A Regional Perspective on the Economic Development of the late Russian Empire

Andrei Markevich[1]
(New Economic School, Moscow, Russia)

This version: May 14th, 2019

Abstract
This paper analyzes the relative impacts of geographical and institutional factors on the economic development of the late Russian empire. I reconstruct gross regional products and labor productivity for all provinces of the empire around 1900 for the first time. My estimates highlight substantial heterogeneity within a middle-income country. I show that both natural advantages – sea access, mineral resources, land abundance – and institutions, in particular the legacy of serfdom, account for the observed variation. I also provide evidence that market potential and specialization externalities played a role.

[1] Address for correspondence: amarkevich@nes.ru; New Economic School, 45 Skolkovskoe Shosse, Skolkovo, Moscow, 143025, Russia.
I thank Paul Castañeda Dower, Jari Eloranto, Kerstin Enflo, Amanda Gregg, Antii Hakkinen, Paul Gregory, Mark Harrison, Alex Klein, Gijs Kessler, Sakari Saaritsa, Steven Nafziger, Albrecht Ritschl, Joan Roses, Max Schulze, Timur Valetov and Ekaterina Zhuravskaya, anonymous referees, and the seminar participants at LSE and 2018 World Economic History Congress in Boston, for insightful comments. I thank Olga Pavlenko, Pavel Bacherikov and Viktor Malein for their excellent research assistance. All possible errors are mine.

1. Introduction

In this paper, I explore the fundamental factors affecting economic development in the late 19th century Russian empire. To do this, I reconstruct nominal and real gross regional products (GRPs) as well as labor productivity for all provinces of the empire and analyze the determinants of their regional variation. My estimates confirm that the average citizen in imperial Russia in the late 19th – early 20th century had a middle-sized income, according to the international standards of that time. But the reconstruction also reveals substantial spatial inequality within the country, comparable to the spatial inequality within such countries of the new world as the US or Mexico. The richest Russian regions – Saint-Petersburg and Moscow – were at the level of the most developed economies at the turn of the century while GDP per capita hardly exceeded the subsistence level in the poorest regions of the empire, like Turgaj or Semirechie,

Given the size of the Russian empire, a territory of which equals roughly one sixth of the global landmass and is currently home to seventeen separate states, geography is a natural candidate to account for this variation. Alternatively, the economic history literature has long explained the backwardness of economic development of this part of the world relative to the Western Europe by pointing to institutional factors (e.g. Gerschenkron 1962, 1965). I show that both factors of physical geography and the institutional characteristics of Russian regions affected their economic development. In particular, I find that sea access, mineral resources and the legacy of serfdom account for almost half of the variation in labor productivity across Russia in the late 19th century. Surprisingly, almost the same factors (if one replaces serfdom by slavery) are major determinants of labor productivity in the US around 1900 (Mitchener and McLean 2003). I also provide evidence that in addition to mineral resources, land abundance was positively correlated with income, which is an important finding in terms of the hypothesis that 19th century Russia was a frontier economy (Markevich and Mikhailova 2013) and the debate on the resource abundance and economic devolvement (e.g. Papyrakis 2017). More broadly, with these findings, I contribute to the debate on the relative importance of geographical and institutional factors for economic growth (e.g. Gallup, Sachs and Mellinger 1998; Acemoglu and Robinson 2012 among many others).

My findings on the negative legacy of serfdom complement recent empirical papers on the consequences of this institution of forced labor (Buggle and Nafziger 2016, forthcoming, Markevich and Zhuravskaya 2018). I extend the geographical coverage of such work, exploring the effect of serfdom on the development of all regions of the Russian empire rather than only the European part of the country. To address potential endogeneity and mismeasurement concerns, I adjust the approaches suggested by Buggle and Nafziger (2016) and Markevich and Zhuravskaya (2018). I rely on exogenous variation in the prevalence of serfdom driven by historical geography of orthodox monasteries and the nationalization of church lands with serfs on them a century before the emancipation. I provide suggestive evidence regarding the channels through which the legacy of serfdom continued to affect economic development almost forty years after its abolition.

This paper is also related to the new economic geography literature. The models from this literature assume increasing returns to scale. They predict that demand factors together with positive externalities from economies of scale and regional specialization, especially in modern sectors, are important determinants of spatial variation in economic development (e.g. Fujita et al. 1999). The Russian development pattern at the early stages of industrialization in the late 19th – early 20th century is in line with these projections. I find that market potential along with economic specialization in particular subsectors of industry or agriculture (but not services) were positively associated with labor productivity. I also provide evidence on the importance of economies of scale in industry in the Russian context.

Last but not least, by presenting a comprehensive snapshot of the economic development across the country this paper provides a new perspective on the history of the late 19th-century Russian Empire. The literature on the development of this part of the world in that time primarily relies on the analysis of national development trends (e.g. Gerschenkron 1962, 1965; Gregory 1982, 1994; Allen 2003; Mironov 2010; Cheremukhin et al. 2018). As Paul Gregory (2001) notes, regional variation in economic development across the large territory of the former Russian empire remains largely unknown and unexplored. The noticeable exception is Lindert and Nafziger (2014), who explore inequality in the Russian empire on the eve of the 1905 revolution and estimate

the distribution of household incomes in 1904 for fifty provinces of the European part of the empire. But, unlike my current endeavor, their calculations do not cover such parts of the Russian empire as Poland (10 provinces), Finland (8 provinces), the Caucasus (11 provinces), Siberia (9 provinces), and Central Asia (9 provinces). The other difference is that I reconstruct gross regional products from the production side by 44 subsectors, which allows me to analyze the spatial variation in the sectoral structure of the economy. Not surprisingly, I find that industrialization was positively correlated to economic development. At the same time, my reconstruction shows that relatively advanced agriculture was an alternative way to prosperity in a number of regions of the empire. The level of development of agriculture in these provinces does not support claims of an agricultural crisis in pre-revolutionary Russia (Gerschenkron 1965; Allen 2003; Grinin et al. 2010).

The paper proceeds as following. The next section outlines the methodology of my reconstruction of regional economic indicators and presents the baseline estimates. In section 3, I review the geographical patterns of economic development which follow from my reconstruction. Section 4 adds an international perspective to my analysis and compares the spatial distribution of national income in Russia and other countries around 1900. Section 5 explores the determinants of the observed variation in labor productivity across the country. The last section summarizes my findings.

2. Reconstructing regional indicators of economic development

First, I estimate the nominal gross regional product for 97 provinces of the Russian empire in 1897. I chose this year because of the timing of the first and only imperial population census. Using various official sources, I estimate value added in 44 subsectors, namely in 14 subsectors in industry, 18 subsectors in agriculture (including primary economic activities such as forestry, fishing and hunting), and in 12 subsectors in services.[2] Second, I construct a relative regional price index of consumption goods and

[2] To be precise, for industry, I reconstruct value added in ten branches of large industry (cotton, wool, linen products; food; leather; woodworking; paper; chemicals; ceramic; mining and quarrying; machine-building and metal-working; and other industrial products) as well as the value added of products manufactured by large industry from materials received from customers, and value added in small industry, construction and public infrastructure. For agriculture, I estimate value added of production of grains, potatoes and beetroot, industrial crops, beans and lentils, as well as value added in meat production, milk production, production of leather, horse breeding, pig farming, sheep and goats farming, production of hay and straw, production of sunflower, mustard, hop and rapeseed, gardening and beekeeping, production of fruits and grapes,

services (equaled to one for a hypothetical average province) in order to be able to switch from nominal to real national income. For this index, I use local prices in 1904 (Central Statistical Agency 1906), which is the closest year with the necessary information on local prices with comprehensive geographical coverage. Third, I estimate labor productivity (measured as output per working age citizen).[3] Sections AI-AIII in the online appendix describe the methodology in full details and report all related tables. Of these, tables A1-A61 in the online appendix present the data as retrieved from the sources, and tables VA1-VA8 report my estimates of value added for individual subsectors. Tables A62 and A63 introduce estimates of the relative price index and working-age population, respectively. Real GRP and labor productivity are in table VA9. Table VA0 in the online appendix reports all estimates for individual provinces in one table.

I start with summary statistics for my baseline estimates of major reconstructed indicators (Table 1). An average province with a population of 1.3 million people produced about 100 million rubles of gross regional product in current prices, i.e. about 82.5 rubles per capita, or about 1,325 1990 international dollars.[4] If one accounts for the differences in local prices of consumption goods, national income per capita in an average province drops to about 79 rubles. Labor productivity in such a province was about 150 rubles. An average province had an economy with a large agricultural sector and smaller but more productive, predominantly urban, industry and service sectors. The

---

fishing, hunting, and forestry. For the service sector, I estimate value added by services produced by railroads, water transport, trucking, value added in communications, trade, banking, production of public goods (public education, health, national defense, public administration), urban and rural housing, and the value of services added by servants.

[3] The 1897 population census suffers from the under-enumeration of employment. In family firms, the census counted only the heads of households as gainfully employed. All other family members were registered as dependents irrespective of whether they actually contributed to the family businesses or not (Gozulov 1936 pp. 205-206). Because of peasant family farms, the problem was the worst in agriculture and less pronounced in other sectors. The under-enumeration of employment in the census makes it difficult to estimate labor productivity by sector properly. To estimate average labor productivity across sectors in a region, I use the population in working age as the denominator. According to various household surveys in the late Russian empire, the working age population (males between the ages of 16-59 and females between the ages of 16-54) fitted almost one-to-one to the number of workers (Svavitskij and Svavitskij 1926). Section AIII in the online appendix provides more details on labor productivity estimates.

[4] To bring an international perspective, I introduce the 1897 ruble / international 1990 US dollar exchange rate (16.0431908) obtained from Maddison's (2010) and Gregory's (1982) estimates of national income per capita in 1900 and 1913 in 1990 USD dollars and 1900 and 1913 Russian rubles, respectively, and Gregory's (1982) deflation index. The estimates of regional incomes in 1990 US dollars are a crude approximation; I use them for comparisons across national borders only.

bulk of the population – 87 percent in the average province – lived in the countryside, and the majority of them were employed in agriculture, even if agriculture produced only about 57 percent of nominal national income in the average province. Industry provided 19 percent of nominal GRP, and services accounted for the remaining quarter of nominal output in an average province.[5]

Table 1 somewhere here

I reconstruct regional incomes from the production side by sector using disaggregated data. An alternative would be to use an indirect short-cut method originally suggested by Geary and Stark (2002) and recently utilized to reconstruct GRPs of the European regions in the 20th century (Roses and Wolf 2018). The short-cut method assumes that regional sectoral wages reflect local productivity in the corresponding sectors, and thus along with sectoral employment, could be used to allocate country level-output. Under these assumptions, the advantage of this method is that it requires few data inputs, namely only those pertaining to historical national accounts, and wages and employment by sector and region.[6] Unfortunately, some of the necessary ingredients, like regional sectoral wages, are unavailable for Russia, as there is no systematic information on wages for the outskirts of the empire. In addition, given the size of the country and a potentially limited level of market integration, the assumption that wages reflect productivity might be problematic for some sectors, like agriculture, which is characterized by an abundance of self-sufficient farms with a low level of market operations. Because of all that, regional income estimates based exclusively on the short-cut method cannot be obtained for the provinces of the Russian empire. Instead, I mostly rely on regional disaggregated data to estimate value added in individual subsectors in each province of the empire. For example, for crops production, I multiply province-level figures on output by local crop-specific market prices to get monetary values in rubles (see a detailed description of the methodology in section AI of the online appendix).[7]

[5] Due to data availability, I cannot construct sector-specific price indexes to estimate sectoral shares of real GRPs.

[6] Conditional on data availability, there are other indirect short-cut methods based on using various proxies that are functionally related to output to predict GRP levels. E.g. Caruana-Galizia and Ma (2016) use cultivated acreage, urban share and population to proxy regional gross products in Chinese provinces in the late 19th – early 20th century.

[7] To be precise, subsectors with value added reconstructed from disaggregated data underlie almost 90 percent of my estimate of national income for the Russian empire. For the rest, I use a "top-down"

Reconstruction from the production side requires a lot of input data. They are available because of a rich tradition of high-quality statistical accounting in late imperial Russia. I follow the approach originally elaborated by Sergei Prokopovich (1918), who pioneered estimates of output of the Russian economy in the beginning of the 20th century, with later adjustments by various authors as discussed in Markevich and Harrison (2011). Where possible, I rely on subsector-specific corrections as discussed in the literature (Yasnopolskij 1891, 1897; Silantiev 1898; Commission on peasant welfare 1901; Groman 1912; Strumilin 1924; Central statistical agency 1924; Vainshtein 1960; Gregory 1982; Salomatina 2014; section AI of the online appendix describes how I use these corrections and my reconstruction methodology in full details). In particular, I adjust for double counting where possible, like in the case of large industry.[8] Similarly, I apply corrections to the original figures as discussed in the related literature (e.g. I adjust cereal output in physical units and livestock figures for under-registration).[9] Because of all that, my baseline estimate might have some potential upward bias. If I relax the assumptions, which push my estimates up, the obtained alternative figures on regional national incomes are not very much different to the baseline ones. Table B1 in the online appendix illustrates that with summary statistics of alternative estimates (I report alternative estimates for each single province in table A65 in the online appendix and discuss their robustness in more details in section AIV in the online appendix).

