Soil Quartile 3 | | 0.152 | 0.167 | | -0.109 | -0.067 |
| | | (0.105) | (0.102) | | (0.096) | (0.108) |
| Podzol Soil Quartile 4 | | 0.158 | 0.192 | | -0.227* | -0.170 |
| | | (0.121) | (0.119) | | (0.119) | (0.132) |
| Distance City in 1600 | | | 20.544 | | | -54.487 |
| | | | (21.718) | | | (40.700) |
| Distance Provincial Capital | | | -0.062 | | | -0.055 |
| | | | (0.038) | | | (0.045) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Country | Country | Country | Province | Province | Province |
| Observations | 17155 | 17155 | 17155 | 17155 | 17155 | 17155 |
| R-squared | 0.39 | 0.40 | 0.41 | 0.40 | 0.41 | 0.41 |
| Number of Clusters | 45 | 45 | 45 | 45 | 45 | 45 |
| $\delta$ for $\beta = 0$ | 16.126 | 9.856 | 2.486 | 2.772 | 1.518 | 1.166 |
| Lower Bound Estimates | -0.462 | -0.626 | -0.465 | -0.530 | -0.606 | -0.444 |

NOTE: The unit of observation is the household. Household controls include the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

20

## TABLE C3: SENSITIVITY TO HOUSEHOLD CONTROLS

| | (ln) Equivalent Expenditures Per Capita | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Serfs % (1858) | -0.413*** | -0.409*** | -0.395*** | -0.395*** | -0.394*** | -0.729*** | -0.721*** | -0.696*** | -0.702*** | -0.718*** |
| | (0.132) | (0.131) | (0.122) | (0.122) | (0.124) | (0.184) | (0.187) | (0.183) | (0.181) | (0.186) |
| Household Size | | 0.047*** | -0.023** | -0.029*** | -0.028*** | | 0.046*** | -0.021** | -0.028*** | -0.027*** |
| | | (0.008) | (0.010) | (0.009) | (0.009) | | (0.008) | (0.010) | (0.009) | (0.009) |
| Share Old | | | -0.472*** | -0.455*** | -0.456*** | | | -0.461*** | -0.443*** | -0.444*** |
| | | | (0.030) | (0.032) | (0.032) | | | (0.031) | (0.033) | (0.033) |
| Share Young | | | 0.003 | 0.029 | 0.027 | | | -0.010 | 0.018 | 0.016 |
| | | | (0.051) | (0.047) | (0.048) | | | (0.049) | (0.045) | (0.046) |
| Share Male | | | | 0.128*** | 0.133*** | | | | 0.134*** | 0.138*** |
| | | | | (0.023) | (0.023) | | | | (0.023) | (0.023) |
| No Religion | | | | | 0.067 | | | | | 0.073 |
| | | | | | (0.056) | | | | | (0.056) |
| Buddhist Religion | | | | | -0.297** | | | | | -0.332*** |
| | | | | | (0.120) | | | | | (0.115) |
| Christian Religion | | | | | 0.064 | | | | | 0.062 |
| | | | | | (0.051) | | | | | (0.050) |
| Jewish Religion | | | | | 0.256 | | | | | 0.238 |
| | | | | | (0.165) | | | | | (0.154) |
| Muslim Religion | | | | | 0.052 | | | | | 0.041 |
| | | | | | (0.079) | | | | | (0.079) |
| Linear Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Country | Country | Country | Country | Country | Province | Province | Province | Province | Province |
| Observations | 17333 | 17333 | 17333 | 17333 | 17155 | 17333 | 17333 | 17333 | 17333 | 17155 |
| R-squared | 0.33 | 0.33 | 0.38 | 0.38 | 0.39 | 0.33 | 0.34 | 0.38 | 0.38 | 0.39 |
| Number of Clusters | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 |

NOTE: The unit of observation is the household. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## TABLE C4: CROP SUITABILITY

| | (ln) Equivalent Expenditures Per Capita | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Wheat | | Rye | | Barley | | Oat | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Serfs % (1858) | -0.566*** | -0.558*** | -0.561*** | -0.573*** | -0.583*** | -0.560*** | -0.563*** | -0.510** |
| | (0.178) | (0.187) | (0.178) | (0.209) | (0.175) | (0.194) | (0.179) | (0.201) |
| Distance City in 1600 | -82.511* | -65.240* | -83.750* | -64.927 | -81.862* | -61.916 | -82.360* | -63.217 |
| | (44.101) | (38.809) | (43.984) | (40.687) | (44.461) | (42.260) | (43.989) | (38.849) |
| Distance Provincial Capital | -0.071 | -0.041 | -0.069 | -0.038 | -0.066 | -0.037 | -0.069 | -0.040 |
| | (0.044) | (0.042) | (0.042) | (0.038) | (0.044) | (0.043) | (0.044) | (0.039) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | | ✓ | | ✓ | | ✓ | |
| Flexible Controls | | ✓ | | ✓ | | ✓ | | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 17155 | 17155 | 17155 | 17155 | 17155 | 17155 | 17155 | 17155 |
| R-squared | 0.40 | 0.41 | 0.40 | 0.41 | 0.40 | 0.41 | 0.40 | 0.41 |
| Number of Clusters | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 |

NOTE: The unit of observation is the household. Household controls include the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE C5: MAIN RESULTS WITH CONSUMER GOODS

| | Consumer Goods | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Serfs % (1858) | -0.299*** | -0.327*** | -0.281*** | -0.496*** | -0.461*** | -0.445*** |
| | (0.096) | (0.076) | (0.082) | (0.169) | (0.150) | (0.144) |
| Distance City in 1600 | | | -17.438 | | | -47.478 |
| | | | (15.972) | | | (31.467) |
| Distance Provincial Capital | | | -0.026 | | | -0.023 |
| | | | (0.016) | | | (0.028) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | | | ✓ | | |
| Flexible Controls | | ✓ | ✓ | | ✓ | ✓ |
| Fixed Effects | Country | Country | Country | Province | Province | Province |
| Observations | 21734 | 21734 | 21734 | 21734 | 21734 | 21734 |
| R-squared | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 | 0.43 |
| Number of Clusters | 45 | 45 | 45 | 45 | 45 | 45 |
| $\delta$ for $\beta = 0$ | -32.802 | -113.289 | 13.079 | 5.225 | 2.860 | 2.453 |
| Lower Bound Estimates | -0.321 | -0.359 | -0.297 | -0.431 | -0.384 | -0.361 |

*Conley S.E. 250km*

| | | | | | | |
|---|---|---|---|---|---|---|
| Serfs % (1858) | [0.064]*** | [0.059]*** | [0.061]*** | [0.126]*** | [0.116]*** | [0.100]*** |

NOTE: The unit of observation is the household. Household controls include the household size, the share of household members aged 0-18, the share of household members aged 60+, the share of male household members, the religious denomination of the household respondent, LiTS wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for country/province fixed effects. Standard errors clustered at the province are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

#### TABLE C6: HOUSEHOLD EXPENDITURE WAVES 2006

| | (ln) Equivalent Expenditures Per Capita | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Serfs % (1858) | -0.543*** | -0.480*** | -0.413** | -0.952*** | -0.946*** | -0.890*** |
| | (0.159) | (0.157) | (0.173) | (0.276) | (0.236) | (0.243) |
| Distance City in 1600 | | | 18.360 | | | -85.886 |
| | | | (29.275) | | | (62.035) |
| Distance Provincial Capital | | | -0.109** | | | -0.027 |
| | | | (0.051) | | | (0.057) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | | | ✓ | | |
| Flexible Controls | | ✓ | ✓ | | ✓ | ✓ |
| Fixed Effects | Country | Country | Country | Province | Province | Province |
| Observations | 5721 | 5721 | 5721 | 5721 | 5721 | 5721 |
| R-squared | 0.40 | 0.41 | 0.41 | 0.42 | 0.44 | 0.44 |
| Number of Clusters | 38 | 38 | 38 | 38 | 38 | 38 |
| $\delta$ for $\beta = 0$ | -34.045 | 4.290 | 2.398 | 2.585 | 1.727 | 1.375 |
| Lower Bound Estimates | -0.617 | -0.509 | -0.399 | -0.772 | -0.772 | -0.672 |
| *Conley S.E. 250km* | | | | | | |
| Serfs % (1858) | [0.159]*** | [0.166]*** | [0.165]** | [0.295]*** | [0.249]*** | [0.217]*** |

NOTE: The unit of observation is the household. Household controls include the household size, the share of household members aged 0-18, the share of household members aged 60+, the share of male household members, the religious denomination of the household respondent, LiTS wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for country/province fixed effects. Standard errors clustered at the province are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE C7: MAIN RESULTS CONTROLLING FOR URBAN/RURAL STATUS

| | (ln) Equivalent Expenditures Per Capita | | | | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Serfs % (1858) | -0.329*** | -0.382*** | -0.347*** | -0.621*** | -0.619*** | -0.577*** |
| | (0.113) | (0.101) | (0.096) | (0.167) | (0.170) | (0.163) |
| Distance City in 1600 | | | 14.413 | | | -54.832 |
| | | | (19.022) | | | (38.977) |
| Distance Provincial Capital | | | -0.044 | | | -0.048 |
| | | | (0.036) | | | (0.039) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | | | ✓ | | |
| Flexible Controls | | ✓ | ✓ | | ✓ | ✓ |
| Fixed Effects | Country | Country | Country | Province | Province | Province |
| Observations | 17155 | 17155 | 17155 | 17155 | 17155 | 17155 |
| R-squared | 0.41 | 0.42 | 0.42 | 0.41 | 0.42 | 0.42 |
| Number of Clusters | 45 | 45 | 45 | 45 | 45 | 45 |
| $\delta$ for $\beta = 0$ | 4.367 | 3.960 | 2.002 | 2.112 | 1.221 | 0.990 |
| Lower Bound Estimates | -0.320 | -0.435 | -0.359 | -0.465 | -0.470 | -0.375 |

*Conley S.E. 250km*

| | | | | | | |
|---|---|---|---|---|---|---|
| Serfs % (1858) | [0.111]*** | [0.104]*** | [0.098]*** | [0.153]*** | [0.131]*** | [0.128]*** |

NOTE: The unit of observation is the household. Household controls include the household size, the share of household members aged 0-18, the share of household members aged 60+, the share of male household members, the religious denomination of the household respondent, LiTS wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for country/province fixed effects. Standard errors clustered at the province are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# FIGURE C2: NIGHT-TIME LIGHTS, HOUSEHOLD EXPENDITURE, AND CONSUMER GOODS
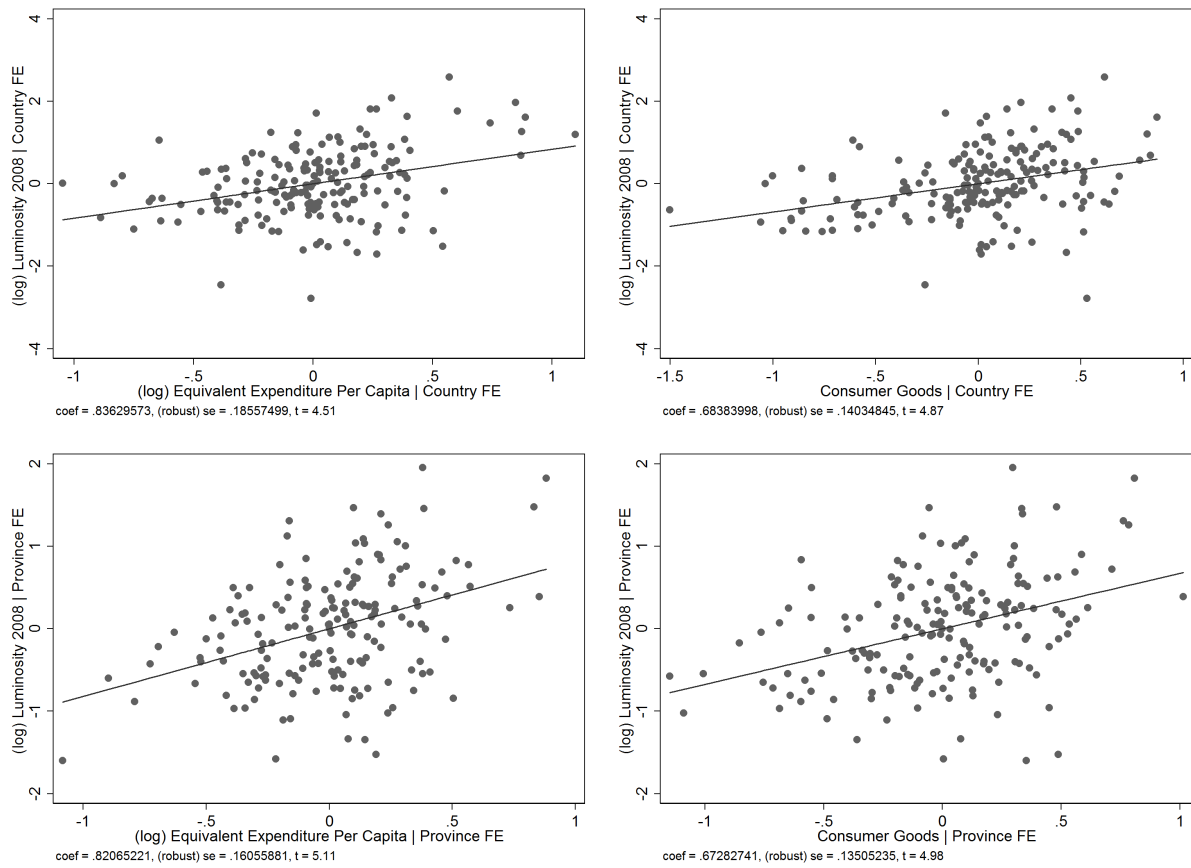


NOTES: These figures plot average district-level night-time luminosity in 2008 against average household expenditure and average consumer goods ownership, conditional on country fixed effects (upper panel), and province fixed effects (lower panel).