Two external comparisons validate my reconstruction. First, the sum of reconstructed GRPs over all provinces except Finland (i.e. the national income of the

---

approach that is similar Geary and Stark's (2002) short-cut method. I either estimate or borrow from the existing literature the value added in the related subsector at the national level, and allocate this value between provinces using regional activity in the corresponding subsector (measured in various ways) as weights. I account for regional variation in monetary indicators used in my estimates (prices, wages, rents) everywhere data allows this. All in all, subsectors, for which I do this, underlie more than 90 percent of my estimate of national income for the Russian empire. See table A0 in the online appendix for an overview of my chosen reconstruction methodologies by subsector. The online appendix provides full description of the methodology.

[8] For large industry, I sum up disaggregated figures on industrial output by subsectors and adjust the obtained sums using industrial subsector specific ratios of value added to gross industrial production in 1913 from the 1918 industrial census (Central statistical agency 1924). See section AI of the online appendix for further details.

[9] The literature debates how large the adjustment should be for cereals. I follow Markevich and Harrison (2011) and apply a correction that makes my GRPs estimates comparable with the previous estimates of the Russian GDP. In section AI of the on-line appendix I discuss potential alternative corrections and show that they do not change the results substantially. The livestock adjustment relies on the approach originally suggested by Vainshtein (1960). See section AI of the online appendix for further details.

Russian empire in 1897 without the Great Duchy of Finland), equals 9,448 million rubles that is close to 9,172 million rubles – the estimate of national income in this year by Gregory (1982), the most solid reconstruction of national income for the former Russian empire.[10] The difference between these two estimates is about three percent, while the reconstruction procedures, data and original sources are independent: in contrast to the present study, Gregory (1982) estimated national income from the income side.

The second external comparison validates my reconstruction in terms of relative ranking. I compare my estimates for the 50 provinces of the European part of the empire with calculations of mean household incomes in these provinces in 1904 made by Lindert and Nafziger (2014).[11] While the methodologies of the two reconstructions, data and sources are independent, the two sets of estimates fit each other well. The coefficient of correlation is high (0.84) and statistically significant at the one percent level.[12] The difference in estimates for individual provinces might be explained by the difference between production-based and income-based measures, i.e. the difference between where output was produced and where it was consumed.[13]

3. Economic geography of the Russian Empire

To the best of my knowledge, my study presents the first systematic reconstruction of the major indicators of economic development with complete geographical coverage at the regional level for the territory of the former Russian empire. For that reason I describe the geographical patterns of economic development of the late Russian empire in some detail, before proceeding to the quantitative analysis of the observed spatial variation. I start with a discussion of the spatial distribution of nominal national income per capita as rendered in figure 1.

[10] Gregory (1982) is one of the sources for Maddison (1995). For Russia, Maddison reconstructs national income in the borders of the former Soviet Union and for selected years only. He does not do that for 1897, the reference year of this study. The closest year with an estimate is 1900, for which Madison calculates GDP per capita equaled to 1237 USD 1990 dollars.

[11] For Finland, I can also make a rough comparison of my results with Enflo (2018). Both studies portray the south of Finland as the wealthiest sub-region of the country around 1900 (about one and a half above the average) and Kuopio region as the poorest. I cannot evaluate the level of similarity between the two reconstructions quantitatively because Enflo (2018) uses modern rather than historical borders.

[12] Figure B1 in the online appendix plots GRP per capita in 1897 vs. mean household income in 1904.

[13] Unfortunately, there is no systematic data on between provinces seasonal migration and remittances in the Russian empire to explore this explanation quantitatively. The bulk of such experience occurred most likely within the same province. The major exceptions were winter rural-urban migration to capitals and other large industrial centers and rural-rural summer migration to the South European provinces during harvest season.

Figure 1 somewhere here

The economic geography of the empire was eclectic. Income per capita varied widely within the empire, both in the European core and in the outskirts of the country. The overall picture is not one of a single group of economically advanced provinces surrounded by a backward periphery, but rather of a multi-focal process of economic development. Provinces with relatively high nominal income per capita were concentrated in several geographical regions. Among these were the Northwest, with Saint Petersburg, the Baltic provinces and the South of Finland; the Central Industrial Region with Moscow; the Polish provinces with Warsaw; the Southern provinces around the Black Sea; the Baku province with its oil fields; the Fergana province in Central Asia; and several provinces in Eastern Siberia and the Far East. Saint Petersburg was the richest province in terms of nominal national income per capita. Surprisingly, the easternmost and comparatively thinly populated provinces of the empire, Primorie and Sakhalin, are the second and third in terms of nominal income per capita. The steppe region, i.e. modern Kazakhstan, and the Black Earth provinces to the South of Moscow, had the lowest nominal income per capita.

As figure 2 shows, the spatial distribution of real income generally mirrors the distribution of nominal income (the coefficient of correlation between the two is 0.85 significant at the one percent level) but with important nuances. In particular, local prices help explaining the Far East paradox. In terms of real income per capita, the wealth of Primorie and other eastern provinces of the empire becomes less pronounced. In general, accounting for the difference in local prices decreases spatial variation in national income. As one could estimate from table 1, the coefficient of variation, i.e. the standard variation divided by the mean, drops from 0.56 to 0.40 when one switches from nominal to real income per capita; the corresponding figures for the US in 1900 are 0.43 and 0.39 as reported by Mitchener and McLean (1999). This disparity supports the claim that US markets were better integrated at this time.

Figure 2 somewhere here

Figure 3 presents the geography of labor productivity in the Russian empire. It mostly repeats what we have already seen in figure 2 regarding the spatial distribution of real income (the coefficient of correlation between real income per capita and labor

productivity is 0.987 significant at the one percent level). This is not a surprise given that the age structure did not vary much across provinces (Trojnitskij 1905). The only exception was, again, the provinces in the east of the empire, thinly populated with a lot of land and resources for extraction and attracting migrants. Likely because of that, their demographic composition was different, with a higher share of the population of working age and with fewer dependents, which translates into comparatively lower labor productivity. In this respect, the late 19th century Russian East was similar to the American West or to Australia in the same period of time, where a different demographic make-up due to immigration helps to account for part of the initial spatial variation in income per capita (Mitchener and McLean 1999, McLean 2007). In the Russian case, the coefficient of variation is 0.36 for labor productivity and 0.40 for real income per capita; the corresponding figures for the US are 0.33 and 0.39 (Mitchener and McLean 1999).

Figure 3 somewhere here

The differences in prices and demographic make-up reduce the variation in regional economic indicators but the remaining variation in labor productivity is substantial. Below, I present evidence on the extent to which this variation was associated with structural change. Industrialization and the level of development of agriculture and services are endogenous to economic growth, so I am not making any causal claims here. In section 5 of this paper, I address the question as to which fundamental reasons could explain the geography of labor productivity in the Russian empire in the late 19th century.

Figure 4 plots industrial output per capita (broadly defined as nominal output in large and small industry, construction and public infrastructure) in each province on the map. The observed spatial distribution generally fits the conventional story of industrialization as the main determinant of economic development in the late 19th century. Indeed, a simple visual comparison of figures 3 and 4 shows that in many provinces higher industrial output transformed into higher labor productivity. According to a classic study by Petr Lyashchenko (1956), there were six industrial centres at that time in the empire, namely the Northwest with Saint Petersburg and the Baltic provinces, the Central Industrial Region with Moscow, Poland with its textile and coal mining industries in Petrokov province, Ekaterinoslav province in the south, Baku, and the Urals. With the possible exception of the Urals, industrialized status and high labor productivity

in these regions are clearly visible in figures 3 and 4. Similarly, the Southern Finnish provinces, the provinces of Eastern Siberia and the Far East, as well as the Chernomorsk and Fergana provinces, were comparatively rich and industrialized.[14] The coefficient of correlation between industrial output per capita and productivity is 0.73 and statistically significant at the one percent level.

Figure 4 is somewhere here

However, from the comparison of figures 3 and 4, one could also easily note that a number of provinces with a relatively high labor productivity (and income per capita if one also consults figure 2) did not produce substantial industrial output in per capita terms, having a rather modest industrial sector. In particular, the southern provinces around the Black Sea as well as in the Low Volga and the North Caucasus regions were relatively productive and rich, but poorly industrialized. As figure 5 demonstrates, these provinces had relatively high labor productivity in agricultural sector (measured as value added in agriculture broadly defined, i.e. including forestry, fishing and hunting, divided by the number of people of working age in the countryside), which offered an alternative route to prosperity.[15] Agriculture was also advanced in the Baltic provinces, in Poland, and in the Yaroslavl province of the European part of the empire. The Siberian (with the exception of Yakutiya) and Far Eastern provinces, with their abundant land and space for extensive fishing and hunting activities, were also developed in terms of labor productivity in agriculture. Labor productivity in this sector was the lowest in the densely populated Black Earth region south of Moscow and in the steppe provinces of modern

[14] My reconstruction shows that the level of industrialization of these provinces was higher than has so far been acknowledged in the literature. Gold and other extractive industries in Siberia and the Far East pushed industrial output up there. The Chernomorsk province was a center of cement production. In the case of Fergana province, it was local craft industry that accounted for the industrial wealth. Finland is usually lacking in reconstructions because of its autonomous status within the Empire.

[15] Figure 5 presents a lower bound of labor productivity in agriculture due to the fact that a portion of rural citizens were employed outside agriculture. Employment of rural population outside agriculture in some industrial provinces, like Moscow, was substantial (e.g. Nafziger 2010) that push true labor productivity in agriculture up in these regions. According to surveys conducted in Smolensk, Vladimir and Vologda provinces in various years between 1893 and 1905, the share of rural population working full-time outside agriculture was 15.3 percent of rural population on average (Svavitskij and Svavitskij 1926). In other provinces, this share was most likely smaller since there was positive selection of provinces into such surveys. Agricultural output per agricultural worker as registered in the 1897 census presents an upper bar of labor productivity in agriculture (figure B2 in the online appendix). As discussed in footnote 3, the census suffers from under-enumeration of employment. The lower and upper bars of labor productivity draw roughly the same picture of variation in agricultural development across regions. The correlation between the two measures is 0.94 (significant at the one percent level).

Kazakhstan, with its substantial nomad population. I explore the relationship between labor productivity and land endowment in further details in the regression section below.

Figure 5 is somewhere here

High agricultural productivity in a number of provinces does not fit the image of an agrarian crisis in the Russian countryside after the emancipation – a widely debated interpretation of Russian agrarian development in the second half of the 19th century (Allen 2003; Gerschenkron 1965; Gregory 1982, 1994; Gatrell 1986; Mironov 2010; Nafziger 2010; Nefedov 2011). According to Gerschenkron (1965) – who put forward the idea of a crisis – the commune institution along with the legacy of serfdom harmed peasant incentives, depressing yields on communal land relative to yields on land in private ownership, and causing mass poverty and decreasing living standards in combination with rapid population growth and slow out-migration from the countryside. I show below that the legacy of serfdom indeed played a role, but its negative influence was limited to the region of its prevalence. My estimates rather fit Gatrell (1986) and Gregory's conjuncture (1994) that there was by no means an agrarian crisis in the outskirts of the empire.

Last but not least, there were a number of rich provinces like Saint Petersburg, Moscow, Warsaw, Sakhalin, and Primorie, where services constituted the largest sector of the local economy. In the European provinces, trade and banking made major contributions to GRP in terms of services, whereas the production of public goods was more important in the outskirts of the empire. The share in GRP of public services like defense or administration was particularly prominent in the sparsely populated Russian Far East, which Russia had started to colonize less than a century before the period under consideration.

4. Spatial distribution of national income in the Russian empire in international comparison

For the purpose of international comparison, in table 2, I present nominal national income per capita in international 1990 USD dollars along with summary statistics of its spatial distribution and spatial Gini-coefficients for Russia and selected countries with available subnational figures for around 1900. My reconstruction confirms that according to western standards, the late Russian empire was a poor country. Russia was more than

three and a half times poorer than the UK, the richest country at that time, more than three times poorer than the US, and more than two times as poor as other leading countries in Europe. In global comparisons, the empire was a middle-income country in terms of national income per capita being close to the level of Portugal or Mexico, and more than twice as rich as China.

Table 2 is somewhere here

Within the Russian empire, as already discussed in the previous section, there was substantial heterogeneity in economic development. GRP per capita was almost eleven times higher in the richest province of the empire (Saint Petersburg) than in the poorest one (Turgaj). This was noticeably higher than in other European countries. Similarly, another indicator of the level of inequality across regions, the spatial Gini index, was larger in Russia than in any single European country at that time. In this respect, the Russian empire was closer to other large countries in terms of territory, such as the US or Mexico. However, as the case of China suggests, where the richest province was only two and a half times wealthier than the poorest one, the size and geographical diversity of the country alone does not suffice to explain this variation at the turn of the century. Resource abundance and labor scarcity in the frontier provinces, driving up incomes per capita in these regions relative to older parts of these countries, is another potential explanation. If one considers full spatial distribution of regional incomes per capita in the countries under discussion, Mexico around 1900 would be the closest analogy to the Russian empire. Figure 6 demonstrates that both countries were comparatively poor and unequal with "long tails" on the right.

Figure 6 is somewhere here

5. Determinants of labor productivity in the Russian empire

To analyze the determinants of regional inequality in economic development, I investigate which fundamental factors accounted for the variation in labor productivity across regions. I explore the relative impacts of geographical and institutional variables on regional labor productivity rather than on real or nominal income per capita in a region in order to rule out the effects of labor inputs or local prices. In the absence of clear guidance from theory which fundamentals should be considered first, I adapt an approach implemented by Mitchener and McLean (2003) for the US. The two countries

had important similarities in the late 19th century. Both were abundant in terms of natural resources that could be important at the early stage of industrialization (e.g. Wright [1990] in application to the US case). And both countries experienced a relatively recent abolition of forced labor in the early 1860s (slavery and serfdom, respectively), institutional legacies which could slow down economic development in various ways (e.g. Mitchener and McLean 2003, Bertocchi and Dimico 2014 for the US case; Markevich and Zhuravskaya 2018 and Buggle and Nafziger forthcoming for the Russian case).

In table 3, in the spirit of Mitchener and McLean (2003), I regress labor productivity in the provinces of the Russian empire (taken in logs) on Russian analogues to the geographical and institutional variables that shaped variation in labor productivity in the US around 1900.[16] However, I do not reproduce exactly the same specification for Russia as the one which Mitchener and McLean (2003) use for the US because of data limitations and the peculiarities of the Russian case. To be precise, I use a sea dummy (rather than an ocean and Great Lakes dummy as in the US case) as an indicator of access to water routes; the share of employment in mining (estimated as the number of employees in mining from the 1897 census divided by the working age population) as a proxy for the abundance of mineral resources in a province; and the share of serfs in 1858 from Troijnitskij (1861), i.e. three years before the abolition of serfdom, (rather than the share of slaves in the US case) as a proxy for the legacy of forced labor. I explore these channels first separately (columns 1-3) and then summarily (column 4 – baseline specification). In the case of serfdom, I add a control of the distance to Moscow (in logs) because, as previous research shows, the prevalence of this institution was closely related to the proximity to Moscow (Buggle and Nafziger forthcoming; Markevich and Zhuravskaya 2018; Nafziger 2013). Remoteness from such a large market as Moscow could also affect the level of development of a province directly.[17] Mitchener and McLean (2003) do not account for the distance to major US markets in the Northeast in

[16] Section AV in the online appendix contains a detailed data description and provides sources for the variables used in the regression analysis. Table B2 in the online appendix presents their summary statistics.
[17] I weigh observations by provincial working age population to account for the relative size of provinces. I do not cluster errors due to the potential problem of the small number of clusters, which is difficult to avoid because I have only 97 observations in my sample. In the robustness section, I verify that my results are not driven by underestimation of standard errors due to potential spatial correlation (Conley 1999; Colella et al. 2018).

their approach. Instead they consider climate characteristic such as the average annual number of cooling degree days in their baseline specification. There is no such measure of historical climate for the Russian empire.[18]

Table 3 is somewhere here

I find that sea dummy and mineral recourse abundance were positively associated with labor productivity, whereas the legacy of serfdom and the distance to Moscow were negatively correlated with labor productivity. All coefficients are statistically significant at least at the five percent level.[19] I illustrate the magnitude of obtained associations using coefficients from my baseline specification (table 3 column 4) and compare them with their US 1900 counterparts (in the US comparisons, I refer to coefficients reported in column 2 of panel A of table 2 in Mitchener and McLean [2003]; I remind the reader that the two compared specifications are not exactly identical in terms of controls).

Provinces with a seashore were on average 40 percent more productive than a mean province in the Russian empire; the magnitude of the coefficient on the ocean and Great lakes dummy was about twenty percent in the US case. This result is in line with Alexander Baykov's conjecture (1954) that Russian backwardness was partially driven by the landlocked geography of a large part of the empire and the scarcity of suitable waterways. An additional percentage point of employment in mining transformed into about thirteen percent increase in labor productivity, which was much more than the equivalent one percent in the US. The larger impact of mineral resources and access to marine trade does not contradict the view that internal markets were less integrated in the Russian empire than in the US around 1900 (an alternative explanation is that Mitchener and McLean [2003] do not control for the distance to major markets). An increase in the share of serfs in 1858 by one percentage point was associated with a bit less than a half percent drop in labor productivity. The corresponding figure for the legacy of US slavery was about two percent. Taken together these factors account for almost half of the variation in labor productivity in Russia (in contrast to 77 percent in the US). As in the US case, other features of physical geography and historical climate were not correlated

[18] Please also note that the coefficient on the average annual number of cooling degree days is not statistically significant in Mitchener and McLean (2003) once they account for the legacy of slavery.

[19] Figures 3 to 5 in the online appendix report conditional scatter for the three major variables of interest obtained from the specification reported in column 4 table 3.

to labor productivity in Russia (table B3 in the online appendix).[20] I also do not find that other institutional characteristics, which I can measure at the province level, were associated with labor productivity (table B4 in the online appendix).[21]

Sea location advantage is externally determined but it could be correlated with other variables for which I do not control. If I allow the effect to vary by sea, the positive association appears for the Black Sea, the Baltic Sea and the Pacific Ocean (column 1 of table B5 in the online appendix). In contrast, the coefficient on a dummy for the arctic seas (which are frozen for more than half the year and where there were few trade routes in that time) is not statistically different from zero. I interpret these results as weak evidence in favor of the causal effect of sea advantages because of better exposure to international trade.[22]

[20] In particular, I show that ruggedness, length of rivers, mean temperature in 1901-1910, mean precipitation in 1901-10, mean relative humidity in 1901-1910, mean vapor pressure in 1901-1910, land suitability for cereals, and longitude and latitude were not associated with labor productivity. Only two out of these nine measures have significant coefficients when taken separately and none has a significant coefficient in a joint specification. In tables B6-B8 of the online appendix, I reproduce the regression analysis presented in this section with geographical controls (there, I include only latitude but not latitude because I already control for distance to Moscow in the baseline specification).

[21] To account for the nuances of local administrative institutions, I use the share of cities with 1892 city charters, *zemstvo* and *oblast* dummies. The *Zemstvo* dummy equals to one for 34 out of 97 provinces where the 1864 reform introduced elected local self-governments (*zemstvos*). The 1870 city reform (with further corrections in 1892) replaced appointed administrations by elected local city councils in about three-quarters of all urban settlements; these cities got city charters. Elected *zemstvo*s and city authorities actively invested into local education, healthcare and the provision of other public goods that could fuel economic development. The *Oblast* dummy equals to one in provinces with an *oblast* status, i.e those provinces, which were under the Ministry of War rather than the Ministry of Interior Affairs in terms of administration. In these territories, military authorities had extra rights, at the expense of civil administration. Some alternative ways to account for the difference in the local institutions, which are feasible because of data availability, are to consider the number of years the province spent as a part of the Russian state by 1897, the share of Orthodox Christians (the dominant religious in the empire), or religious diversity (measured by the Hirschman-Herfindahl index). The coefficients on all these institutional variables are not statistically different from zero in pairwise regressions.

My list of institutional characteristics hardly covers all potential institutions that could affect the economic development of the late Russian empire. In particular, I cannot properly measure and explore the effects of the prevalence and strength of different types of communes, a much debated topic (e.g. Gerschenkron [1965]; Gregory [1994] and Nafziger [2010]; Chernina et al. [2014]; Castaneda Dower and Markevich [2018] and forthcoming). A particular type of commune (e.g. repartition, hereditary, Cossack, Siberian) dominated in each single province which makes it difficult to distinguish their effect from other regional geographical covariates. I also do not have excludable instruments for the prevalence of communes of different types.

[22] In columns 2 and 3 in table B5 in the online appendix, I report results with alternative measures of access to seas and recourse abundance, namely distance to the nearest sea ports (in logs) and the share of employment in mining relative to workers registered by 1897 census. I do not use the later in the baseline specifications because of the under-enumeration of employees in the census. I prefer the sea dummy to minimal distance to major ports because the location of ports might be endogenous.

Similarly, the share of the labor force employed in mining and especially the prevalence of serfdom might be endogenous. In addition, employment in mining might be imprecisely measured because of under-registration issues in the 1897 imperial census. The legacy of serfdom proxy could also suffer from measurement errors because the share of serfs in 1858 reflects the geography of serfdom in one particular year rather than measuring its geography over the more than two hundred years of its existence. All that, not to mention the potential omitted variables problem, could bias the obtained coefficients on the share of employment in mining and the share of serfs.

I use the instrument variables approach to overcome these difficulties and run the same specification as reported in column 4 of table 2. In particular, in column 5 of table 2, I use the number of major deposit fields per unit of area discovered by 1920, the earliest year for which such information is available, as an instrument for employment in the mining sector.[23] The employed instrument is valid under the assumption that pioneering works in mining and, correspondingly, the amount of uncovered mineral resources in a province was not correlated with the regional level of economic development. That was arguably true due to the practice of scientific expeditions specifically designed by the Russian Academy of Science and imperial ministries to discover and exploit regions all over the empire. The coefficient on the share of employment in mining in the IV regression remains positive and statistically significant while its magnitude is about one and a half larger than in the OLS specification. This is in line with the presence of a measurement problem (an alternative explanation of the difference between OLS and IV estimates is reverse causality, i.e. conditional on the presence of natural resources, the mining sector was more developed in more productive regions due to higher potential returns). I report first-stage results in panel B of table 2 below the corresponding second-stage results. Deposit fields are positively correlated with employment in mining and the magnitude of F-statistic is 42, i.e. well above 10, which is considered as the rule-of-thumb benchmark level suggesting no weak instrument problem (Stock and Yogo 2005). To address the identification strength of my instrument in a general case when the data are allowed to be heteroskedastic, I implement a two-

[23] To be precise, I account for copper, gold, diamond, feldspar, gypsum, iron, kaolin, lead, limestone, manganese, phosphorous, potash, and zinc deposit fields.

step-weak-iv test recently developed by Sun (2018) and based on a theoretical approach suggested by Andrews (2018). The distortion cutoff level obtained by this test is five percent, indicated a lack of a weak instrument problem.[24]

To instrument the legacy of serfdom, I adjust the approaches suggested by Buggle and Nafziger (2016) and Markevich and Zhuravskaya (2018), who use either the number of monasteries in the past or the share of serfs that belonged to monasteries and clergy before the nationalization of church lands by the Russian state (most of which took place in 1764). The nationalization changed the legal status of former monasterial serfs to state peasants that might be viewed as a negative exogenous shock for the prevalence of serfdom. Monasteries accumulated the majority of their lands before the introduction of serfdom and monasterial serfs before the nationalization did not differ systematically from other serfs. As discussed in Buggle and Nafziger (2016) and Markevich and Zhuravskaya (2018), the geography of monasteries and their ownership were most likely not related to other unobserved factors that might influence Russian economic development around the emancipation or in the long-run other than via the prevalence of serfdom channel. However, these authors focus only on European provinces of the Russian empire, which were overwhelmingly orthodox, with a positive number of orthodox monasteries, and accordingly, shares of monasterial serfs before the nationalization in almost each single province. In contrast, serfdom was never introduced or abolished before the arrival of Russians in many non-orthodox territories outside the European core of the empire, which they conquered comparatively late, like Poland, Finland or Central Asia; there were also no orthodox monasteries and monasterial serfs in these territories. Likely because of that, the IVs used by Buggle and Nafziger (2016) and Markevich and Zhuravskaya (2018) – the number of monasteries before the nationalization and the share of former monasterial serfs – extended to the whole territory of the empire and taken separately do not show a statistically significant negative correlation with the prevalence of serfdom. To account for both the size and the probability of the emergence of monasteries' serfs ownership in a province, i.e. for both intensive and extensive margins of nationalization of church property in the 18th century, I use two instruments for the prevalence of serfdom. Namely, I use the share of former

[24] To be precise, I use Stata command *twostepweakiv* constructed by Sun (2018).

monasterial serfs in a province from Markevich and Zhuravskays (2018) extended to the whole territory of the empire with the data from Beskrovnii et al. (1972) and a dummy for the presence of monasteries in a province at the moment of the nationalization of church property in 1764 obtained from Zverinskij (1890-1897).[25]

After instrumenting, the coefficient on the legacy of serfdom remains negative and statistically significant.[26] The IV point estimate is about two and a half times larger than its OLS counterpart. One should use caution interpreting the size of the effect because of the weak instrument problem. I report the first-stage results in panel B of table 2 below the relevant second-stage specification. With two instruments, Kleibergen-Paap rk Wald F statistic is 5.964 and the cut-off level obtained by two-step-weak-iv test is thirteen percent. I report LC_2sls 95 percent confidence interval for the IV coefficient obtained by two-step-weak-iv test that is robust to weak instruments. The interval [-2.53, -0.44] overlaps with the confidence interval for the OLS estimate. The difference between OLS and IV point estimates might also arise because of potential mis-measurement of historical serfdom or omitted characteristics of the provinces important for economic development that leads to the underestimation of the effect in the OLS specification (e.g. it might be quite possible that there were more serfs in provinces with better economic prospects due to stronger incentives for landowners to privatize rents).