### C.2   Cultural Attitudes and Preferences

Did economic exploitation over several centuries shape peoples beliefs and attitudes, perhaps fostering a "culture of serfdom," with persistent implications for economic development? Several recent studies have documented that institutions can impact cultural norms in the long-run (for an overview see Nunn (2012)), which can persist over generations. Moreover, it is possible that various institutional restrictions, social and economic inequality, or persistent limitations on urban development under and after serfdom generated long-lasting norms and beliefs that undermined income growth in modern post-Soviet economies. We report in the main text that the persisting (spatial) economic inequality caused by historical serfdom is also reflected in stronger preferences for redistribution, and a greater willingness to engage in collective action.

In Table C8 we explore this possibility further and consider inter-personal trust, as well as preferences for a market economy, and democracy, as cultural attitudes that are likely associated with economic growth (e.g. Algan and Cahuc, 2013). Overall, we find no evidence that these economically relevant cultural outcomes are associated to serfdom. Trust in others showed no relationship to historical serfdom. Moreover, preferences for a market economy (versus a planned economy) or for democracy (versus autocracy) are not statistically different between areas with a greater or lesser history of serfdom. While cultural channels have been emphasized in the literature on persistent effects of past labor coercion (i.e. Nunn and Wantchekon (2011)), we find no support that culture is a mechanism from serfdom to growth in the Russian and former Soviet case. The results on redistribution and collective action suggest that cultural attitudes are rather an additional reflection of the underlying economic conditions that serfdom helped to create.

In Table C11 we test whether respondents believe that luck is the main reason for poverty and investigate differences in preferences for income redistribution.

Table C9 documents that serfdom does not have any effect on membership in the communist party.

In Table C10 we investigate whether places where serfdom was more widespread are more hostile towards outsiders. Respondents were asked whether they would accept a neighbor with a different race, someone with an immigration background, someone with a different religion, someone who is Jewish, or someone that speaks a different language. Individuals living in areas where serfdom was higher in the past only show a higher aversion towards Jewish neighbors, but are not different otherwise.

27

TABLE C8: TRUST, AND ATTITUDES TOWARDS ECONOMIC AND POLITICAL INSTITUTIONS

| | Trust | | | Prefer Market Economy | | | Prefer Democracy | | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Serfs % (1858) | 0.028 | 0.031 | 0.078 | -0.015 | -0.006 | -0.002 | 0.052 | 0.093 | 0.118 |
| | (0.108) | (0.110) | (0.090) | (0.092) | (0.096) | (0.107) | (0.110) | (0.113) | (0.114) |
| Distance City in 1600 | | -4.133 | 3.928 | | -18.215 | -18.730 | | -6.588 | -13.123 |
| | | (15.321) | (19.508) | | (19.520) | (20.325) | | (17.063) | (20.051) |
| Distance Provincial Capital | | -0.002 | 0.001 | | -0.005 | 0.009 | | -0.024 | -0.018 |
| | | (0.014) | (0.016) | | (0.019) | (0.022) | | (0.015) | (0.019) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | ✓ | | ✓ | ✓ | | ✓ | ✓ | |
| Flexible Controls | | | ✓ | | | ✓ | | | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 21218 | 21218 | 21218 | 19609 | 19609 | 19609 | 19811 | 19811 | 19811 |
| R-squared | 0.03 | 0.03 | 0.03 | 0.06 | 0.06 | 0.06 | 0.05 | 0.05 | 0.05 |
| Number of Clusters | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 |

NOTE: The unit of observation is the individual. Household controls include the age, age squared, and gender of the respondent, the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### TABLE C9: COMMUNISM

| | At Least One Communist Family Member | |
|---|---|---|
| | (1) | (2) |
| Serfs % (1858) | 0.005 | -0.175 |
| | (0.164) | (0.178) |
| Distance City in 1600 | -26.602 | -64.766** |
| | (23.893) | (30.046) |
| Distance Provincial Capital | 0.028 | 0.003 |
| | (0.023) | (0.026) |
| Household Controls | ✓ | ✓ |
| Linear Controls | ✓ | |
| Flexible Controls | | ✓ |
| Fixed Effects | Province | Province |
| Observations | 8727.00 | 8727.00 |
| R-squared | 0 | 0 |
| Number of Clusters | 39 | 39 |

NOTE: The unit of observation is the individual. Household controls include the age, age squared, and gender of the respondent, the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

At Least One Communist Family Member

(1)                    (2)

## TABLE C10: XENOPHOBIA

| Reject as neighbor someone... | with different race | | who is an immigrant | | with different religion | | who is Jewish | | that speaks a different language | | Average Xenophobia | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Serfs % (1858) | -0.135 | -0.056 | 0.028 | 0.005 | 0.113 | 0.145 | 0.338** | 0.368** | 0.070 | 0.103 | 0.656 | 0.883 |
| | (0.168) | (0.205) | (0.159) | (0.178) | (0.155) | (0.172) | (0.155) | (0.174) | (0.214) | (0.232) | (1.093) | (1.245) |
| Distance City in 1600 | 4.698 | 40.165* | 14.658 | 21.791 | 25.713* | 45.342** | 41.541** | 73.750*** | 8.798 | 39.949** | 134.639 | 313.743*** |
| | (19.277) | (20.598) | (20.816) | (24.458) | (15.195) | (17.885) | (18.850) | (20.155) | (17.014) | (18.492) | (98.728) | (107.959) |
| Distance Provincial Capital | 0.032* | 0.047* | 0.034 | 0.051* | -0.003 | 0.001 | -0.004 | 0.003 | -0.013 | -0.004 | 0.037 | 0.103 |
| | (0.017) | (0.026) | (0.020) | (0.028) | (0.013) | (0.016) | (0.013) | (0.018) | (0.015) | (0.021) | (0.089) | (0.133) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | |
| Flexible Controls | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 8727 | 8727 | 8727 | 8727 | 8727 | 8727 | 8727 | 8727 | 8727 | 8727 | 8727 | 8727 |
| R-squared | 0.05 | 0.05 | 0.09 | 0.09 | 0.03 | 0.04 | 0.05 | 0.05 | 0.06 | 0.07 | 0.05 | 0.06 |
| Number of Clusters | 39 | 39 | 39 | 39 | 39 | 39 | 39 | 39 | 39 | 39 | 39 | 39 |

NOTE: The unit of observation is the individual. Household controls include the age, age squared, and gender of the respondent, the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE C11: PREFERENCES FOR REDISTRIBUTION

| | Poor Unlucky? | | Less or More Inequality? | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Serfs % (1858) | 0.051 | 0.087** | -1.255 | -1.869** |
| | (0.038) | (0.042) | (0.864) | (0.878) |
| Distance City in 1600 | -2.917 | -7.433 | -61.030 | 24.334 |
| | (8.968) | (8.428) | (144.045) | (127.606) |
| Distance Provincial Capital | 0.006 | 0.015** | -0.008 | 0.027 |
| | (0.006) | (0.007) | (0.117) | (0.121) |
| Household Controls | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | | ✓ | |
| Flexible Controls | | ✓ | | ✓ |
| Fixed Effects | Province | Province | Province | Province |
| Observations | 13000 | 13000 | 14946 | 14946 |
| R-squared | 0.03 | 0.03 | 0.08 | 0.09 |
| Number of Clusters | 45 | 45 | 40 | 40 |
| $\delta$ for $\beta = 0$ | 1.936 | -16.709 | 7.205 | -12.595 |
| Lower Bound Estimates | 0.043 | 0.114 | -1.798 | -2.642 |
| *Conley S.E. 250km* | | | | |
| Serfs % (1858) | [0.035] | [0.040]** | [0.623]** | [0.643]*** |

NOTE: The unit of observation is an individual. Household controls include the age, age squared, and gender of the respondent, the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### C.3 Protest and Collective Action

Table C12 investigates the effects of serfdom on the willingness to protest and engage in collective action.

TABLE C12: PROTEST AND COLLECTIVE ACTION

| | Demonstrated | | Striked | | Sign Petition | | Average Protest | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Serfs % (1858) | 0.118 | 0.089 | 0.116 | 0.128 | 0.080 | 0.059 | 0.105 | 0.092 |
| | (0.093) | (0.084) | (0.087) | (0.085) | (0.091) | (0.105) | (0.083) | (0.085) |
| Distance City in 1600 | | -38.821** | | -22.253 | | -30.981** | | -30.685** |
| | | (15.935) | | (15.271) | | (14.226) | | (13.404) |
| Distance Provincial Capital | | -0.001 | | 0.000 | | -0.000 | | -0.000 |
| | | (0.019) | | (0.018) | | (0.019) | | (0.017) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | | ✓ | | ✓ | | ✓ | |
| Flexible Controls | | ✓ | | ✓ | | ✓ | | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 20731 | 20731 | 20731 | 20731 | 20731 | 20731 | 20731 | 20731 |
| R-squared | 0.07 | 0.07 | 0.09 | 0.10 | 0.09 | 0.10 | 0.10 | 0.11 |
| Number of Clusters | 45 | 45 | 45 | 45 | 45 | 45 | 45 | 45 |

NOTE: The unit of observation is the individual. Household controls include the age, age squared, and gender of the respondent, the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## C.4 Inequality and Public Goods Provision

Table C13 investigates differences in historical and contemporary inequality. To measure contemporary inequality, we constructed Gini indices using expenditure and consumer good ownership of households living in a common historical district.

Table C14 investigates the effects of serfdom on the availability of public goods.

### TABLE C13: INEQUALITY

| | Historical | | | | Modern (2006-2016) | | | | | |
| | Land Owned by Nobles 1905 | | Land Inequality 1905 | | Gini Expenditure | | | Gini Consumer Goods | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| Serfs % (1858) | 28.099*** | 27.970*** | 0.323*** | 0.308*** | -0.008 | -0.006 | -0.006 | 0.018 | 0.030 | 0.041 |
| | (4.929) | (4.704) | (0.052) | (0.049) | (0.008) | (0.007) | (0.008) | (0.035) | (0.036) | (0.032) |
| Distance City in 1600 | 0.170 | 2.142 | -0.208 | -0.176 | -0.023 | -0.019 | -0.029 | -0.023 | -0.038 | -0.037 |
| | (9.977) | (9.449) | (0.144) | (0.139) | (0.022) | (0.024) | (0.024) | (0.062) | (0.086) | (0.093) |
| Distance Provincial Capital | -1.627** | -1.728** | -0.010 | -0.010 | 0.001 | 0.001 | 0.002 | 0.009 | 0.008 | 0.007 |
| | (0.654) | (0.674) | (0.009) | (0.009) | (0.002) | (0.002) | (0.002) | (0.006) | (0.006) | (0.006) |
| Land Gini 1905 | | | | | | | -0.002 | | | -0.046 |
| | | | | | | | (0.012) | | | (0.039) |
| Linear Controls | ✓ | | ✓ | | ✓ | | | ✓ | | |
| Flexible Controls | | ✓ | | ✓ | | ✓ | ✓ | | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 490 | 490 | 467 | 467 | 183 | 183 | 179 | 185 | 185 | 180 |
| R-squared | 0.79 | 0.80 | 0.71 | 0.72 | 0.55 | 0.62 | 0.62 | 0.35 | 0.38 | 0.38 |
| Number of Clusters | 50 | 50 | 49 | 49 | 45 | 45 | 43 | 45 | 45 | 43 |
| $\delta$ for $\beta = 0$ | 1.309 | 1.199 | -29.931 | 3.762 | 1.562 | 0.803 | 0.792 | 1.419 | 32.802 | -17.028 |
| Lower Bound Estimates | 29.385 | 28.007 | 0.412 | 0.343 | -0.008 | -0.002 | -0.002 | 0.009 | 0.034 | 0.055 |
| *Conley S.E. 300km* | | | | | | | | | | |
| Serfs % (1858) | [4.736]*** | [4.556]*** | [0.060]*** | [0.054]*** | [0.006] | [0.005] | [0.007] | [0.033] | [0.030] | [0.027] |

NOTE: The unit of observation is a district. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## TABLE C14: PUBLIC GOODS

| Availability of: | Water | | Landline | | Heating | | Sewage | | PC Public Goods | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Serfs % (1858) | -0.044 | -0.011 | -0.174 | -0.315** | -0.059 | -0.062 | -0.203 | -0.242 | -0.849 | -1.094* |
| | (0.123) | (0.106) | (0.162) | (0.122) | (0.188) | (0.168) | (0.185) | (0.191) | (0.678) | (0.638) |
| Distance City in 1600 | -25.205 | -26.043 | -0.616 | 5.114 | -34.042 | -14.100 | -46.265 | -40.014 | -134.270 | -126.614 |
| | (20.139) | (28.440) | (23.730) | (23.089) | (27.289) | (30.598) | (30.611) | (36.998) | (112.974) | (154.235) |
| Distance Provincial Capital | -0.059*** | -0.046** | -0.046** | -0.045** | -0.112*** | -0.070** | -0.136*** | -0.074 | -0.567*** | -0.337* |
| | (0.020) | (0.019) | (0.023) | (0.018) | (0.029) | (0.029) | (0.043) | (0.048) | (0.158) | (0.169) |
| Household Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Linear Controls | ✓ | | ✓ | | ✓ | | ✓ | | ✓ | |
| Flexible Controls | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 21710 | 21710 | 21605 | 21605 | 21490 | 21490 | 6119 | 6119 | 6113 | 6113 |
| R-squared | 0.16 | 0.17 | 0.15 | 0.16 | 0.22 | 0.24 | 0.22 | 0.24 | 0.26 | 0.28 |
| Number of Clusters | 45 | 45 | 45 | 45 | 45 | 45 | 38 | 38 | 38 | 38 |

NOTE: The unit of observation is the household. Household controls include the household size, share of household members aged 0-18, share of household members aged 60+, share of male household members, religious denomination of the household respondent, Lits Survey Wave fixed effects. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as the remaining linear controls. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# D  Additional Results: City Population

## D.1  Cross-Sectional Estimation in the Period 1897 - 2002

Figure D1 illustrates estimated coefficients and confidence intervals of the effect of serfdom on city population obtained from cross-sectional regressions.

Figure D2 plots average city population for cities with above and below median serfdom intensity.

Figure D3 contains coefficients from estimating city population on an above/below median serfdom indicator × year fixed effects, conditional on province dummies × year fixed effects, distance to Moscow × year fixed effects, and cereal suitability × year fixed effects.

In Table D1, we document negative effects of serfdom on city population conditional on linear controls.

In addition, Table D2 documents the absence of any convergence over time. If there was (beta-) convergence in city sizes, we would expect cities with larger populations to grow slower - not faster. Interestingly, the positive relationship between initial population size and growth in population only appears in the Soviet period (i.e., after 1926). This could be an indication of the well-known urban bias of Soviet planners.

Figure D4 plots the corresponding coefficients.

FIGURE D1: ESTIMATED COEFFICIENTS CROSS-SECTION 1897-2002
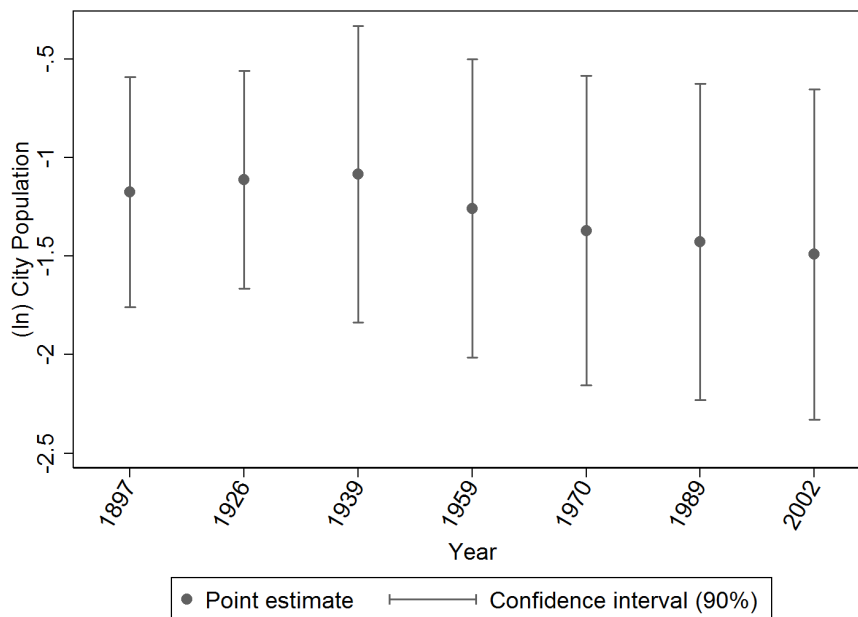


NOTES: This figure plots the coefficient from regressions of (log) city population on serfdom, conditional on province fixed effects and flexible controls. The corresponding regression results are reported in Table 6 of the main text.

FIGURE D2: AVERAGE CITY POPULATION OF CITIES WITH ABOVE/BELOW MEDIAN SERFDOM 1897-2002
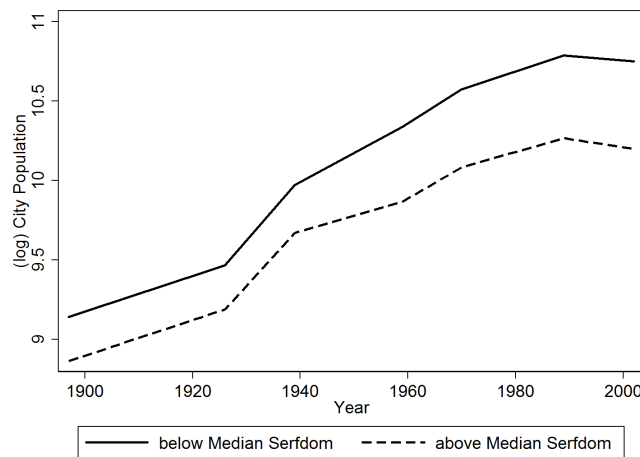


NOTES: This figure plots average city population for cities whose intensity of serfdom is above the median of the distribution (dashed line), and for cities with below median serfdom intensity (solid line). The sample corresponds to the 366 cities used in Table 6.

FIGURE D3: GAP BETWEEN ABOVE & BELOW MEDIAN SERFDOM AREAS, CONDITIONAL ON PROVINCE × YEAR FE 1897-2002
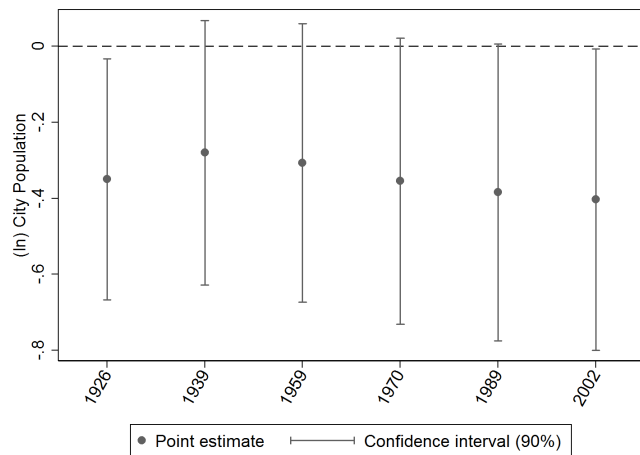


NOTES: This figure plots the coefficients from regressing (log) city population on a dummy for above median serfdom intensity × year fixed effects, conditional on province dummies × year fixed effects, distance to Moscow × year fixed effects, and cereal suitability × year fixed effects. The sample corresponds to the 366 cities used in Table 6.

36

## TABLE D1: PERSISTENCE THROUGH THE SOVIET PERIOD: LINEAR CONTROLS

| Log City Population in | 1897 | 1926 | 1939 | 1959 | 1970 | 1989 | 2002 | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Serfs % (1858) | -1.080*** | -0.929** | -0.791 | -0.953* | -1.061* | -1.127* | -1.196* | -0.036 |
| | (0.349) | (0.348) | (0.494) | (0.534) | (0.565) | (0.601) | (0.613) | (0.558) |
| Log City Population 1897 | | | | | | | | 1.074*** |
| | | | | | | | | (0.065) |
| Linear Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Distances: City & Prov. Capital | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 366 | 366 | 366 | 366 | 366 | 366 | 366 | 366 |
| R-squared | 0.32 | 0.29 | 0.26 | 0.28 | 0.26 | 0.25 | 0.25 | 0.64 |
| Number of Clusters | 33 | 33 | 33 | 33 | 33 | 33 | 33 | 33 |
| $\delta$ for $\beta = 0$ | 6.720 | 13.070 | 14.850 | 13.660 | 11.990 | 12.920 | 11.570 | |
| Lower Bound Estimates | -1.100 | -0.967 | -0.818 | -0.996 | -1.099 | -1.173 | -1.240 | |
| *Conley S.E. 300km* | | | | | | | | |
| Serfs % (1858) | [0.341]*** | [0.356]*** | [0.509] | [0.535]* | [0.572]* | [0.613]* | [0.617]* | [0.487] |

NOTE: The unit of observation is a city. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## TABLE D2: POPULATION GROWTH

| Growth in City Population | 1897-1926 | 1926-1939 | 1939-1959 | 1959-1970 | 1970-1989 | 1989-2002 | 1926-1989 | 1897-2002 |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Serfs % (1858) | 0.031 | 0.085 | -0.132 | -0.070 | -0.011 | -0.051 | -0.214 | -0.250 |
| | (0.164) | (0.221) | (0.111) | (0.052) | (0.075) | (0.031) | (0.357) | (0.478) |
| Log City Population 1897 | -0.027 | | | | | | | 0.056 |
| | (0.029) | | | | | | | (0.060) |
| Log City Population 1926 | | 0.050* | | | | | 0.091 | |
| | | (0.028) | | | | | (0.060) | |
| Log City Population 1939 | | | 0.039** | | | | | |
| | | | (0.017) | | | | | |
| Log City Population 1959 | | | | 0.033** | | | | |
| | | | | (0.013) | | | | |
| Log City Population 1970 | | | | | 0.034** | | | |
| | | | | | (0.014) | | | |
| Log City Population 1989 | | | | | | 0.008** | | |
| | | | | | | (0.004) | | |
| Distance Provincial Capital | -0.087*** | -0.032 | -0.043** | 0.009 | -0.003 | -0.000 | -0.110* | -0.213*** |
| | (0.027) | (0.030) | (0.019) | (0.016) | (0.016) | (0.008) | (0.058) | (0.067) |
| Distance City in 1600 | 0.161 | 0.459 | 0.197 | 0.094 | -0.057 | -0.138* | 0.780 | 0.826 |
| | (0.389) | (0.348) | (0.223) | (0.273) | (0.206) | (0.073) | (0.821) | (0.950) |
| Flexible Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 366 | 366 | 366 | 366 | 366 | 366 | 366 | 366 |
| R-squared | 0.25 | 0.40 | 0.40 | 0.29 | 0.29 | 0.32 | 0.29 | 0.30 |
| Number of Clusters | 33 | 33 | 33 | 33 | 33 | 33 | 33 | 33 |