Under the assumption of excludability of my instruments, the obtained results suggest a causal link between the legacy of serfdom and labor productivity in 1897 in the Russian empire, which is in line with findings by Buggle and Nafziger (forthcoming) on the negative long-run consequences of serfdom. An alternative interpretation would be that it was the 1861 emancipation reform that worsened the economic prospects of those Russian regions that had relied more on forced labor before its abolition. Such a view would be in parallel with the story on economic crises in the postbellum US South. However, findings by Markevich and Zhuravskaya (2018) on the positive effects of the abolition of serfdom rules out such an interpretation for the Russian case. In table 4 I

[25] I also present suggestive evidence addressing potential concerns regarding the excludability assumption of my instruments. In particular, Hansen J statistics (0.612) does not reject the validity of overidentification restrictions for my two instruments.

[26] In column 4 of table B5 in the online appendix I simultaneously instrument both the share of employment in mining and the legacy of serfdom. The results hold.

present more suggestive evidence regarding the channels via which the legacy of serfdom slowed down economic development.

Table 4 is somewhere here

To explore the potential mechanisms through which the legacy of serfdom impacted economic development, I add additional regressors into my baseline specification from column 4 of table 3. In particular, I account for regional variation in the size of serf estates, the distribution of land by tenure type after the implementation of the emancipation land reform, and the post-emancipation development characteristics of the Russian provinces such as exposure to migration, urbanization and literacy. All these variables are likely endogenous to economic development, and I do not have excludable instruments for them, so the results presented in table 4 should be viewed as suggestive.

I start with inserting the share of large estates (with more than twenty serfs, obtained from Troijnitskij [1861]) in a province before emancipation as an additional control (column 1). The coefficient on this control is negative and statistically significant at the five percent level. This does not contradict a conjecture that the legacy of serfdom was more pronounced in large estates, suggesting parallels between the Russian empire and Latin America in terms of the negative effect of inequality and large ownership on development (Engerman and Sokoloff 2000, Dell 2010). Other things equal, an increase in the share of large serf estates by one percentage point was associated with a drop in labor productivity by about 0.21 percent. By contrast, the coefficient on the share of serfs becomes twice as small and statistically insignificant. The two variables are jointly significant. Under the assumption that the share of large estates in a province was hundred percent, one cannot reject a conjecture that the combined effect of these two variables (-0.41=(-0.21)+(-0.2)) was as large as the effect of the legacy of serfdom from the baseline specification (-0.44). Again, there are two possible interpretations of this result. The legacy of serfdom could be more pronounced in large estates because wealthy landlords continued to dominate local markets and de facto kept their power over peasants after emancipation. Another possibility is that, for some reason, the emancipation reform disorganized production specifically in large estates. In the latter case, however, one should also explain why this disorganization of production had a downward effect on labor productivity which lasted more than thirty years.

In column 2 of table 4, I explore whether the legacy of serfdom affected labor productivity via the difference in land tenure introduced by the land reform subsequent to the emancipation. As a part of the abolition of serfdom, former serfs had to redeem lands from their former lords. The exact amount of redeemed land was the outcome of a local bargaining process subject to legal norms, which varied across regions. However, one feature was common. The law upheld the institution of the commune and stipulated that local communes rather than individual former serfs became the legal owners of the land ceded by the gentry. I control for the share of such communal landholdings along with the share of state lands in a province (leaving the share of private lands as a control group). The coefficients on both these variables are negative and statistically significant at the one percent level. While the point estimate for commune lands is almost one and a half times larger than for state lands, the conjecture that the effects of these two variables are equal to each other cannot be rejected. This means that a higher share of lands with private tenure in a province was associated with higher labor productivity, whereas commune and state lands were equally less productive. These results do not contradict Alexander Gerschenkron's (1965) hypothesis that the commune and communal land ownership had a negative impact on Russian economic development. Gerschenkron (1965) argued that the commune regulated peasant use of land, e.g. crop rotation, time and duration of field-works, which could hinder the adoption of innovations. Also, in repartition communes (about 80 percent of all communes in European Russia) periodical redistribution of land within the commune undermined peasant incentives to invest in the quality of the land. An alternative interpretation would be that the land reform transferred too little land to peasant farmers, which limited their opportunities to employ advanced technologies and to explore economies of scale, thus resulting in lower efficiency. Also, it might be that in more productive regions landlords managed to keep a higher share of lands in their possession after the emancipation. Whatever the true channel was, this channel does not fully account for the negative legacy of serfdom since the coefficient on the share of serfs in column 2 remains negative, statistically significant and even increases in magnitude relative to the baseline specification.

Besides landownership, the legacy of serfdom could work via other channels associated with the commune. In particular, after the emancipation peasants continued to

face constraints on migration. They could not leave their villages either permanently or temporarily without consent of the commune, which might have resulted in the misallocation of labor away from more productive regions and sectors, as well as low labor productivity in labor abundant areas (Chernina et al. 2014). At the same time, a number of studies argue that due to mutual tax responsibility it was in the best interest of the commune to allow out-migration from the village to lower population pressure on the land, which could have reduced the importance of these legal constraints (Gregory 1994 and Nafziger 2010). There is no regular statistics on temporarily peasant migration (*otkhod*), but I can use the regional share of in-migrants as well as the share of out-migrants known from the 1897 census to control for more permanent type of resettlement (column 3). As one could expect, there were on average more in-migrants in provinces with higher labor productivity and more out-migrants in less productive provinces (the coefficient on out-migrants is not statistically different from zero). However, it is difficult to make any causal conclusion from the observed relationships due to the endogeneity of migration measures. Nevertheless, given that controlling for migration reduces the coefficient on the share of serfs almost by forty percent, one could argue that the negative legacy of serfdom for labor productivity was partially driven by post-emancipation constraints on peasant resettlement.

The columns 4 and 5 present evidence that the legacy of serfdom operated via urbanization and human capital channels, correspondingly. The urbanization level and literacy rates were positively associated with labor productivity. Accounting for these variables reduces the coefficient on the share of serfs substantially. These results support findings by Buggle and Nafziger (forthcoming) on the negative long-run consequences of serfdom on economic development through slower structural change.[27]

On top of the direct effects of spatial physical characteristics and human-made institutions, geography and social infrastructure can affect economic development in various combinations. In table 5, I explore to what extent the so-called second-order geography resulting from human interventions shaped labor productivity in the Russian

[27] Accounting for the urbanization or literacy channels also substantially decreases the magnitude of coefficients on the geographical characteristics, namely the sea dummy and distance to Moscow (the coefficient on the sea dummy in the later case also becomes not statistically different from zero). One interpretation of these results is that prospects to trade either with foreign markets or the largest internal market contributed to urbanization and the accumulation of human capital.

empire (on top of the factors discussed above). The new economic geography literature argues that good access to markets, specialization and economies of scale effects could boost economic development. There is a growing amount of evidence on the importance of these forces in application to other countries in the late 19th – 20th century (Wolf 2007; Klein and Crafts 2011; Combes et al. 2011; Martinez-Galarraga 2012; Liu and Meissner 2015; Donaldson 2018). In line with this literature, I find that access to markets in other provinces and foreign countries measured as a sum of their regional and national incomes weighted by inverse distance was positively associated with labor productivity in the late 19th century Russian empire (column 1).[28] To overcome a possible endogeneity of market potential, I follow the literature and use the sum of inverted distances to other regions and foreign countries as an instrument (column 2). I report the first-stage results in panel B of table 5 below the corresponding column.[29] The IV coefficient remains statistically significant and the same in terms of magnitude. Under the assumption of instrument validity, these results imply that a one percent increase in market potential caused about a 0.21 percent increase in labor productivity.

Table 5 is somewhere here

In column 3, on top of market potential, I control for population density (in log) as a proxy for the potential of the province's own market. In this specification, the coefficient for density is positive and statistically significant while the coefficient for the market potential variable reflecting access to markets in other provinces and foreign countries is insignificant and changes its sign. These results suggest that the potential of the province's own market was more important. An alternative interpretation would be that population density proxies other channels of positive externalities of scale. Both interpretations, however, should be viewed as suggestions only, because I do not have an excludable instrument for density, which is endogenous.

[28] I cannot estimate market potential using the gravity equation framework because of a lack of internal trade flows data. To be precise, I calculate the internal market potential as the inverse distance-weighted sum of the nominal incomes of the other provinces of the Russian empire (known form the present study). To account for foreign markets I calculate the inverse distance-weighted sum of GDPs of foreign countries taken from Maddison's dataset (2010). I sum up the two figures to get total market potential of a province and take the log from this. I do not add the potential of the province's own market because of the possible correlation with labor productivity.

[29] F-statistic is high (195.6) and the distortion cut-off level obtained by two-step-weak-iv test is five percent.

Whatever the interpretation is, the effect of density was driven by urban rather than rural density (column 4). A positive and statistically significant coefficient on urban density points to the importance of urban agglomerations. I interpret rural density as an inverse measure of land abundance. The negative coefficient on this measure is in line with a view of the late Russian empire as an open frontier economy, where land abundance resulted in higher labor productivity (to account for the difference in land quality I control for land suitability for cereals in this specification). A simple redistribution of labor relative to land across provinces, like the post-1906 migration from the European to the Asian parts of the empire triggered by the Stolypin reform, could spur on economic growth (Chernina et al. 2014). Given these results and the results from table 4 on the consequences of serfdom and resettlement, one might speculate that it was the legacy of serfdom that prevented such redistribution. In this sense, serfdom not only slowed down economic development of those regions where it was present, but was also likely the reason why some regions remained relatively sparely populated and rich (due to high marginal returns on labor) at the end of the imperial period. A more than 50 percent drop in the magnitude of the coefficient for the share of serfs in column 4 relative to the magnitude of this coefficient in the baseline specification supports such an interpretation.

I provide evidence how regional specialization and the economies of scale effect were related to labor productivity in columns 5 and 6, respectively. I measure specialization with Krugman specialization indexes constructed separately for agriculture, industry and services from the present study's estimates of values added in the corresponding subsectors.[30] The Krugman index is a measure of relative

[30] For each sector (industry, agriculture and services), I construct the Krugman specialization index in two steps. First, for each province, I estimate a set of its pairwise specialization metrics reflecting the difference between this province and each province from the set of other provinces in terms of the composition of value added in the corresponding sector. At the second step, I take the average over these pairwise metrics for each of the corresponding sectors separately.

To be precise, I calculate pairwise metric of similarity between provinces j and k as $M_{jk} = \sum_{i=1}^{n} \left| \frac{VA_{ij}}{VA_j} - \frac{VA_{ik}}{VA_k} \right|$ where VA_{ij} is the value added in subsector i=1,..,n for province j, and VA_j is the value added in the corresponding sector for province j, and similarly for province k (if the pairwise metric is equal to zero, this means that two provinces are completely the same in terms of composition of value added; if the pairwise metric is equal to two, the provinces are completely different). At the second step, I estimate the index for r province j as $I_J = \frac{\sum_{k \neq j} M_{jk}}{n-1}$. Table A64 in the online appendix presents the estimates of the Krugman

specialization reflecting how much a particular region differs from the others in terms of the composition of the corresponding sector. A positive and statistically significant coefficient on the industry specialization index is in line with the presence of Marshallian externalities in this sector. Similarly, there is evidence that they were also in operation in agriculture (but at a smaller scale as point estimates suggest), which well fits this paper's finding that highly productive agriculture constituted a leading sector of the economy in a number of provinces of the empire. In contrast, specialization in particular subsectors in services was correlated negatively with labor productivity. More productive regions developed all services rather than specializing in particular types of services.[31] I use the mean firm revenue in industry in a region in 1894 from Gregg (2019) to proxy scales. Economies of scale in industry were positively associated with labor productivity. The results presented in column 5 and 6 should be considered with caution because specialization indexes and firm size are endogenous to labor productivity. I do not have good candidates to instrument these variables.

5.1. Robustness of determinants

This subsection briefly describes the results of the sensitivity tests that I ran to check the robustness of my findings on the determinants of spatial variation in economic development of the Russian empire. I present tables with full-scale results in the online appendix.

First, I explore how sensitive my results are to an addition of geographical controls (tables B6 – B8 in the online appendix). With an exception of the effect of market potential and population density that were likely correlated with geography, the results are stable. Second, I show that my estimates are robust if I correct standard errors for spatial correlation as discussed in Conley (1999) and Collela et al. (2018). In table B9 – B11, I allow for correlation across space within a radius of 1000 kilometers, the distance equals to more than one fifth of the North-South and more than one tenth of the West-East dimensions of the territory of the former Russian empire (I experimented with smaller and larger thresholds that did not change the results). Third, I repeat my analysis

specialization indexes.

[31] I also notice that controls for specialization undercut the significance of the coefficients for the share of serfs and distance to Moscow and reduce the magnitude of the coefficients for the sea dummy and the share of employment in mining. This implies that geographical and institutional characteristics shaped regional industrial structures.

for the alternative estimates of labor productivity (based on my estimates of nominal GRPs when I simultaneously relax the assumptions, which might push my baseline estimations up, as discussed in section 2 and section AIV in the online appendix). Tables B12 – B14 in the online appendix present the results of the same set of exercises I run in table 3 – 5, but with the new dependent variable. Forth, in tables B15 – B17 of the online appendix, I show that my results are consistent if I use labor productivity in levels rather than in logs.

6. Conclusion

This paper documents and explores the regional variation in economic development of the Russian empire around 1900. My reconstruction confirms that Russia in that time was one of the poorest countries in Europe and a middle-income country in the world, but also reveals substantial spatial variation in national income and labor productivity across regions. The capital province was the wealthiest and the most productive one, but the geographical pattern of economic activities was not as simple as a single core and a withering periphery. Rather, there were several advanced territories, in various parts of the empire, with an arguably good potential for growth. In terms of the variation within the country, the Russian empire at the turn of the century was comparable to countries of the New World, such as the US and Mexico, at the same moment of history.

My analysis of the determinants of the observed variation within the country shows that both geography and institutions mattered. In a striking parallel to the US case as explored by Mitchener and McLean (2003), I find that access to international waters, natural resources and the legacy of forced labor were important factors accounting for the variation in labor productivity and income per capita in the Russian empire. Based on the estimated coefficients on proxies for these factors, I implement a naïve counter-factual exercise for the purpose of illustrating the relative importance of these channels.[32] If, other things equal, each province in the Russian empire (rather than twenty three out of ninety-nine provinces) would have had a maritime border, labor productivity in the

[32] I calculate counter-factual figures for each province, multiplying the estimated coefficients from table 3 on the sea dummy, the share of the labor force in mining, and the share of serfs by regional values of the corresponding variables and deduct the obtained products from the true figures (I use the OLS coefficient from column 4 for the sea dummy; the IV coefficient for the share of labor force in mining from column 5 and either OLS or IV coefficients for the share of serfs from columns 4 and 6). Then, I take the averages over obtained regional counter-factual estimates.

average imperial province in 1897 would have been 34.5 percent higher (and real income per capita would increase by 35.2 percent). The absence of mineral resources and a mining sector would lead to a drop in labor productivity of 8.3 percent (and a drop in real income per capita of 8 percent). In a counter-factual Russia without serfdom, labor productivity in 1897 would have been between 8.9 to 27.2 percent higher (depending whether one uses OLS or IV coefficients) and real income per capita would have been 9.7 to 27.4 percent larger.

References:

Acemoglu, Daron and James Robinson (2012). *Why Nations Fail: the Origins of Power, Prosperity and Poverty,* New York: Crown Publisher.

Aguilar-Retureta, José (2015). 'Regional income distribution in Mexico: new long-term evidence, 1895-2010'. *UB Economics Working Papers 2015/323.*

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

Andrews, I. (2018). Valid two-step identification-robust confidence sets for GMM. *Review of Economics and Statistics,* 100: 337–348.

Baykov, Alexander (1954). 'The Economic Development of Russia', *Economic History Review*, 7(2): 137-49.

Beskrovnii L.G., Ya.E. Vodarskii, and V.M. Kabuzan. (1972). *Perepisi naseleniya Rossii: itogovie materiali podvornikh perepisei I revizii naseleniya Rossii, 1646 – 1858* [Russian Population Censuses: results of household censuses and tax censuses in Russia, 1646 - 1858]. Moscow: Institut istorii AN SSSR.

Bertocchi, Graziella and Arcangelo Dimico (2014). 'Slavery, education, and inequality', European Economic Review, 70: 197-209.

Buggle, Johannes and Steven Nafziger (2016). 'Long-Run Consequences of Labor Coercion: Evidence from Russian Serfdom', *Working paper.*

Buggle, Johannes and Steven Nafziger (forthcoming). 'The Slow Road from Serfdom: Labor Coercion and Long-Run Development in the Former Russian Empire', *Review of Economics and Statistics*.

Caruana-Galizia, Paul and Ye Ma (2016). 'Chinese Regions in the Great Divergence: Provincial Gross Domestic Product per capita, 1873–1918', *Australian Economic History Review*, 56(1): 21-45.

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

Castañeda Dower, Paul and Andrei Markevich (forthcoming). "The Stolypin Reform and Agricultural Productivity in Late Imperial Russia," *European Economic History Review.*

Central statistical agency (1906). *Goroda Rossii v 1904 g.* [Russian cities in 1904]. Saint Petersburg.

Cheremukhin A., M. Golosov, S. Guriev and A. Tsyvinski (2018). 'The Industrialization and Economic Development of Russia through the Lens of a Neoclassical Growth Model', *Review of Economic Studies*, 84(2): 613-649.

Chernina, Eugenia, Paul Castañeda Dower and Andrei Markevich, (2014). 'Property Rights and Internal Migration: The Case of the Stolypin Agrarian Reform in the Russian Empire', *The Journal of Development Economics,* 110: 191-215.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli and Mathias Thoeniget (2018). 'Inference with Arbitrary Clustering', *Working paper*.

Combes, Pierre-Philippe, Miren Lafourcade, Jacques-François Thisse and Jean-Claude Toutain (2011). 'The rise and fall of spatial inequalities in France: A long-run perspective', *Explorations in Economic History,* 48: 243-71.

Commision on Peasnats Welfare (1901). *Materialy komissii dlya viyasneniya voprosa o blagosostoyanii krestyan s 1861 g.* [Materials of the commission to study wealth of peasants since 1861]. Saint Petersburg.

Conley, Timothy G. (1999). 'GMM Estimation with Cross Sectional Dependence,' *Journal of Econometrics* 92, 1-45.

Dell, Mellisa (2010). 'The Persistence Effects of Peru's Mining *Mita*', *Econometrica*, 78(6): 1863–1903.

Donaldson, Dave (2018). 'Railroads of the Raj: Estimating the Impact of Transportation Infrastructure', *American Economic Review*, 108(4-5): 899‑934.

Enflo, Kerstin (2019). 'Balancing east and west: evidence from Finland's regional GDPs, 1880-2010' in Rosés, Joan Ramón and Nikolaus Wolf (Eds.). *The Economic Development of Europe's Regions A Quantitative History since 1900*. London: Routledge.

Engerman Stanley L. and Kenneth L. Sokoloff (2000). 'Institutions, Factor Endowments, and Paths of Development in the New World', *Journal of Economic Perspectives*, 14(3): 217-232.

Fujita, M., Paul Krugman and Anthony J. Venables (1999), *The Spatial Economy*, MIT Press.

Gallup, Jonh Luke, Jeffrey D. Sachs and Andrew D. Mellinger (1998). 'Geography and Economic Development', *NBER Working Paper* No. 6849.

Gatrell, Peter (1986). *The Tsarist Economy: 1850-1917*. New York: St. Martin's Press.

Geary, Frank and Tom Stark (2002). 'Examining Ireland's Post-Famine Economic Growth Performance', *The Economic Journal,* 112(482): 919-935.

Gerschenkron, Alexander, (1962). *Economic backwardness in historical perspective, a book of essays*, Cambridge, Massachusetts: Belknap Press of Harvard University Press.

Gerschenkron Alexander (1965). 'Agrarian Policies and Industrialization: Russia 1861-1917', The Cambridge Economic History of Europe (Cambridge, 1965), vol. 6 p. 2.

Gozulov A.I. (1936). *Perepisi naseleniya SSSR i kapitalisticheskikh stran (opit istoriko-metodologicheskoj kharakteristiki proizvodsvta perepisej naseleniya)* [Population censuses in the USSR and capitalist countries]. Moscow: TsSUNKhU press, 1936.

Gregg, Amanda. (2019). "Factory Productivity and the Concession System of Incorporation in Late Imperial Russia," *Working Paper.*

Gregory Paul (1982). *Russian National Income, 1885-1913*. Cambridge: Cambridge University Press.

Gregory, Paul (1994). *Before Command: The Russian Economy from Emancipation to Stalin*. Princeton, NJ: Princeton University Press.

Gregory, Paul (2001). '*Ekonomicheskaya istoriya Rossii: chto mi znaem o nej i chego ne*

znaem. Otzenka ekonomista' [Russian economic history; what we know and what we do not know. Economist's notes] in *Ekonomicheskaya istoriya. Ezhegodnik 2000* [Economic history. Annual volume 2000]. Moscow: Rosspen.

Grinin L.E., A.V. Korotaev, and A.V. Malkov (2010). *O prichinakh russkoj revolutsii*. [About the causes of the Russian revolution]. Moscow: LKI.

Groman V.V. (1912) *Obzor stroitelnoj deyatelnosti v Rossii* [Review of construction industry in Russia]. Saint Petersburg.

Klein, Alexander and Nicholas Crafts (2011). 'Making sense of the manufacturing belt: determinants of U.S. industrial location, 1880–1920', *Journal of Economic Geography*, 12(4): 775–807.

Klein, Alexander (2013). 'New State-Level Estimates of Personal Income in the United States, 1880–1910.' *In Research in Economic History*, pp. 191-255. Emerald Group Publishing Limited.

Lindert, Peter H. and Nafziger, Steven (2014). 'Russian Inequality on the Eve of Revolution', *Journal of Economic History*, 74(3): 767-798.

Liu, Dan and Christopher M. Meissner (2015). "Market potential and the rise of US productivity leadership', *Journal of International Economics*, 96: 72-87.

Lyashchenko P.I. (1956). *Istroiya narodnogo khozyastva SSSR* [History of national economy of the USSR]. Moscow.

Maddison, Angus. (1995). *Monitoring the World Economy 1820–1992*. OECD Development Centre, Paris.

Maddison, Angus. (2010). World population, GDP and per capita GDP, 1-2003 AD. Available at: http://www.ggdc.net/maddison/oriindex.htm.

Markevich, Andrei and Mark Harrison (2011). 'Great War, Civil War, and Recovery: Russia's National Income, 1913 to 1928', *Journal of Economic History,* 71(3): 672-703.

Markevich, Andrei and Tatyana Mikhailova (2013). 'Economic Geography of Russia', in Alekseev, M., and Weber, S. (Eds.) *The Oxford Handbook of the Russian Economy*. Oxford: Oxford University Press.

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

Martinez-Galarraga, Julio (2012). 'The determinants of industrial location in Spain, 1856–1929', *Explorations in Economic History,* 49(2): 255-275.

McLean, Ian W. (2007). 'Why was Australia so rich?' *Explorations in Economic History,* 44: 635–656.

Mironov B.N. (2010). *Blagososotoyanie naseleniya I revolutsii v imperskoj Rossii*, Moscow: Novij Khronograf.

Mitchener, Kris James and Ian W. McLean (1999). 'U.S. Regional Growth and Convergence, 1880-1980', *Journal of Economic History,* 59(4):,1016-1042.

Mitchener, Kris James and Ian W. McLean (2003). 'The Productivity of US States since 1880', *Journal of Economic Growth* 8(1): 73-114.

Nafziger, Steven (2010). 'Peasant communes and factor markets in late nineteenth-century Russia', *Explorations in Economic History*, 47(4): 381-402.