NOTE: The unit of observation is a city. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as linear controls of latitude and longitude of the district, the area covered by forest, ruggedness, the distance to the coast, and the distance to Moscow. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

FIGURE D4: ESTIMATED COEFFICIENTS POPULATION GROWTH
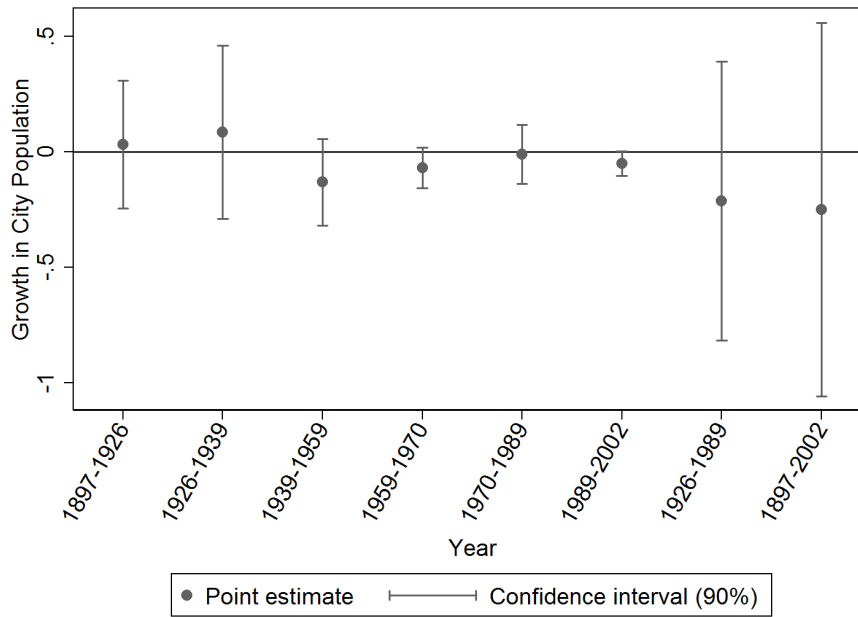
NOTES: This figure plots the coefficient from regressions of average city population growth on serfdom, conditional on province fixed effects and flexible controls. The corresponding regression results are reported in Table D2.

## D.2 Panel Estimation in the Period 1800 - 2002

## D.3 Balanced Sample

Table D4 reports from the flexible estimations. Figure D5 illustrates the corresponding coefficients and confidence intervals.

Figure D6 illustrates coefficients from regressions on a dummy for above median serfdom intensity × year fixed effects, conditional on province dummies × year fixed effects, distance to Moscow × year fixed effects, and cereal suitability × year fixed effects.

Table D4 reports panel estimation with standard errors clustered at the province.

TABLE D3: FLEXIBLE ESTIMATION

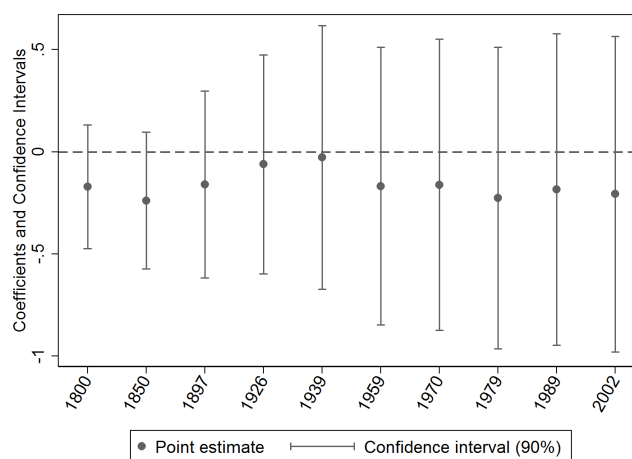|  | Log City Population | |
|---|---|---|
|  | (1) | (2) |
| Serfs % (1858) × Year 1850 | -0.378* | -0.172 |
|  | (0.201) | (0.271) |
| Serfs % (1858) × Year 1897 | -0.720* | 0.200 |
|  | (0.366) | (0.425) |
| Serfs % (1858) × Year 1926 | -1.378*** | 0.039 |
|  | (0.477) | (0.529) |
| Serfs % (1858) × Year 1939 | -1.517*** | 0.097 |
|  | (0.547) | (0.654) |
| Serfs % (1858) × Year 1959 | -1.995*** | -0.208 |
|  | (0.610) | (0.726) |
| Serfs % (1858) × Year 1970 | -2.076*** | -0.213 |
|  | (0.625) | (0.751) |
| Serfs % (1858) × Year 1979 | -2.178*** | -0.302 |
|  | (0.635) | (0.771) |
| Serfs % (1858) × Year 1989 | -2.155*** | -0.265 |
|  | (0.637) | (0.770) |
| Serfs % (1858) × Year 2002 | -2.209*** | -0.313 |
|  | (0.643) | (0.781) |
| Controls × Year FE |  | ✓ |
| Year FE | ✓ | ✓ |
| City FE | ✓ | ✓ |
| Observations | 982 | 982 |
| R-squared | 0.76 | 0.82 |
| Number of Clusters | 99 | 99 |

NOTE: The unit of observation is a city-year. Controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Standard errors clustered at the city in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

FIGURE D5: ESTIMATED COEFFICIENT FLEXIBLE PANEL ESTIMATION 1800-2002
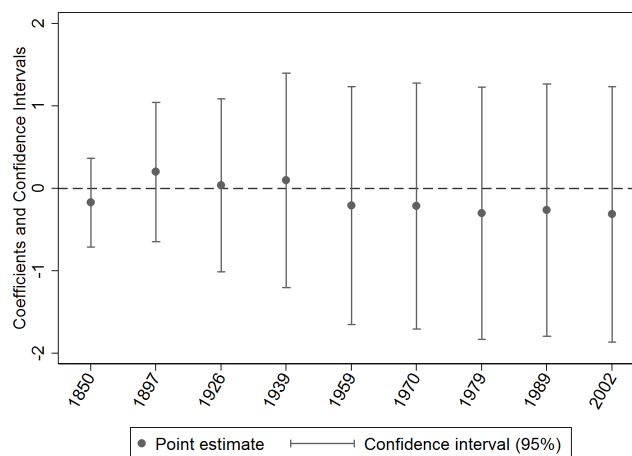


NOTES: This figure plots the coefficients from regressions of city population on serfdom, conditional on city and year fixed effects, and controls × year fixed effects. The corresponding regression results are reported in Column (7) of Table 7.

FIGURE D6: GAP BETWEEN ABOVE & BELOW MEDIAN SERFDOM AREAS, CONDITIONAL ON PROVINCE × YEAR FE 1800-2002



NOTES: This figure plots the coefficients from regressing (log) city population on a dummy for above median serfdom intensity × year fixed effects, conditional on province dummies × year fixed effects, distance to Moscow × year fixed effects, and cereal suitability × year fixed effects. The sample corresponds to the 99 cities used in Table 7.

TABLE D4: PANEL FIXED EFFECTS WITH STANDARD ERRORS CLUSTERED AT PROVINCE

| | Log City Population | | | | | |
| | Full | | 1800-1897 | Full | | |
| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Serfs % (1858) × Post Emancipation | -1.588*** | -0.041 | 0.208 | 0.259 | | |
| | (0.469) | (0.684) | (0.385) | (0.637) | | |
| Serfs % (1858) × Soviet (1922-1991) | | | | -0.398*** | | |
| | | | | (0.127) | | |
| Serfs % (1858) × Year 1850 | | | | | -0.378* | -0.172 |
| | | | | | (0.210) | (0.262) |
| Serfs % (1858) × Year 1897 | | | | | -0.720* | 0.200 |
| | | | | | (0.386) | (0.457) |
| Serfs % (1858) × Year 1926 | | | | | -1.378** | 0.039 |
| | | | | | (0.514) | (0.602) |
| Serfs % (1858) × Year 1939 | | | | | -1.517*** | 0.097 |
| | | | | | (0.536) | (0.787) |
| Serfs % (1858) × Year 1959 | | | | | -1.995*** | -0.208 |
| | | | | | (0.584) | (0.896) |
| Serfs % (1858) × Year 1970 | | | | | -2.076*** | -0.213 |
| | | | | | (0.591) | (0.911) |
| Serfs % (1858) × Year 1979 | | | | | -2.178*** | -0.302 |
| | | | | | (0.605) | (0.934) |
| Serfs % (1858) × Year 1989 | | | | | -2.155*** | -0.265 |
| | | | | | (0.627) | (0.959) |
| Serfs % (1858) × Year 2002 | | | | | -2.209*** | -0.313 |
| | | | | | (0.643) | (0.980) |
| Controls × Post Emancipation | | ✓ | ✓ | ✓ | | |
| Controls × Year FE | | | | | | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| City FE | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 982 | 982 | 294 | 982 | 982 | 982 |
| R-squared | 0.75 | 0.79 | 0.75 | 0.79 | 0.76 | 0.82 |
| Number of Clusters | 33 | 33 | 33 | 33 | 33 | 33 |

NOTE: The unit of observation is the city-year. Controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Standard errors clustered at the province in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### D.4 Unbalanced Sample

The main analysis uses a balanced panel of 99 cities for which population was recorded already back in the 19th century. This balanced sample does not take into account the creation of new cities. If new cities were created in different areas then old cities, and as a result exhibit a different growth performance during the 20th century, and if the creation of new cities was systematically (positively or negatively) associated with serfdom, the omission of these cities could create a sample selection bias. To investigate the likelihood of a sample selection bias, we first check whether cities in the baseline balanced sample are significantly different along the observable characteristics compared to cities out of sample. Table D5 documents that sampled cities are on average larger, located in areas with a higher river density, higher agricultural suitability, higher mean temperature, and in closer proximity to Moscow. These differences are not surprising, given the historical process of city formation in favorable areas around the capital of the Russian Empire.

We then test whether the total number of cities that are not included in the sample per district is significantly related to serfdom. To do this, we count the number of "out of sample" cities per district. Column 1 of Table D6 shows the raw correlation between the number of new cities and serfdom, which suggests that the creation of new cities is negatively associated with the intensity of serfdom. Once we condition on the baseline controls (or province fixed effects), however, there is no longer a significant relationship between serfdom and the number of out of sample cities.

As a next step, we look at the growth performance of sample and non-sample cities in Figures D7 and D8. For this we computed average city population (taking out the mean), and average city growth in both samples. The Figures show that the evolution of population is very similar in both samples during the period 1897-2002. The only period when there is a small growth advantage of non-sampled cities is between 1939 and 1959, but otherwise population co-evolves.

Finally, we test whether the effects of serfdom are different when we use a balanced and an unbalanced sample. For this, we estimate flexible models in the balanced and unbalanced sample in Table D7 (i.e. adding cities in the period 1897 - 2002). The estimated coefficients obtained in both samples have the same sign and are very similar in magnitude. Our overall conclusion does not change: there is no catch-up of cities in areas were serfdom was more widespread in the period 1800 - 2002. These persistent differences in city population are also illustrated in Figure D9. Together, these couple of exercises document that the restriction to a balanced sample did not cause a sample selection bias.