Papyrakis, Elissaios (2017). 'The Resource Curse - What Have We Learned from Two Decades of Intensive Research: Introduction to the Special Issue', *The Journal of*

Development Studies, 53(2): 175-185.
Prokopovich S.N. Ed. (1918). *Opyt ischisleniya narodnogo dokhoda 50 gubernii Evropejskoj Rossii v 1900-1913* [Estimations of national income of 50 provinces ofEuropean provinces in 1900-1913]. Moscow.
Rosés, Joan Ramón and Nikolaus Wolf (Eds.). (2019). *The Economic Development of Europe's Regions A Quantitative History since 1900*. London: Routledge.
Salomatina, Sofia (2014). "Rekonstruktsiya dannikh v istoricheskoj statistike: otdeleniya kommercheskokh bankov Rossijskoj imperii v 1897 g." [Reconstruction of historical statistics: branches of commercial banks in the Russian empire in 1897], *Vestnik Permskogo universiteta. Seria "Istoriya",* 26(3): 87-98.
Silantiev A.A. Ed. (1898). *Obzor promislovikh okhot v Rossii* [Review of hunting in Russia]. Saint Petersburg.
Stock, J. H., and M. Yogo. (2005). 'Testing for weak instruments in linear IV regression'. In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, ed. D. W. K. Andrews and J. H. Stock, 80–108. New York: Cambridge University Press.
Strumilin S.G. (1924). Tovarooborot v Rossii. In *Ekonomicheskoe obozrenie*. Vol. 19. Moscow.
Svavitskij Z.M. and N.A. Svavitskij (1926) *Zemskie podvornie perepisi, 1880-1913. Pouezdnie itogi* [Zemstvo household surveys, 1880-1913. District-level results]. Moscow: TSU.
Sun, Liyang (2018). Implementing valid two-step identification-robust confidence sets for linear instrumental-variables models, *The Stata Journal* 18(4): 803–825.
Trojnitskij A.G. (1861). *Krepostnoe naselenie po 10 narodnoj perepisi* [Serf population according to the 10th tax census]. Saint Petersburg.
Trojnitskij N. A. (ed.). (1905). *Obshchij svod po imperii resultatov razrabotki dannikh pervoj vseobshchej perepisi naseleniya, proizvedennoj 28.01.1897* [The First Population 1897 Census of the Russian Empire. General results] Vols.1-2. Saint Peterburg.
Trojnitskij N. A. (ed.). *Pervaya Vseobshchaya perepis' naseleniya Rossijskoj imperii 1897 goda. (Gubernskie itogi)*. [The First Population 1897 Census of the Russian Empire (Province-level results)] Vols.1-89. Saint Petersburg, 1903-1905
Vainshtein A.L. (1960). 'Iz istorii predrevolutsionnoj statistiki zhivotnovodstva' [From history of pre-revolutionary statistics of husbandry] in *Ocherki po istorii statistiki SSSR*. Moscow.
Wolf, N. (2007) 'Endowments vs. market potential: What explains the relocation of industry after the Polish reunification in 1918?'' *Explorations in Economic History*, 44: 22–42.
Wright, Gavin (1990). 'The Origins of American Industrial Success, 1879-1940', *The American Economic Review,* 80(4): 651-668.
Yasnopolskij N.P. (1891, 1897) *O geographisheskom raspredelenii gosudarstvennikh dokhodov i raskhodov Rossii* [On geography of state incomes and expenditures], Kiev. Vol. 1-3.
Zverinskij V.V. (1890-1897). *Material dlya istoriko-topographicheskogo issledovaniya o pravoslavnikh monastiryakh Rossijskoj imperii* [Materials for historical and geographical studies on orthodox monasteries in Russian empire]. Saint-Petersburg. Vol. 1-3.

Table 1. Provinces of the Russian empire in 1897: Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Gross regional products in thousand roubles (nominal)	97	100141.8	89506.2	5697.9	641476.7
Gross regional product per capita in roubles, nominal	97	82.6	45.9	28.2	303.7
Gross regional product per capita in roubles, real	97	78.9	31.5	35.5	226.3
Labor productivity (output per citizen in the working age), roubles	97	150.7	53.8	68.98	356.7
Share of value added in agriculture (nominal)	97	0.57	0.16	0.05	0.82
Share of value added in industry (nominal)	97	0.19	0.13	0.03	0.62
Share of value added in services (nominal)	97	0.25	0.09	0.08	0.67
Population	97	1321843	846421.6	28113	3559229
Share of population in the working age	97	0.53	0.05	0.47	0.74
Relative price index of consumption goods	97	1.02	0.26	0.76	2.35
Urban share	97	0.13	0.099	0	0.67

Sources: Population and urban share - Trojnitskij N. A. (ed.). Pervaya Vseobshchaya perepis' naseleniya Rossijskoj imperii 1897 goda. (Gubernskie itogi). [The First Population 1897 Census of the Russian Empire (Province-level results)] Vols.1-89. Saint Petersburg, 1903-1905; other indicators - the current study estimates.

Table 2. The Russian empire around 1900 in international comparisons

	GDP	Number	GRP per capita, 1990 USD						Spatial
Country	per capita	of regions	Mean	St. Dev.	Min	Region with the lowest GRP	Max	Region with the largest GRP	Gini
Russia (1897)	1215	97	1325	736	453	Turgaj	4873	Saint Petersburg	0.25
US (1900)	4073	48	4113	1768	1437	North Carolina	8263	Montana	0.22
UK (1900)	4606	10	4333	871	3098	Northern Ireland	6489	London, UK	0.09
Germany (1900)	2 831	36	2822	611	1850	Tübingen, Regierungsbezirk	4420	Bremen, Germany	0.11
France (1900)	2 966	22	2726	854	1187	Corse	5688	Ile de France	0.16
Portugal (1900)	1 492	5	1535	545	1159	Algarve	2493	Norte	0.12
Mexico (1900)	1366	29	1557	771	560	Guerrero	4248	Baja C. North	0.26
China (1893)	545	24	531	115	270	Yunnan	716	Heilungkiang	n.a.

Sources: Russia and Russian provinces (1897) – the current study; US regions (1900) - Klein, Alexander (2013). "New State-Level Estimates of Personal Income in the United States, 1880–1910." In Research in Economic History, pp. 191-255. Emerald Group Publishing Limited; UK, Germany, France, Portugal (1900) - Rosés Joan R. and Nikolaus Wolf. Eds. (2018). 'The Economic Development of Europe's Regions: A Quantitative History since 1900', Oxon&New York: Routledge; China (1893) - Caruana-Galizia and Ye Ma (2016). 'Chinese Regions in the Great Divergence: Provincial Gross Domestic Product per Capita, 1873–1918 Australian Economic History Review, Vol. 56, No. 1; Mexico (1900) - José Aguilar-Retureta (2015). 'Regional income distribution in Mexico: new long-term evidence, 1895-2010' UB Economics Working Papers 2015/323. National GDPs per capita for Russia, US, UK, Germany, France and Portugal are a weighted mean of corresponding GRPs. National GDPs per capita for China and Mexico are from Angus Maddison dataset.

Table 3. Determinants of labor productivity in the Russian Empire: Geography and Institutions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	1897 labor productivity (log)					
	OLS	OLS	OLS	OLS	IV, 2nd stage	IV, 2nd stage
Sea dummy	0.45***			0.40***	0.38***	0.34***
	[0.102]			[0.105]	[0.120]	[0.119]
Share of employment in mining		13.93**		13.23***	20.59***	14.06***
		[5.804]		[4.716]	[7.369]	[4.590]
Share of serfs in 1858			-0.59***	-0.44***	-0.45***	-1.12***
			[0.144]	[0.124]	[0.118]	[0.407]
LC_2sls 95 percent confidence interval						[-2.52627,-.440696]
Log distance to Moscow			-0.18***	-0.22***	-0.23***	-0.33***
			[0.045]	[0.039]	[0.036]	[0.074]
Constant	4.87***	4.91***	6.34***	6.41***	6.50***	7.34***
	[0.038]	[0.045]	[0.344]	[0.287]	[0.257]	[0.595]
Observations	97	97	97	97	97	97
R-squared	0.262	0.080	0.152	0.467	0.447	0.331
Panel B: First stages of 2SLS regressions						
Dependent var:					Share of employment in mining	Share of serfs in 1858
Number of deposit fields per sq km					124.53***	
					[19.291]	
Share of (former) monasterial serfs						-0.46
						[0.367]
Dummy for the presence of monasteries in 1764						0.20***
						[0.057]
Sea dummy					0.00	-0.05
					[0.002]	[0.055]
Share of employment in mining						0.00
						[2.382]
Share of serfs in 1858					0.00	
					[0.003]	
Log distance to Moscow					0.00	-0.14**
					[0.001]	[0.056]
Observations					97	97
R-squared					0.302	0.481
F-stat					41.67	5.964

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. For weak instruments, LC_2sls 95 percent confidence interval is computed using the Stata command twostepweakiv *** p<0.01, ** p<0.05, * p<0.1.

Table 4. Labor productivity and the legacy of serfdom

	(1)	(2)	(3)	(4)	(5)
VARIABLES	1897 labor productivity (log)				
Sea dummy	0.42***	0.41***	0.31***	0.25***	0.30***
	[0.107]	[0.106]	[0.096]	[0.067]	[0.095]
Share of employment in mining	14.06***	19.92***	10.54**	13.59***	13.28***
	[4.769]	[4.769]	[4.654]	[3.785]	[4.472]
Share of serfs in 1858	-0.20	-0.63***	-0.27**	-0.17**	-0.28**
	[0.157]	[0.149]	[0.105]	[0.082]	[0.123]
Log distance to Moscow	-0.23***	-0.26***	-0.21***	-0.11***	-0.11**
	[0.036]	[0.036]	[0.029]	[0.026]	[0.045]
Share of large estates in 1858	-0.21**				
	[0.098]				
Share of land with state tenure		-0.69***			
		[0.199]			
Share of land with commune tenure		-1.06***			
		[0.324]			
Share of in-migrants			2.52***		
			[0.857]		
Share of out-migrants			-0.84		
			[0.585]		
Share of urban populatin				1.58***	
				[0.207]	
Literacy					0.88***
					[0.290]
Constant	6.60***	7.34***	6.32***	5.42***	5.50***
	[0.264]	[0.361]	[0.244]	[0.198]	[0.365]
Observations	97	60	89	97	89
R-squared	0.496	0.671	0.571	0.684	0.554

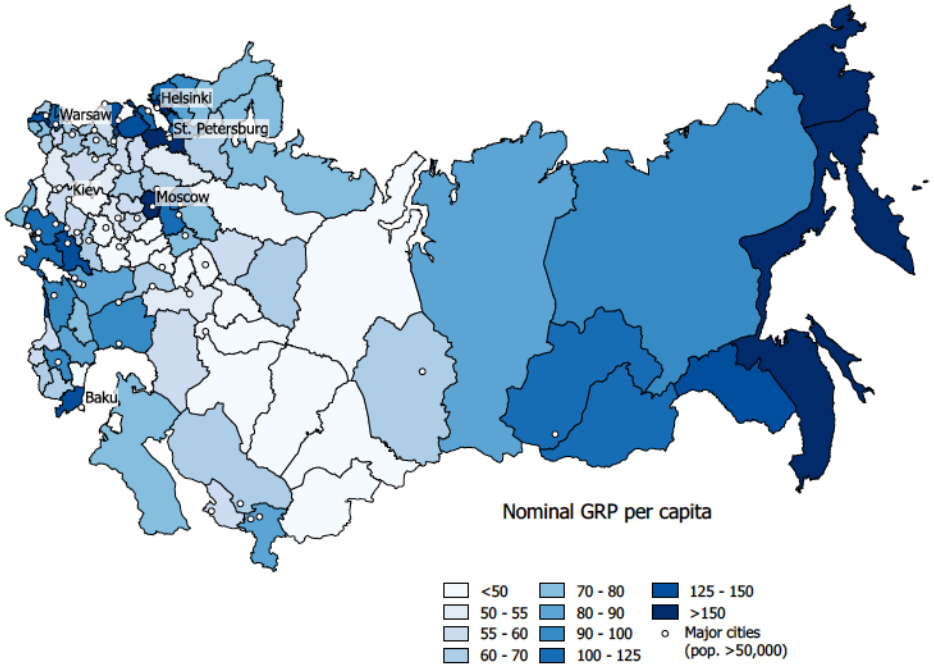
Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. Labor productivity and the new economic geography

	(1)	(2)	(3)	(4)	(5)	(6)
	1897 labor productivity (log)					
VARIABLES	OLS	IV, 2nd stage	OLS	OLS	OLS	OLS
Sea dummy	0.39***	0.39***	0.42***	0.25***	0.31***	0.37***
	[0.103]	[0.099]	[0.100]	[0.066]	[0.074]	[0.094]
Share of employment in mining	14.80***	14.83***	15.33***	15.17***	5.07*	13.11***
	[4.377]	[4.243]	[3.953]	[3.625]	[2.990]	[4.206]
Share of serfs in 1858	-0.48***	-0.48***	-0.48***	-0.24***	-0.12	-0.45***
	[0.122]	[0.118]	[0.128]	[0.092]	[0.118]	[0.124]
Log distance to Moscow	-0.20***	-0.20***	-0.17***	-0.12***	-0.04	-0.17***
	[0.047]	[0.046]	[0.050]	[0.022]	[0.039]	[0.037]
Market potentilal (log)	0.21**	0.21*	-0.03	-0.08		
	[0.095]	[0.112]	[0.149]	[0.104]		
Density (log)			0.08**			
			[0.040]			
Urban density (log)				0.25***		
				[0.028]		
Rural density (log)				-0.23***		
				[0.045]		
Land suitability for cereals (log)				0.02**		
				[0.010]		
Specialization index for agriculture					0.64*	
					[0.362]	
Specialization index for industry					1.10***	
					[0.209]	
Specialization index for services					-0.42**	
					[0.167]	
Mean industrial firm revenue (log)						0.15***
						[0.046]
Constant	4.11***	4.06***	6.17***	6.95***	4.85***	4.44***
	[1.129]	[1.317]	[1.543]	[1.048]	[0.284]	[0.612]
Observations	97	97	97	97	97	79
R-squared	0.491	0.491	0.515	0.712	0.672	0.564
Panel B: First stages of 2SLS regressions						
Dependent var:		Market potentilal (log)				
Sums of reverse distances		0.88***				
		[0.063]				
Sea dummy		-0.01				
		[0.032]				
Share of employment in mining		1.79				
		[2.201]				
Share of serfs in 1858		-0.02				
		[0.086]				
Log distance to Moscow		0.10***				
		[0.037]				
Observations		97				
R-squared		0.836				
F-stat		195.6				

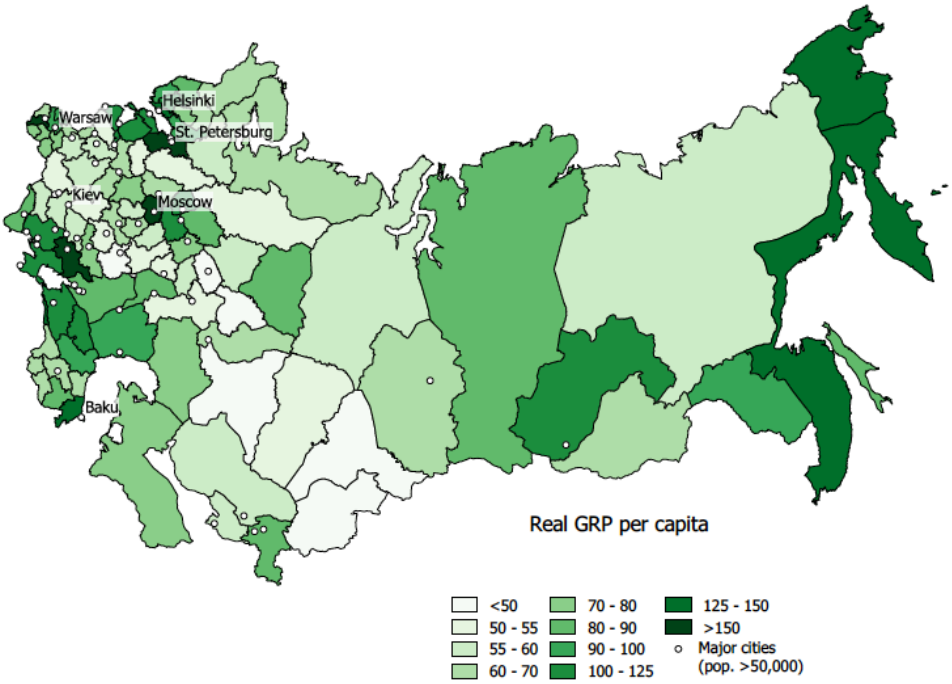
Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1. Nominal gross regional product per capita in the provinces of the Russian empire in 1897 (1897 rubles)



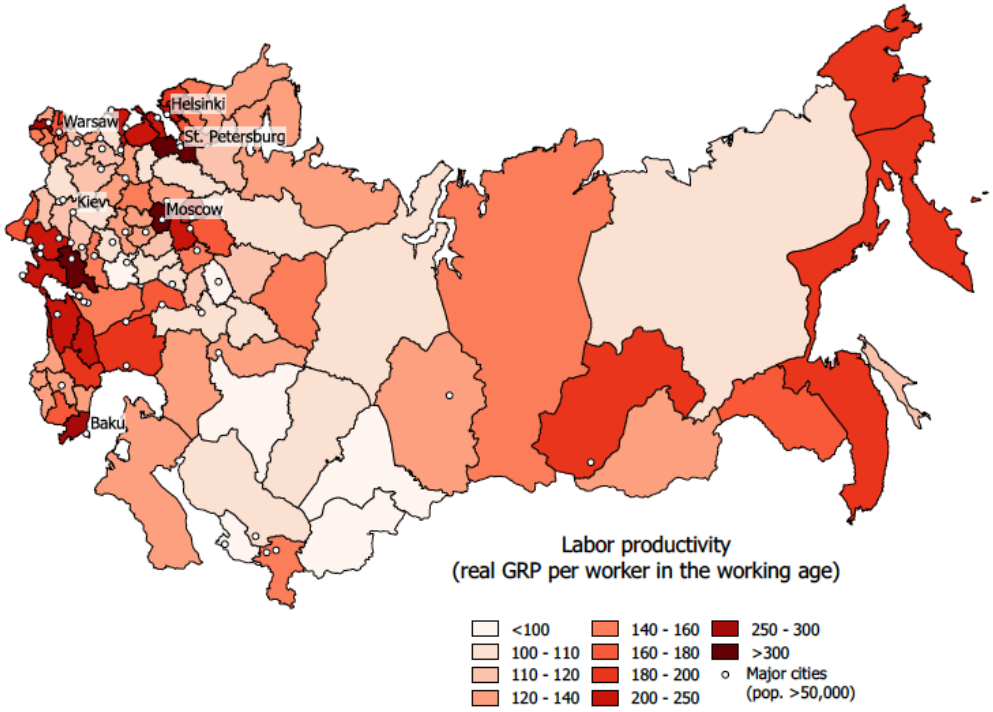
Source: the current study.

Figure 2. Real gross regional product per capita in the provinces of the Russian empire in 1897 (1897 rubles)



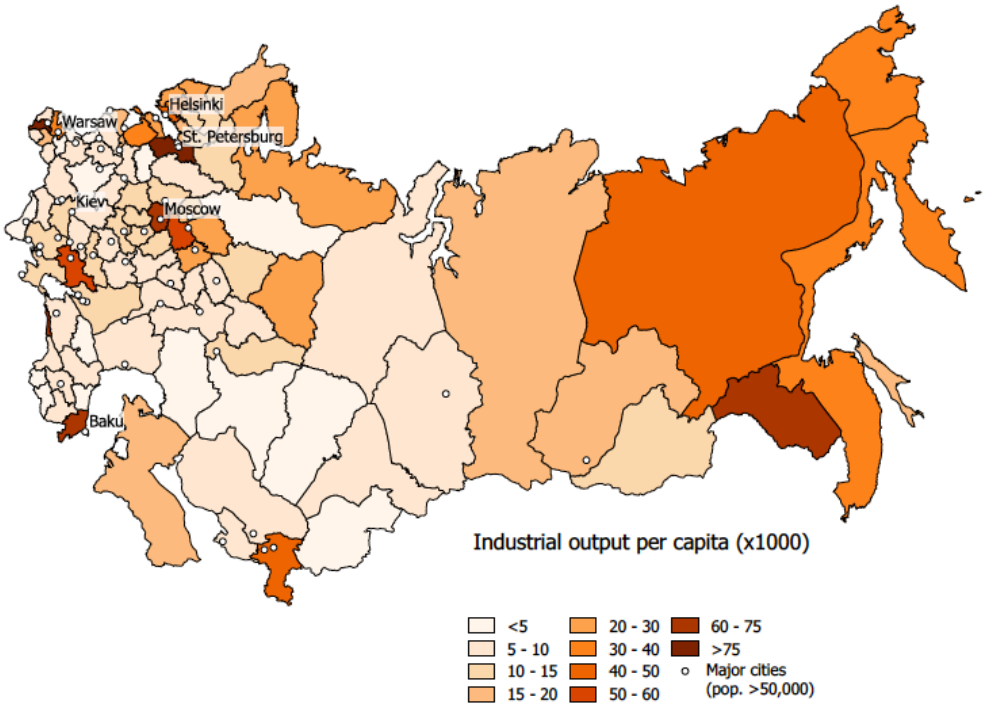
Source: the current study.

Figure 3. Labor productivity in the provinces of the Russian empire in 1897 (1897 rubles)



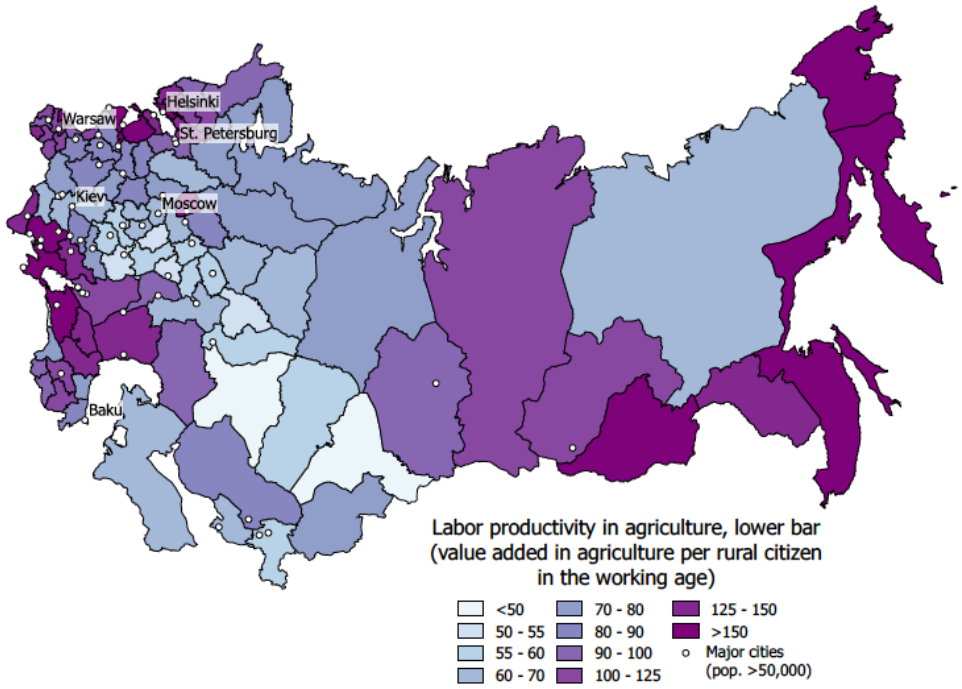
Source: the current study.

Figure 4. Industrial gross value added per capita in the provinces of the Russian empire in 1897 (1897 rubles)



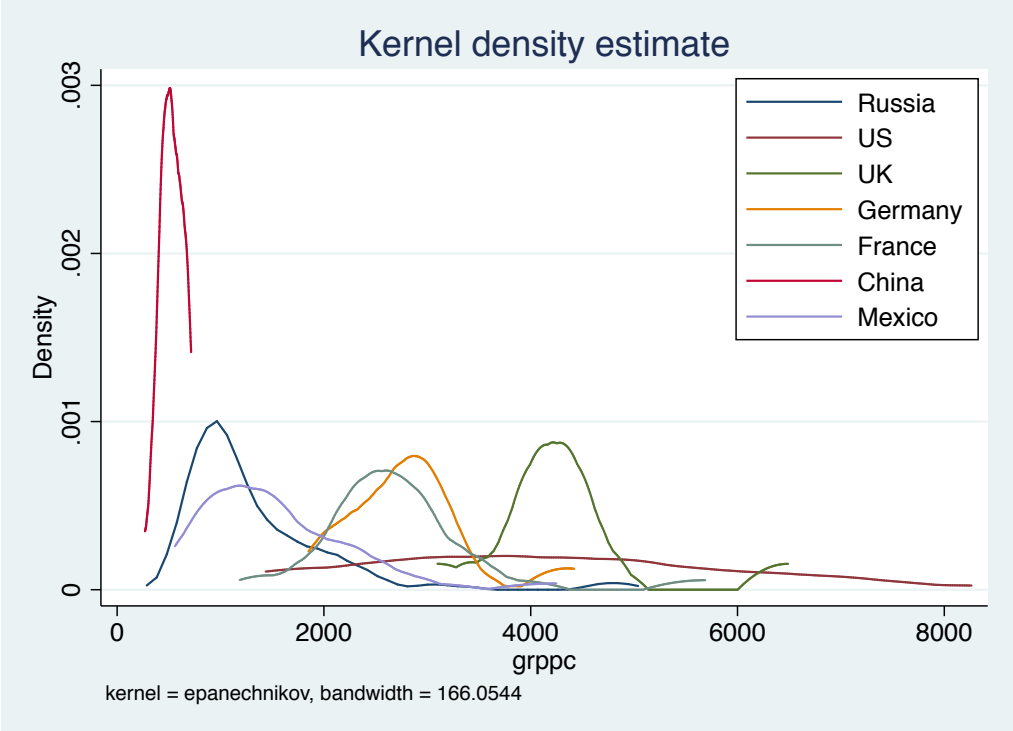
Source: the current study.

Figure 5. Labor productivity in agriculture in the provinces of the Russian empire in 1897 (1897 rubles)



Source: the current study.
Note: Low bound of estimates of labor productivity in agriculture are presented.