TABLE D5: CHARACTERISTICS OF CITIES IN BALANCED AND UNBALANCED SAMPLE

| | Mean Population | River Density | Coal Territory | Cereal Suitability | Mean Temp. Apr-Sep | Mean Prec. Apr-Sep | Podzol Soil | Distance to Moscow |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Sample City | 1.209*** | 0.003*** | 0.012 | 0.609*** | 0.053* | 0.836 | -0.033 | -1.848** |
| | (0.114) | (0.001) | (0.073) | (0.158) | (0.030) | (1.299) | (0.059) | (0.841) |
| Observations | 683 | 683 | 683 | 683 | 683 | 683 | 683 | 683 |
| R-squared | 0.18 | 0.02 | 0.00 | 0.03 | 0.01 | 0.00 | 0.00 | 0.02 |

NOTE: The unit of observation is the city. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE D6: RELATIONSHIP BETWEEN "NEW CITIES" AND SERFDOM

| | Number of Out of Sample Cities | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Serfs % (1858) | -18.074*** | 5.504 | -1.821 | 2.474 |
| | (6.409) | (10.712) | (7.335) | (7.867) |
| Linear Controls | | ✓ | | ✓ |
| Fixed Effects | | | Province | Province |
| Observations | 282 | 282 | 282 | 282 |
| R-squared | 0.04 | 0.23 | 0.34 | 0.42 |
| Number of Clusters | 34 | 34 | 34 | 34 |

NOTE: The unit of observation is the district. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Heteroskedastic-robust standard errors are in parentheses, clustered at the province. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

FIGURE D7: AVERAGE CITY POPULATION IN SAMPLE AND NON-SAMPLE CITIES
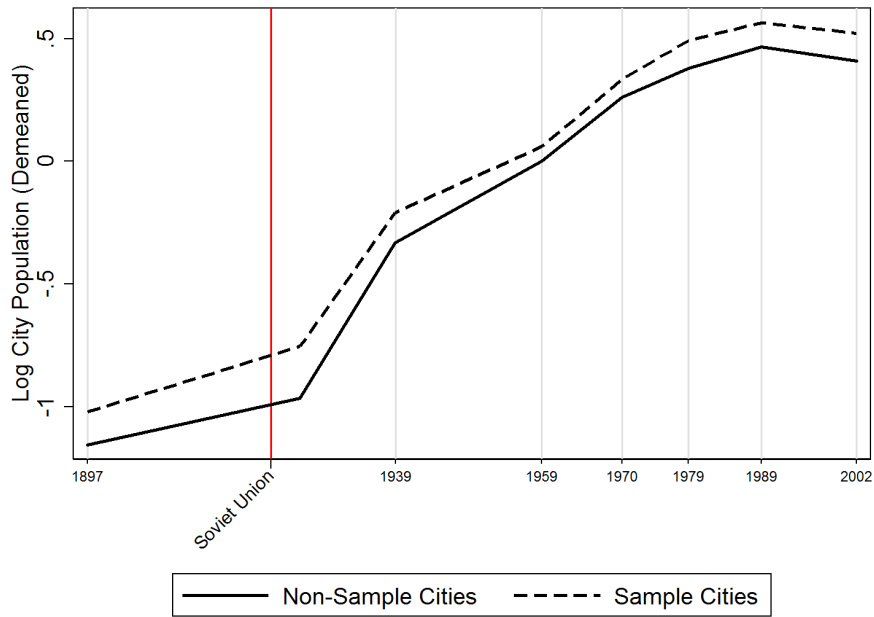


NOTES: This figure plots average demeaned (log) city population for cities included in the balanced sample (dashed line) and those not included in the balanced sample (solid line).

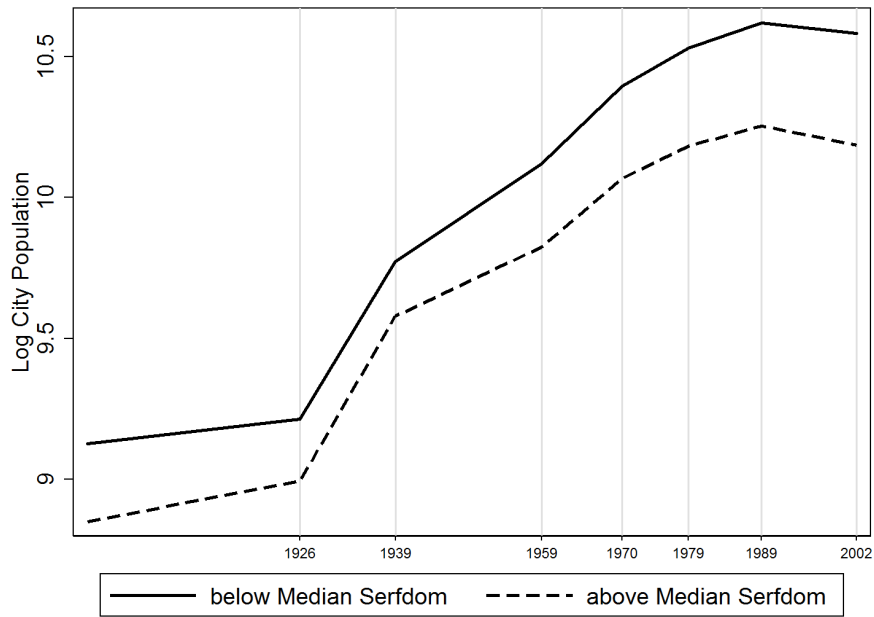FIGURE D8: CITY POPULATION GROWTH IN SAMPLE AND NON-SAMPLE CITIES
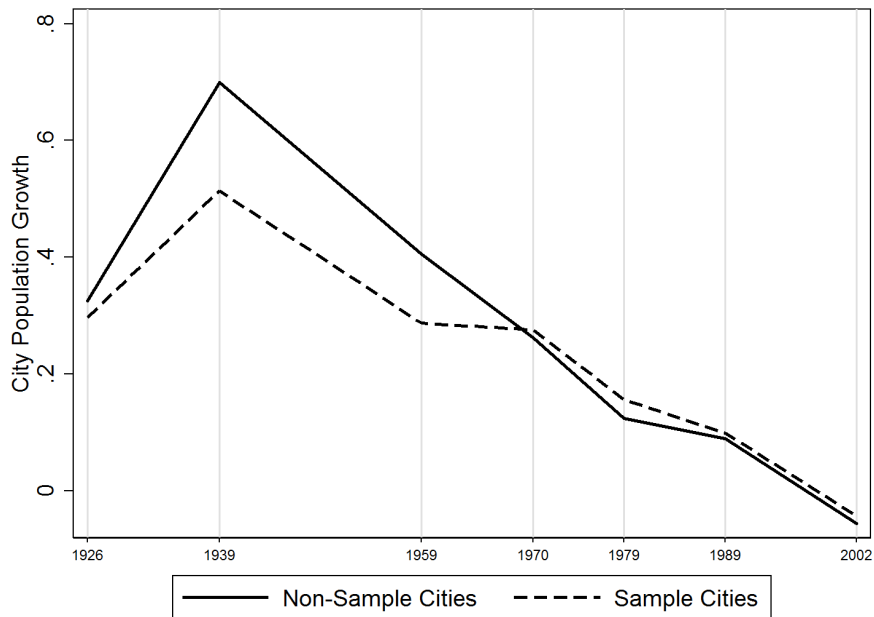


NOTES: This figure plots average city population growth for cities included in the balanced sample (dashed line) and those not included in the balanced sample (solid line).

### TABLE D7: PERSISTENCE THROUGH THE SOVIET PERIOD: PANEL FIXED EFFECTS (1800 - 2002) IN UNBALANCED SAMPLE OF RUSSIAN CITIES

| | Log City Population | | |
| --- | --- | --- | --- |
| | Balanced | Unbalanced | |
| | (1) | (2) | (3) |
| Serfs % (1858) × Year 1850 | -0.378* | -0.378* | -0.172 |
| | (0.201) | (0.199) | (0.254) |
| Serfs % (1858) × Year 1897 | -0.720* | -1.437*** | -0.027 |
| | (0.366) | (0.509) | (0.575) |
| Serfs % (1858) × Year 1926 | -1.378*** | -1.359*** | 0.246 |
| | (0.477) | (0.524) | (0.593) |
| Serfs % (1858) × Year 1939 | -1.517*** | -1.605*** | 0.144 |
| | (0.547) | (0.540) | (0.610) |
| Serfs % (1858) × Year 1959 | -1.995*** | -1.842*** | 0.081 |
| | (0.610) | (0.549) | (0.621) |
| Serfs % (1858) × Year 1970 | -2.076*** | -1.921*** | -0.155 |
| | (0.625) | (0.548) | (0.624) |
| Serfs % (1858) × Year 1979 | -2.178*** | -1.981*** | -0.349 |
| | (0.635) | (0.550) | (0.628) |
| Serfs % (1858) × Year 1989 | -2.155*** | -2.033*** | -0.415 |
| | (0.637) | (0.550) | (0.629) |
| Serfs % (1858) × Year 2002 | -2.209*** | -2.081*** | -0.514 |
| | (0.643) | (0.551) | (0.631) |
| Controls × Year FE | | | ✓ |
| Year FE | ✓ | ✓ | ✓ |
| City FE | ✓ | ✓ | ✓ |
| Observations | 982 | 5031 | 5031 |
| R-squared | 0.76 | 0.67 | 0.71 |
| Number of Clusters | 99 | 683 | 683 |

NOTE: The unit of observation is the city-year. Controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow, the distance to the Province capital, and the distance to the closest city in 1600. Standard errors clustered at the city in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

FIGURE D9: SERFDOM AND CITY POPULATION: UNBALANCED SAMPLE



NOTES: This figure plots average city population growth for cities included in the balanced sample (dashed line) and those not included in the balanced sample (solid line).

### D.5 Province-Level Results

Table D8 reports results from panel estimations using data that is aggregated at the province level.

TABLE D8: PERSISTENCE THROUGH THE SOVIET PERIOD:
PANEL FIXED EFFECTS (1800 - 2002) ACROSS PROVINCES

| | Mean City Size | | Total Urban Population | Mean City Size | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Serfs % (1858) × Post Emancipation | -2.581*** | -1.393 | -1.666 | | |
| | (0.484) | (0.863) | (4.464) | | |
| Serfs % (1858) × Year 1850 | | | | -0.745** | -0.566 |
| | | | | (0.312) | (0.812) |
| Serfs % (1858) × Year 1897 | | | | -1.414*** | -0.635 |
| | | | | (0.516) | (1.062) |
| Serfs % (1858) × Year 1926 | | | | -2.302*** | -0.547 |
| | | | | (0.519) | (1.220) |
| Serfs % (1858) × Year 1939 | | | | -2.866*** | -1.869 |
| | | | | (0.523) | (1.483) |
| Serfs % (1858) × Year 1959 | | | | -3.319*** | -2.034 |
| | | | | (0.626) | (1.500) |
| Serfs % (1858) × Year 1970 | | | | -3.370*** | -2.032 |
| | | | | (0.652) | (1.520) |
| Serfs % (1858) × Year 1979 | | | | -3.439*** | -2.022 |
| | | | | (0.678) | (1.543) |
| Serfs % (1858) × Year 1989 | | | | -3.455*** | -2.128 |
| | | | | (0.687) | (1.562) |
| Serfs % (1858) × Year 2002 | | | | -3.461*** | -2.137 |
| | | | | (0.702) | (1.579) |
| Controls × Post Emancipation | | ✓ | ✓ | | |
| Controls × Year FE | | | | | ✓ |
| Year FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| City FE | ✓ | ✓ | ✓ | ✓ | ✓ |
| Observations | 330 | 330 | 330 | 330 | 330 |
| R-squared | 0.89 | 0.92 | 0.73 | 0.90 | 0.96 |
| Number of Clusters | 33 | 33 | 33 | 33 | 33 |

NOTE: The unit of observation is the province-year. Controls include latitude and longitude of the province, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow, the (average) distance to the Province capital, and the distance to the closest city in 1600. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

48

# E    Additional Results: Military Factories

## E.1    Firm Location and Firm Growth 1939 - 1989

Figure E3 illustrates estimated coefficients and confidence intervals of the effect of serf-dom on the number of military factories, conditional on flexible controls.

Table E1 documents negative effects of serfdom on the number of military firms conditional on linear controls.

Table E2 estimates negative binomial regressions suitable for over-dispersed count variables, such as the number of military firms per district. We similarly find negative effects of serfdom on the number of military firms.

In Table E3, we investigate the effect of past serfdom on the growth rate of military firms in different sub-periods. The persistent negative difference in the number of firms in places with higher incidence of serfdom - documented in the main text - is also reflected by the insignificant relationship between serfdom and industrial growth.

Figure E4 illustrates the associated coefficients and confidence intervals.

FIGURE E1: THE PERSISTENCE OF INDUSTRY LOCATION 1939-1989
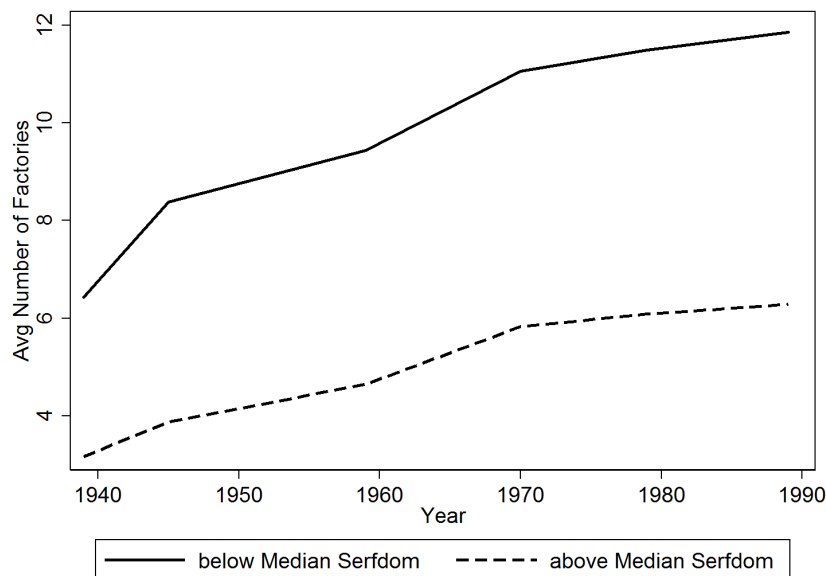


NOTES: This figure plots the average number of firms for cities whose intensity of serfdom is above the median of the distribution (dashed line), and for cities with below median serfdom intensity (solid line).

## FIGURE E2: GAP BETWEEN ABOVE & BELOW MEDIAN SERFDOM AREAS, CONDITIONAL ON PROVINCE × YEAR FE 1939-1989
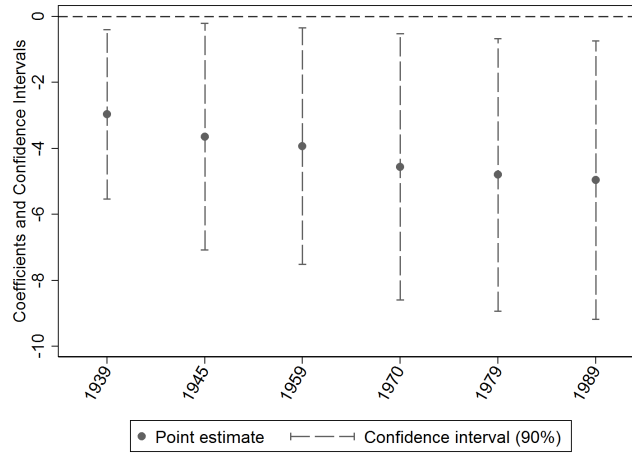


NOTES: This figure plots the coefficients from regressing the number of firms per city-year on a dummy for above median serfdom intensity × year fixed effects, conditional on province dummies × year fixed effects, distance to Moscow × year fixed effects, and average temperature during the growing season × year fixed effects.

## FIGURE E3: ESTIMATED COEFFICIENTS: SERFDOM AND MILITARY FACTORIES
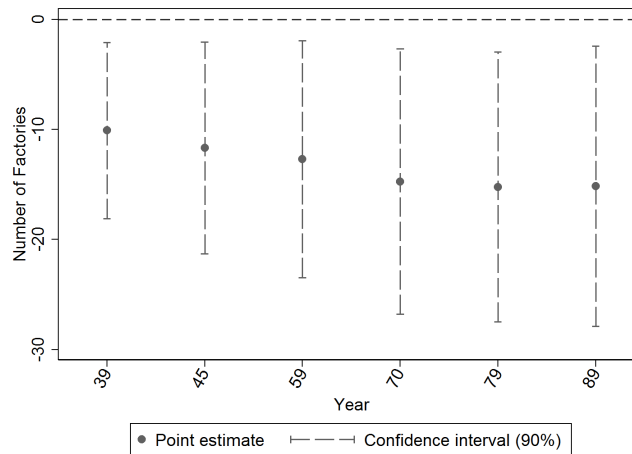


NOTES: This figure plots the coefficient from regressions of the number of firms per city on serfdom, conditional on province fixed effects and flexible controls. The corresponding regression results are reported in Table 10 of the main text.

## TABLE E1: SERFDOM AND INDUSTRY LOCATION IN THE SOVIET PERIOD: LINEAR CONTROLS

| Number of: | Military Firms | | | | | | | Non-Military Firms |
|---|---|---|---|---|---|---|---|---|
| | 1939 | 1945 | 1959 | 1970 | 1979 | 1989 | | 1989 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| Serfs % (1858) | -10.094** | -11.680** | -12.686* | -14.731** | -15.231** | -15.167* | 1.646 | -7.629* |
| | (4.747) | (5.711) | (6.393) | (7.158) | (7.274) | (7.551) | (2.188) | (4.084) |
| Distance Provincial Capital | -2.098 | -2.539 | -2.994 | -3.903 | -4.238 | -4.350 | -0.857 | -1.436*** |
| | (1.771) | (1.919) | (2.228) | (2.565) | (2.586) | (2.646) | (0.814) | (0.437) |
| Distance City in 1600 | -6.016 | -5.497 | -8.854 | -13.155 | -12.425 | -13.975 | -3.955 | -0.595 |
| | (13.744) | (15.782) | (17.505) | (20.161) | (20.521) | (20.982) | (9.123) | (4.481) |
| Number of Factories in 1939 | | | | | | | 1.666*** | |
| | | | | | | | (0.134) | |
| Linear Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 278 | 278 | 278 | 278 | 278 | 278 | 278 | 2656 |
| R-squared | 0.21 | 0.23 | 0.21 | 0.21 | 0.21 | 0.21 | 0.92 | 0.02 |
| Number of Clusters | 41 | 41 | 41 | 41 | 41 | 41 | 41 | 36 |
| $\delta$ for $\beta = 0$ | 2.330 | 1.610 | 1.740 | 1.950 | 1.970 | 1.890 | | 20.290 |
| Lower Bound Estimates | -9.681 | -10.451 | -11.432 | -13.055 | -13.411 | -12.934 | | -7.788 |
| *Conley S.E. 250km* | | | | | | | | |
| Serfs % (1858) | [4.323]** | [5.541]** | [6.224]** | [6.990]** | [7.116]** | [7.422]** | [1.982] | [3.851]** |

NOTE: The unit of observation is a city. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. The restricted model used to compute the Altonji ratios controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## TABLE E2: NEGATIVE BINOMIAL REGRESSIONS

| Number of Factories in: | 1939 | 1945 | 1959 | 1970 | 1979 | 1989 | |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| **Panel A - Linear Controls** | | | | | | | |
| main | | | | | | | |
| Serfs % (1858) | -1.600** | -1.632** | -1.477** | -1.395** | -1.366** | -1.245* | -0.180 |
| | (0.669) | (0.725) | (0.730) | (0.668) | (0.655) | (0.655) | (0.330) |
| Distance Provincial Capital | -0.429** | -0.490*** | -0.480*** | -0.531*** | -0.537*** | -0.525*** | -0.288*** |
| | (0.179) | (0.176) | (0.158) | (0.160) | (0.154) | (0.149) | (0.097) |
| Distance City in 1600 | -1.301 | -0.632 | -0.988 | -1.116 | -1.002 | -1.296 | -0.445 |
| | (2.245) | (2.294) | (1.946) | (1.968) | (1.911) | (1.875) | (1.420) |
| Number of Factories in 1939 | | | | | | | 0.092*** |
| | | | | | | | (0.015) |
| Linear Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province |
| Observations | 278 | 278 | 278 | 278 | 278 | 278 | 278 |
| Number of Clusters | 41 | 41 | 41 | 41 | 41 | 41 | 41 |
| **Panel B - Flexible Controls** | | | | | | | |
| main | | | | | | | |
| Serfs % (1858) | -1.638** | -1.592** | -1.530** | -1.521** | -1.480** | -1.350** | -0.239 |
| | (0.676) | (0.723) | (0.716) | (0.691) | (0.686) | (0.668) | (0.344) |
| Distance Provincial Capital | -0.400** | -0.487** | -0.474** | -0.511*** | -0.516*** | -0.516*** | -0.261** |
| | (0.198) | (0.200) | (0.189) | (0.186) | (0.177) | (0.171) | (0.104) |
| Distance City in 1600 | -0.727 | 0.379 | -0.013 | -0.568 | -0.500 | -0.802 | -0.634 |
| | (2.272) | (2.355) | (2.160) | (2.160) | (2.073) | (2.001) | (1.491) |
| Number of Factories in 1939 | | | | | | | 0.091*** |
| | | | | | | | (0.016) |
| Flexible Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province |
| Observations | 278 | 278 | 278 | 278 | 278 | 278 | 278 |
| Number of Clusters | 41 | 41 | 41 | 41 | 41 | 41 | 41 |

NOTE: The unit of observation is a city. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as linear controls of latitude and longitude of the district, the area covered by forest, ruggedness, the distance to the coast, and the distance to Moscow. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE E3: GROWTH OF FACTORIES

| Growth of Factories in: | 1939-1959 | 1959-1970 | 1970-1989 | 1939-1989 |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Serfs % (1858) | -0.390 | -0.022 | 0.058 | -1.121 |
| | (0.313) | (0.196) | (0.135) | (1.171) |
| Number of Factories in 1939 | 0.002 | | | -0.010 |
| | (0.004) | | | (0.014) |
| Number of Factories in 1959 | | -0.002 | | |
| | | (0.002) | | |
| Number of Factories in 1970 | | | -0.001 | |
| | | | (0.001) | |
| Distance Provincial Capital | -0.136** | -0.137** | 0.005 | -0.307 |
| | (0.064) | (0.055) | (0.025) | (0.330) |
| Distance City in 1600 | -0.084 | -0.089 | -0.242 | -7.872 |
| | (1.019) | (0.942) | (0.405) | (7.024) |
| Flexible Controls | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province |
| Observations | 256 | 265 | 267 | 233 |
| R-squared | 0.29 | 0.31 | 0.25 | 0.40 |
| Number of Clusters | 41 | 40 | 41 | 40 |

NOTE: The unit of observation is a city. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as linear controls of latitude and longitude of the district, the area covered by forest, ruggedness, the distance to the coast, and the distance to Moscow. Standard errors clustered at the province in parentheses. $* \ p < 0.10$, $** \ p < 0.05$, $***$ $p < 0.01$.

FIGURE E4: ESTIMATED COEFFICIENTS: GROWTH OF FACTORIES
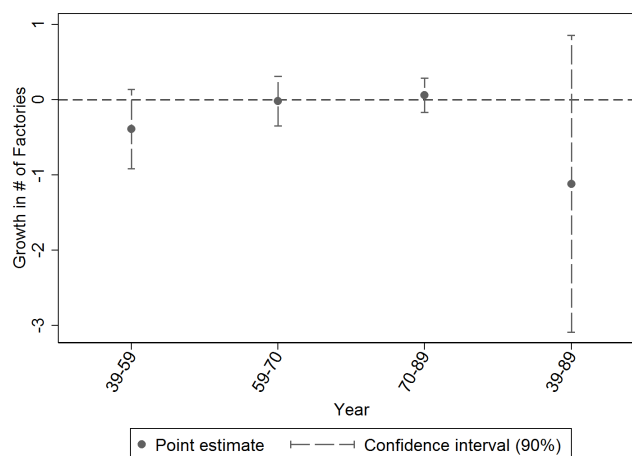


NOTES: This figure plots the coefficient from regressions of the growth rate of the number of firms per city on serfdom, conditional on province fixed effects and flexible controls.

## E.2   Firm Location and Population Growth 1939 - 1989

In Table E5, we study whether localities with higher level of historical serfdom grew faster during the Soviet era if they experienced a larger allocation of military firms early in the period. There is a positive and significant interaction effect between serfdom and the number of factories in 1939, i.e. parts of the negative effect of serfdom on city population growth is mitigated by a larger allocation of firms in 1939.

TABLE E4: SERFDOM, INDUSTRY LOCATION AND URBANIZATION

| | Log Population 1989 | | Population Growth | | | |
| | | | '39-'89 | | '39-'50 | '39-'70 |
| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Serfs % (1858) | -0.573** | -0.678** | -0.054** | -0.064** | -0.025 | -0.041** |
| | (0.232) | (0.273) | (0.024) | (0.027) | (0.015) | (0.018) |
| Serfs % (1858) × Number of Factories in 1939 | 0.030*** | | 0.003*** | | 0.002** | 0.003*** |
| | (0.010) | | (0.001) | | (0.001) | (0.001) |
| Number of Factories in 1939 | -0.002 | | -0.000 | | -0.000 | -0.000 |
| | (0.006) | | (0.001) | | (0.000) | (0.000) |
| Serfs % (1858) × Log Number of Factories 1939 | | 0.242* | | 0.024* | | |
| | | (0.126) | | (0.012) | | |
| Log Number of Factories 1939 | | 0.132 | | 0.014 | | |
| | | (0.088) | | (0.008) | | |
| (Log) Population 1939 | 0.971*** | 0.846*** | -0.013* | -0.024*** | -0.010*** | -0.013** |
| | (0.062) | (0.077) | (0.006) | (0.008) | (0.004) | (0.005) |
| Distance Provincial Capital | -0.019 | -0.037 | -0.002 | -0.004 | -0.000 | -0.001 |
| | (0.057) | (0.060) | (0.006) | (0.006) | (0.004) | (0.005) |
| Distance City in 1600 | 0.273 | 0.545 | 0.028 | 0.061 | 0.014 | 0.015 |
| | (0.829) | (0.890) | (0.084) | (0.087) | (0.044) | (0.059) |
| Linear Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province |
| Observations | 278 | 231 | 278 | 231 | 278 | 278 |
| R-squared | 0.87 | 0.89 | 0.33 | 0.42 | 0.39 | 0.36 |
| Number of Clusters | 41 | 40 | 41 | 40 | 41 | 41 |

NOTE: The unit of observation is a city. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### E.3 Serfdom, WWII, and Industry Location

In Table E5, we investigate whether the intensity of World War II was higher in areas with a higher intensity of serfdom, which could potentially explain the spatial distribution of military firms. We do not find any association between serfdom and (war-related) city population changes between 1939 and 1959, nor with the likelihood that the city was occupied by Germans during WWII. Controlling for these variables, as well as their interaction, we still find a negative and significant effect of serfdom on the number of military firms between 1959 and 1989.

In addition, Figure E5 displays a negative, but weak and insignificant, correlation between the change in the Jewish population in an oblast between 1939 and 1959 and the average intensity of serfdom.

#### TABLE E5: SERFDOM, INDUSTRY LOCATION AND WWII

| | Pop Growth 1939-1959 | Nazi Occupation | # Factories 1959 | # Factories 1970 | # Factories 1979 | # Factories 1989 |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Serfs % (1858) | -0.202 | 0.151 | -13.198* | -15.224** | -15.708** | -15.640** |
| | (0.128) | (0.126) | (6.561) | (7.337) | (7.449) | (7.727) |
| Pop Growth 1939-1959 | | | 10.768** | 12.099** | 12.127** | 12.513** |
| | | | (4.937) | (5.508) | (5.480) | (5.646) |
| German Occupation Dummy | | | 8.074* | 8.832 | 8.902 | 9.154 |
| | | | (4.471) | (5.559) | (5.712) | (5.899) |
| Pop Growth 1939-1959 × German Occupation Dummy | | | -16.100** | -17.465** | -17.044** | -17.366** |
| | | | (6.283) | (7.511) | (7.485) | (7.668) |
| Linear Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province |
| Observations | 322 | 322 | 278 | 278 | 278 | 278 |
| R-squared | 0.37 | 0.89 | 0.23 | 0.23 | 0.23 | 0.23 |
| Number of Clusters | 41 | 41 | 41 | 41 | 41 | 41 |

NOTE: The unit of observation is a city. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

#### FIGURE E5: SERFDOM AND CHANGE IN JEWISH POPULATION 1939-1959 ACROSS OBLASTS



coef = -.01143438, (robust) se = .00816625, t = -1.4
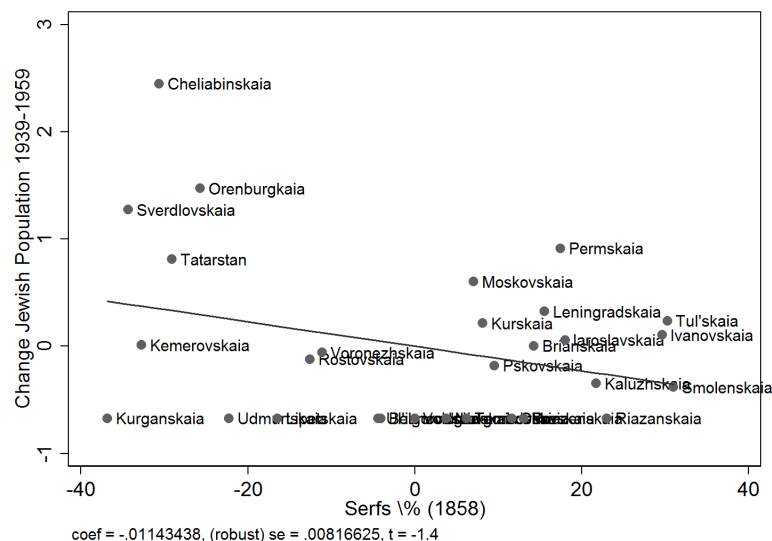
NOTES: This figure plots the relationship between the change in the Jewish population during 1939 and 1959 per oblast and average serfdom.

# F Firms 1989

Figure F1 shows the correlations between the number of military establishments per district in 1989, and the number of non-military firms in the same year aggregated to the district. Table replicates the firm-level analysis of Table conditional on linear controls.

FIGURE F1: MILITARY AND NON-MILITARY FIRMS 1989



coef = .41467583, (robust) se = .05454358, t = 7.6



coef = .44160652, (robust) se = .04523239, t = 9.76
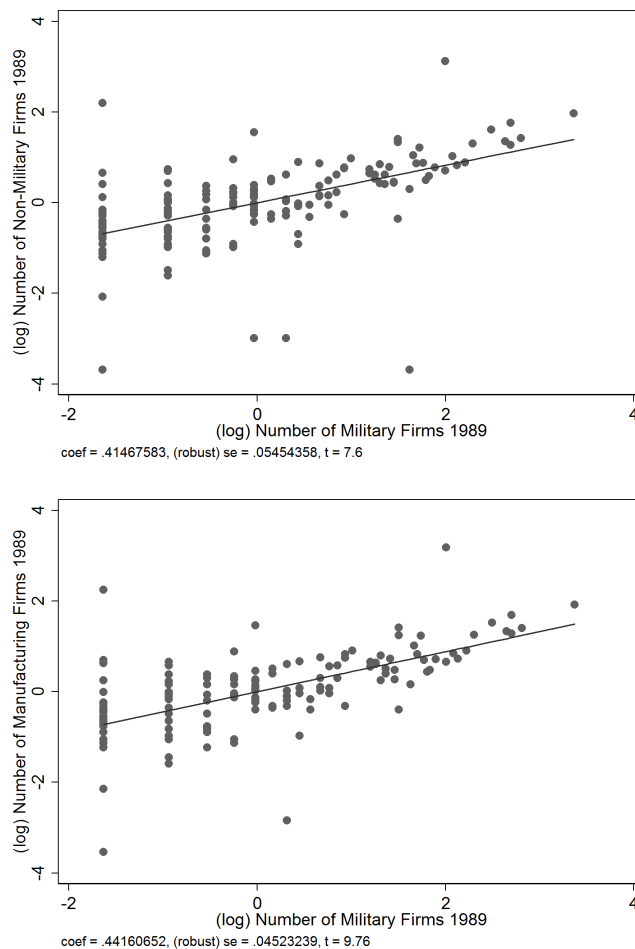
NOTES: These Figure plots the relationship between a) the (log) number of all non-military firms per district in 1989 against the (log) number of military firms per district in the same year, and b) the (log) number of manufacturing firms per district in 1989 against the (log) number of military firms per district in the same year.

56

## TABLE F1: FIRM-LEVEL ANALYSIS 1989: LINEAR CONTROLS

| | Agriculture | Manufacturing | (log) Employment | (log) Turnover | (log) Turnover per Worker |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Serfs % (1858) | 0.032 | -0.044** | -0.245** | -0.325*** | -0.080** |
| | (0.020) | (0.017) | (0.090) | (0.110) | (0.032) |
| **Flexible Controls** | | | | | |
| Distances: City & Prov. Capital | ✓ | ✓ | ✓ | ✓ | ✓ |
| SIC Fixed Effects | | | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province |
| Observations | 14081 | 14081 | 13983 | 13862 | 13852 |
| R-squared | 0.03 | 0.02 | 0.56 | 0.63 | 0.77 |
| Number of Clusters | 36 | 36 | 36 | 36 | 36 |
| $\delta$ for $\beta = 0$ | 1.640 | 1.150 | 2.200 | 5.280 | -1.920 |
| Lower Bound Estimates | 0.025 | -0.030 | -0.180 | -0.304 | -0.119 |
| *Conley S.E. 300km* | | | | | |
| Serfs % (1858) | [0.016]** | [0.019]** | [0.093]*** | [0.113]*** | [0.037]** |

NOTE: The unit of observation is a firm. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Distances are the distance to the nearest city in 1600, and the distance to the Provincial capital. SIC Fixed Effects are dummies for industrial classifications of firms using the 5-Digit Standard Industrial Classification Codes. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# G   Additional Results: District-Level Outcomes and Mechanisms

## G.1   Structural Change and Urbanization

Table G1 documents negative effects of serfdom on indicators of structural change and urbanization conditional on linear controls.

TABLE G1: STRUCTURAL CHANGE AND URBANIZATION: LINEAR CONTROLS

| | Pre-Soviet | | | | Soviet | | Post-Soviet | | |
| | Urbanization Rate | | Factories per 1,000 ppl, 1868 | Log Production per Worker, 1868 | Road Density | Gulag | Log Pop. Density 2000 | Log Light Density | |
| | 1863 | 1913 | | | | | | 1994 | 2008 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
|---|---|---|---|---|---|---|---|---|---|
| Serfs % (1858) | -15.634*** | -15.808*** | -47.209 | -0.583* | -0.008*** | -0.293** | -0.962*** | -0.588** | -0.820*** |
| | (5.014) | (5.463) | (32.863) | (0.296) | (0.003) | (0.127) | (0.320) | (0.240) | (0.281) |
| Distance City in 1600 | -12.996* | -23.047** | -1.765 | 0.361 | -0.012** | -0.333 | -1.603** | -1.063* | -1.470* |
| | (7.687) | (9.066) | (47.779) | (0.964) | (0.005) | (0.329) | (0.746) | (0.561) | (0.741) |
| Distance Provincial Capital | -4.903*** | -5.640*** | -11.308*** | -0.252*** | -0.001*** | -0.021 | -0.360*** | -0.307*** | -0.313*** |
| | (0.649) | (0.662) | (3.799) | (0.069) | (0.000) | (0.031) | (0.061) | (0.048) | (0.054) |
| Linear Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 483 | 490 | 483 | 434 | 490 | 490 | 490 | 490 | 490 |
| R-squared | 0.38 | 0.37 | 0.32 | 0.29 | 0.51 | 0.37 | 0.61 | 0.63 | 0.54 |
| Number of Clusters | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| $\delta$ for $\beta = 0$ | 6.950 | 5.580 | 1.930 | 2.500 | 124.240 | 1.810 | 14.520 | -5.220 | 77.930 |
| Lower Bound Estimates | -14.955 | -14.368 | -36.777 | -0.465 | -0.010 | -0.203 | -1.014 | -0.762 | -0.905 |
| *Conley S.E. 300km* | | | | | | | | | |
| Serfs % (1858) | [4.922]*** | [5.049]*** | [31.971] | [0.218]*** | [0.003]*** | [0.144]** | [0.302]*** | [0.248]** | [0.234]*** |

NOTE: The unit of observation is a district. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## G.2 Employment by Sectors 1897

Table G2 documents effects of serfdom on employment shares in primary and secondary sectors, and indicators of structural change, as well as hetereogeneity by type of serfdom, conditional on linear controls.

Besides primary and secondary employment, Table G3 tests for effects of serfdom on employment in other sectors. The results suggest that employment in service, education, and commercial opportunities was significantly reduced in areas were serfdom was more widespread.

TABLE G2: EMPLOYMENT AND HETEROGENEITY IN STRUCTURAL CHANGE BY TYPE OF SERFDOM: LINEAR CONTROLS

| | Primary Empl. 1897 | | Secondary Empl. 1897 | | Industry Empl. 1897 | | Factories Per 1,000 ppl, 1868 | | Log Light Density, 2008 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Serfs % (1858) | 0.064 | | 0.003 | | -0.012 | | -47.209 | | -0.820*** | |
| | (0.042) | | (0.023) | | (0.022) | | (32.863) | | (0.281) | |
| Obrok % (1858) | | -0.019 | | 0.029*** | | 0.021** | | -5.386 | | 0.031 |
| | | (0.011) | | (0.009) | | (0.009) | | (7.783) | | (0.080) |
| Barschina % (1858) | | 0.033** | | -0.014 | | -0.014* | | -10.260 | | -0.273*** |
| | | (0.014) | | (0.009) | | (0.007) | | (6.177) | | (0.082) |
| Household Serfs % (1858) | | 0.005 | | -0.007 | | -0.008* | | -6.740 | | -0.113 |
| | | (0.010) | | (0.005) | | (0.004) | | (4.952) | | (0.087) |
| H0: Barschina = Obrok (*p-value*) | | 0.00 | | 0.00 | | 0.01 | | 0.48 | | 0.00 |
| Linear Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Distances: City & Prov. Capital | | ✓ | | ✓ | | ✓ | | ✓ | | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 490 | 468 | 490 | 468 | 490 | 468 | 483 | 466 | 490 | 468 |
| R-squared | 0.48 | 0.51 | 0.56 | 0.61 | 0.53 | 0.57 | 0.32 | 0.33 | 0.54 | 0.57 |
| Number of Clusters | 50 | 49 | 50 | 49 | 50 | 49 | 50 | 49 | 50 | 49 |

NOTE: The unit of observation is a district. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

TABLE G3: DIFFERENCES IN OTHER SECTORS

| | Service | | Education & Research | | Commerce | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Serfs % (1858) | -0.006*** | -0.005*** | -0.003** | -0.003** | -0.017*** | -0.018*** |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.005) | (0.005) |
| Distance City in 1600 | -0.006* | -0.006* | -0.002 | -0.001 | -0.004 | 0.000 |
| | (0.004) | (0.003) | (0.002) | (0.002) | (0.014) | (0.014) |
| Distance Provincial Capital | -0.003*** | -0.003*** | -0.001*** | -0.001*** | -0.006*** | -0.006*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.001) | (0.001) |
| Linear Controls | ✓ | | ✓ | | ✓ | |
| Flexible Controls | | ✓ | | ✓ | | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province |
| Observations | 490 | 490 | 490 | 490 | 490 | 490 |
| R-squared | 0.40 | 0.44 | 0.49 | 0.51 | 0.58 | 0.59 |
| Number of Clusters | 50 | 50 | 50 | 50 | 50 | 50 |
| $\delta$ for $\beta = 0$ | 7.660 | 5.550 | 5.950 | 4.180 | 2.390 | 2.550 |
| Lower Bound Estimates | -0.005 | -0.005 | -0.002 | -0.002 | -0.014 | -0.016 |
| *Conley S.E. 300km* | | | | | | |
| Serfs % (1858) | [0.001]*** | [0.001]*** | [0.001]*** | [0.001]*** | [0.005]*** | [0.005]*** |

NOTE: The unit of observation is a district. Linear controls include latitude and longitude of the district, the area covered by forest, ruggedness, cereal suitability, growing-season temperature and precipitation, river density, share of podzol soil, the distance to the coast, and the distance to Moscow. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as linear controls of latitude and longitude of the district, the area covered by forest, ruggedness, the distance to the coast, and the distance to Moscow. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

### G.3 Light Density in Different Periods

In Table G4, we document that the negative effect of past serfdom on log light density in the contemporary era is not limited to the years 1994 and 2008 that we use in the main text. We find negative effects for all years from 1990 to 2012, and present here a selection of years (1994, 1999, 2005, 2008, 2012). In addition, the odd columns of Table G4 report that the results are robust even to conditioning on past population density measured in 1858.

TABLE G4: SERFDOM AND LIGHT DENSITY IN DIFFERENT PERIODS

| Log Light Density in | 1994 | | 1999 | | 2005 | | 2008 | | 2012 | |
|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| Serfs % (1858) | -0.580** | -0.416* | -0.664** | -0.470** | -0.713** | -0.507* | -0.820*** | -0.614** | -0.667** | -0.476* |
| | (0.247) | (0.232) | (0.260) | (0.234) | (0.319) | (0.291) | (0.295) | (0.268) | (0.283) | (0.266) |
| Distance City in 1600 | -0.721 | -0.771 | -0.866 | -0.943 | -1.160 | -1.202 | -1.221* | -1.234* | -1.028 | -1.048 |
| | (0.539) | (0.533) | (0.611) | (0.640) | (0.706) | (0.729) | (0.698) | (0.724) | (0.714) | (0.741) |
| Distance Provincial Capital | -0.308*** | -0.234*** | -0.349*** | -0.278*** | -0.328*** | -0.250*** | -0.313*** | -0.233*** | -0.310*** | -0.233*** |
| | (0.051) | (0.049) | (0.058) | (0.053) | (0.062) | (0.055) | (0.061) | (0.055) | (0.063) | (0.057) |
| (ln) Pop Density 1858 | | 0.627*** | | 0.662*** | | 0.704*** | | 0.702*** | | 0.650*** |
| | | (0.080) | | (0.094) | | (0.105) | | (0.103) | | (0.103) |
| Flexible Controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Fixed Effects | Province | Province | Province | Province | Province | Province | Province | Province | Province | Province |
| Observations | 490 | 478 | 490 | 478 | 490 | 478 | 490 | 478 | 490 | 478 |
| R-squared | 0.65 | 0.70 | 0.64 | 0.69 | 0.58 | 0.64 | 0.57 | 0.63 | 0.60 | 0.65 |
| Number of Clusters | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 | 50 |
| $\delta$ for $\beta = 0$ | -10.730 | 2.410 | 25.260 | 1.650 | 24.170 | 2.040 | 17.250 | 2.330 | 14.860 | 1.940 |
| Lower Bound Estimates | -0.719 | -0.327 | -0.754 | -0.296 | -0.783 | -0.356 | -0.888 | -0.462 | -0.713 | -0.324 |
| *Conley S.E. 300km* | | | | | | | | | | |
| Serfs % (1858) | [0.283]** | [0.247]* | [0.275]** | [0.222]** | [0.325]** | [0.270]* | [0.303]*** | [0.248]** | [0.285]** | [0.239]** |

NOTE: The unit of observation is a district. Flexible controls include eight dummies for cereal suitability, and four dummies for quartiles of growing season temperature, growing-season precipitation, the share of podzol soil, and river density, as well as linear controls of latitude and longitude of the district, the area covered by forest, ruggedness, the distance to the coast, and the distance to Moscow. The restricted model used to compute $\delta$ and the lower bound estimates controls for province fixed effects. Standard errors clustered at the province in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# H   Additional Results: Country-Level

FIGURE F1: PEASANT EMANCIPATION AND SHARE OF LABOR IN AGRICULTURE
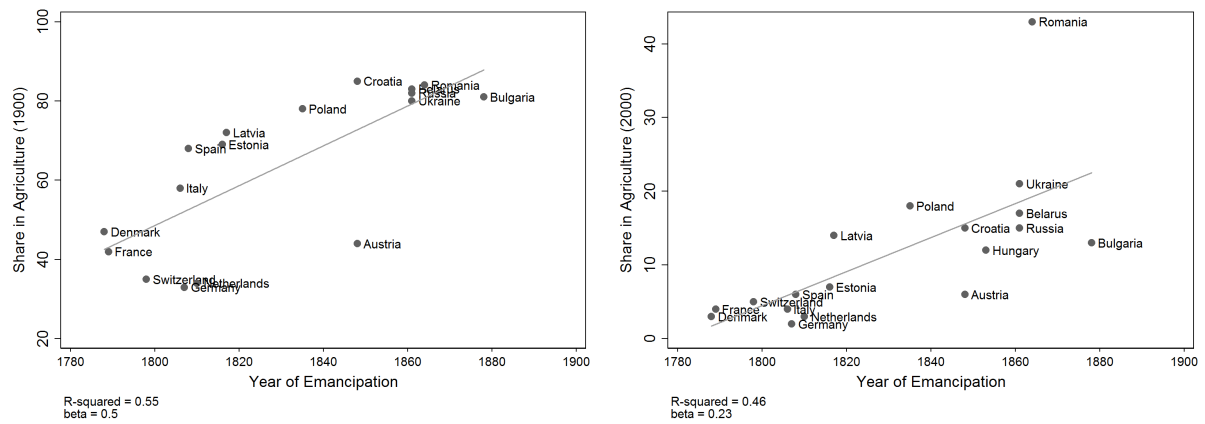


NOTES: This figure plots the share of labor in agriculture in 1900 and 2000 against the year of peasant emancipation in European countries. See Supplementary Appendix for data description.

# References

**Acemoglu, Daron, Tarek A Hassan, and James A Robinson**, "Social Structure and Development: A Legacy of the Holocaust in Russia," *Quarterly Journal of Economics*, 2011, *126* (2).

**Algan, Yann and Pierre Cahuc**, "Trust, Growth and Well-being: New Evidence and Policy Implications," in "Handbook of Economic Growth, Vol. 2B," North Holland, Elsevier, 2013.

**Atkin, Nicholas, Michael Biddiss, and Frank Tallett**, *The Wiley-Blackwell Dictionary of Modern European History Since 1789*, Wiley-Blackwell, 2011.

**Bairoch, Paul, Jean Batou, and Pierre Chèvre**, *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats (la)*, Librairie Droz, 1988.

**Blum, Jerome**, *The End of the Old Order in Rural Europe*, Princeton University Press, 1978.

**Bushen, A., ed.**, *Nalichnoe naselenie Imperii za 1858 god*, Vol. 2 of *Statisticheskiia tablitsy Rossiiskoi Imperii* 1863.

**Davies, Norman**, *God's Playground: A History of Poland, Vol. 2: 1795 to the Present (Volume 2)*, Columbia University Press, 2005.

**Dower, Paul Castaneda, 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), 125–147.

**Fal'bork, G. and V. Charnoluskii, eds**, *Nachal'noe narodnoe obrazovanie v Rossii.*, Imperatorskoe volÊźnoe ekonomicheskoe obshchestvo, 1900-1905.

**Jarrett, Mark**, *The Congress of Vienna and its legacy: War and great power diplomacy after Napoleon*, Vol. 80, IB Tauris, 2013.

**Magas, Branka**, *Croatia Through History*, Saqi Books, 2008.

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

**Mikhailova, Tatiana**, "Gulag, WWII and the Long-Run Patterns of Soviet City Growth," 2012. mimeo.

**Nafziger, Steven**, "The Political Economy of Primary Education in Tsarist Russia," 2012. mimeo.

__ , "Russian Serfdom, Emancipation, and Land Inequality: New Evidence," 2013. mimeo.

**Nunn, Nathan**, "Culture and the historical process," *Economic History of Developing Regions*, 2012, *27* (sup1), S108–S126.

__ **and Leonard Wantchekon**, "The Slave Trade and the Origins of Mistrust in Africa," *American Economic Review*, 2011, *101* (7), 3221–52.

**O'Neill, Kelly**, "Rivers (named) of the Russian Empire," in "The Imperiia Project: Economic and Cultural Infrastructures of the Russian Empire Dataverse," History Department, Harvard University, 2016.

**Satyanath, Shanker, Nico Voigtländer, and Hans-Joachim Voth**, "Bowling for Fascism: Social Capital and the Rise of the Nazi Party," *Journal of Political Economy*, 2017, *125* (2), 478–526.

**Troinitskii, N. A.**, *The Serf Population in Russia According to the 10th National Census*, Oriental Research Partners, 1982 (1861).

— **, ed.**, *Pervaia vseobshchaia perepis' naseleniia Rossiiskoi Imperii, 1897 g.*, Vol. Multiple, Central Statistical Committee, Ministry of the Interior, 1905.

**Tsentraf'nyi statisticheckikh komitet, Ministerstvo vnutrennykh del**, *Statisticheskii Vremennik Rossiiskoi Imperii*, Vol. 6 of *II* 1872.

— , *Sel'skiia uchilishcha v Evropeiskoi Rossii i Privislianskikh guberniiakh*, Vol. 1 of *Statisticheskii vremennik Rossiiskoi imperii (Series III)* 1884.

— , *Statistika zemliavladeniia 1905 g.*, Vol. Multiple 1906.

— , *Statisticheskii ezhegodnik Rossii 1913 g.* 1914.

**Wingender, Asger**, "Structural transformation in the 20th Century: A new database on agricultural employment around the world," 2014.

**Zagorov, Slavcho**, *The agricultural economy of the Danubian countries, 1935-45*, Vol. 8, Stanford University Press, 1955.