Figure 6. GRPs per capita distributions in selected countries around 1900 (1990 USD)



Source: same as for table 2.
Note: Calculated as kernel density approximation.

Supplementary Online Appendix for *A Regional Perspective on the Economic Development of the late Russian Empire*
[Not intended for publication in the journal]

AI. 1897 nominal GRPs reconstruction methodology

A. Industry

1. *Large industry: Cotton, wool, linen products; Food; Leather; Woodworking; Paper; Chemicals; Ceramic; Mining and quarrying; Machine-building and metal-working; Other industrial products and Products manufactured from materials received from customers*
2. *Small industry*
3. *Construction*
4. *Public Infrastructure*

B. Agriculture

1. *Cereals and other crops: grain*; *potatoes and beetroot; industrial crops; beans and lentils*
2. *Husbandry: Meat production; Milk production; Leather production; Horse breeding; Pig farming; Sheep and goat farming*
3. *Revenues from sales of hay and straw to urban areas*
4. *Sunflower, mustard, hop and rapeseed*
5. *Beekeeping and gardening*
6. *Fruits and grape*
7. *Fishing*
8. *Hunting*
9. *Forestry*

C. Services

1. *Railroads*
2. *Water transport*
3. *Trucking*
4. *Communications: imperial post and telegraph, local post, telephone*
5. *Trade*
6. *Banking*
7. *Public goods: public education, health, national defense, public administration, public infrastructure*
8. *Housing: urban and rural*
9. *Servants*

AII. 1897 real GRPs reconstruction methodology

AIII. 1897 regional labor productivity reconstruction methodology

AIV. Robustness of GRPs estimates

AV. Data description and sources for variables used in the regression analysis

Tables and Figures

For appendix tables A1-A65 and VA0-VA10 consult separate .xls file
Figures B1-B2.
Tables B1-B17.

AI. 1897 nominal GRPs reconstruction methodology

I reconstruct nominal gross regional product for all provinces of the Russian empire in 1897 from the production side. To be precise, I reconstruct value added in 14 subsectors in industry, in 18 subsectors in agriculture, and in 12 subsectors in services. Table A0 in the online appendix lists all subsectors and outlines the reconstruction methodology for each of them. Table VA0 presents the reconstruction results by subsector in levels and per capita, nominal gross regional product and nominal GRP per capita (as well as real GRP per capita and labor productivity – see sections below on corresponding reconstruction methodologies). Tables A1 and A2 report province-level 1897 census figures on population and occupations. The bulk of the figures used for estimates are originally from official statistical publications, recently collected for the 'Electronic repository of Russian historical statistics, 18th – 21st centuries', available at http://ristat.org (Kessler and Markevich 2019, forthcoming). In my reconstructions of value added, I follow the methodology originally elaborated by Sergei Prokopovich (1918), who pioneered estimates of output of the Russian economy calculating national income of fifty European provinces of the Russian empire around 1900, with later adjustments by various authors as discussed in Markevich and Harrison (2011). Prokopovich (1918) and Markevich and Harrison (2011) operate at the national level only. In this work, I extend their approach to the province-level and reconstruct gross regional products for each and every province of the Russian empire in 1897. Where possible, I rely on subsector-specific methodologies (Yasnopolskij 1891, 1897; Silantiev 1898; Commission on peasant welfare 1901; Groman 1912; Strumilin 1924; Central statistical agency 1924; Vainshtein 1960; Gregory 1982; Salomatina 2014). For some of the within sub-sector estimates, I use regional employment shares by occupation, i.e. occupation relative rather than absolute values, which allows me to avoid the potential problem of under-enumeration of employees in the 1897 census (discussed in the main part of this paper), which is relevant for absolute figures and between-subsectors relative figures (but likely not relevant for within sub-sector relative figures).[1] Below I provide

[1] In particular, the under-registration problem was most pronounced in agriculture relative to other sectors (Gozulov 1936). There is no evidence that (under)registration practices differed between provinces within sectors or sub-sectors. A comparison of

full details on the implemented reconstruction methodology and historical sources.

A. Industry

1. Large industry: Cotton, wool, linen products; Food; Leather; Woodworking; Paper; Chemicals; Ceramic; Mining and quarrying; machine-building and metal-working – MBMW; Other industrial products and Products manufactured from materials received from customers

Because of available data, I follow the late 19th century Russian statistical tradition and distinguish so-called large industry (factories with more than fifteen workers or with an engine) from small industry (also called cottage industry). To reconstruct value added in large industry, I use monetary values of manufacturing goods from Blau (1900) that is an official volume on industrial production in 1897. The volume reports province-level data on output of particular manufacturing products including production of industrial plants subject to pay excise tax as well as on output in mining. Table A3 in the online appendix reproduces statistics on the monetary values of more than 500 individual products for all provinces of the Russian empire except Finland. To correct for double counting of intermediate goods, I aggregate individual product figures by industrial subsectors and adjust the obtained sums using industrial subsector specific ratios of value added to gross industrial production in 1913 from the 1918 industrial census (Central statistical agency 1924), reported in table A5. The 1918 census collected retrospective data on industrial output and inputs for the years before, during and after the First World War. An alternative approach would be to use 1900 ratios of value added to gross industrial production from Varzar (1903), which is closer to the 1897 benchmark year. However, 1900 figures are of problematic quality. As Varzar (1903) himself noted, a number of industrial firms reported actual costs of production rather than market value of produced

---

geographical variation in labor productivity in agriculture measured in different ways illustrates that. Figure 5 in the main text presents an upper bar of estimates of labor productivity in agriculture, when the number of rural citizens in the working age stands for the denominator. Figure B2 presents a lower bar of labor productivity in agriculture, when the number of registered workers stands for the denominator. The revealed spatial patterns are mostly the same. The correlation coefficient between the two measures is 0.94 (significant at one percent level).

Subsectors, for which I use occupational figures, contributed about 10 percent of GDP.

goods. In the robustness section, AIV of this appendix, I discuss how sensitive my results are to this choice. I present estimates with correction ratios from Varzar (1903) in table A65.

Because of available data, I distinguish ten branches of industry following the late 19th century Russian official classification of branches of industry. In addition to ten branches, I distinguish "products manufactured from materials received from customers", for which the original volume reports payments to factories for their works implemented for customers, i.e. value added. I do not apply any adjustments to this category because the cost of "materials received from customers", which were the major intermediate inputs for these products, is already deducted from the cost of these products.

Blau (1900) does not cover Finnish provinces. To reconstruct value added in industry in Finland, I use official statistical volumes by the Central Statistical Bureau in Helsinki (1900, 1901). I report gross industrial output in Finland from these volumes in table A4. To switch from gross industrial output to value added, I correct the original figures for the double counting of intermediate goods in a similar way as discussed above. I apply 2.666814778074 exchange rate for all figures reported in Finnish marks here and everywhere below (during the gold standard period, 1 mark equaled to 0.290322 and 1 ruble equaled to 0.774235 grams of gold).

Table VA1 in the online appendix presents the results of my reconstruction of value added in large industry.

2. Small industry

There are no comprehensive statistics on output of small (cottage) industry in the Russian empire. I use the materials of the 1901 Commission on peasant welfare (1903) to approximate value added in this sector. In particular, the commission collected data on the number of workers in craft industry and their total income in 50 European provinces of the empire (table A6). Because labor was almost the only input factor in cottage industry, I proxy value added in this subsector by total income of craftsmen. I apply this assumption to the 50 European provinces for which I have data on the income of craftsmen disaggregated by province.

For the other provinces (with an exception of Finland), I assume that income (and value added) per craftsman registered in the 1897 population census (Trojnitskij 1905)

was the same as in the 50 European provinces of the empire. Accordingly, I estimate value added in small industry by multiplying the number of registered craftsmen in a corresponding province by their average income in the 50 European provinces. I get average income per registered craftsmen in the 50 European provinces by dividing total income in the craft industry from table A6 by the number of registered craftsmen in these provinces from table A7 (craftsmen in provinces in regions 1 to 13). Note that there were less craftsmen registered by the census than by the 1901 commission because the commission took into account primary and non-primary occupations in crafts.

Unfortunately, 1897 census volumes report registered craftsmen only by groups of provinces, i.e. regions, and not by individual provinces (table A7). In table A8, for each province outside the 50 European provinces, I report the number of craftsmen in the region to which the province belonged and total incomes of craftsmen in this region. I assume that large industry and small industry complimented rather than competed with each other (the coefficient of correlation between the two in 50 European provinces is 0.4 significant at the one-percent level); I use provincial shares of value added in large industry (table VA1) as weights to distribute total income of craftsmen in a region between provinces in this region (table VA2).

I cannot apply the procedure described above to Finland because the 1897 census did not cover the Great Duchy of Finland. Instead, I approximate value added in craft industry in the Finnish provinces by multiplying the Finnish rural population by value added in craft industry per rural citizen in Estlyandiya province. I chose Estlyandiya because this province was a neighbor of Finland (1), close in terms of language and culture (2) and without the institution of the commune (3).

3. Construction

There are few data on output in this subsector in the late Russian empire. I employ 1897 census data on occupations in construction. I follow Prokopovich (1918) assuming that on average, at the national level, value added per worker in large industry and construction was the same. I estimate national average value added per worker in large industry by dividing total value added in large industry in the empire (without Finland) from table VA1 by the total number of employees in industry registered by the 1897 census from table A2. To account for regional variation in productivity in construction, I

use the earliest available figures on regional variation in wages in this sector from Groman (1912) reported in table A9. For each province, I report the ratio of local wages in construction to the national average (table A10). Finally, to get value added in construction, I multiply the number of employees in construction registered in the 1897 census by value added in larger industry by registered workers in industry and by ratio of local wages in construction to the national average wage in construction.

To reconstruct value added in construction in Finland, I use official figures on output in this sector from the Central Statistical Bureau in Helsinki (1900, 1901) as reported in table A4.

Table VA2 in the online appendix presents the results of my reconstruction of value added in small industry and construction.

4. Public Infrastructure

I reconstruct value added in public infrastructure from budget expenditures (table VA7). I describe the methodology for all types of budget expenditures altogether below.

B. Agriculture

1. Cereals and other crops: grain (wheat, rye, oats, barley, buckwheat, corn, millet and other cereals, rice), *potatoes and beetroot, industrial crops* (tobacco, flax, hemps, cotton, sesam and other industrial crops)*, beans and lentils*

I follow Prokopovich (1918) in estimating value added in production of crops. I multiply crop output in a province by local prices of the corresponding crop and by various adjustment factors. Figures on output of crops reported in tables A11 are from the Central statistical agency (1898b and 1902), supplemented by statistics from governors' reports. Governors' reports often measured crops output in *chetverti* rather than in *puds,* i.e. in measures of capacity rather than weight. I use province-specific conversion factors for grain from Mironov (1985) presented in table A13 to account for geographical variation in crop density and to switch from *chetverti* to *puds*. Output of crops in Finland (table A12) is from the Central Statistical Bureau (1900).

Prices (table A14) are from the Ministry of Agriculture and State Property (1897-1898). Historians recognize these statistics as the most reliable source on local prices in the late Russian empire (Mironov 1985). In addition, I use prices from annual governor's reports. I approximate missing prices with regional averages (e.g. I use the average price

for winter wheat in the Central industrial region in 1897 to approximate the price for winter wheat in Tver province). If prices for a particular crop are unknown for the whole region, I approximate them with the average price of this crop in the neighboring region. Finnish prices for four major cereals (wheat, rye, barley and oats) are from Central Statistical Bureau (1900). For the other crops in the case of Finland, I use prices in neighbor provinces.

There is a long-lasting debate on potential underestimation of production of grain and other crops by Russian imperial statisticians (e.g. Ivantsov 1915 among many others). There is no clear conclusion as to the scale of the necessary adjustment or to its variation over space. I use 1.19 adjustment coefficient for grain and 1.283 for potatoes (table A16), following Markevich and Harrison (2011), to get estimates consistent with the previous literature. These adjustments are an upper bound of discussed corrections. I apply these corrections to all provinces of the empire with an exception of the provinces in the Great Duchy of Finland which had a separated statistical agency and followed different practices of data collection than the rest of the empire. In the robustness section, AIV of this appendix, I discuss how robust my estimates are in regard to this assumption. I present estimates with no grain adjustment in table A65.

Table VA3 in the online appendix presents the results of my reconstruction of value added in the production of grains, potatoes and beetroot, industrial, beans and lentils.

2. Husbandry: Meat production, Milk production, Leather production, Horse breeding, Pig farming, Sheep and goat farming

I follow Prokopovich (1918) and reconstruct value added in husbandry from province-level livestock statistics. Table A15 reports data on horses and other draft animals (camels, buffalos, mules, donkeys) cattle, pigs, sheep, goats and deer extracted from the Ministry of Internal Affairs (1900), the Central Statistical Bureau (1900) and annual governors' reports. I use the Ministry of Agriculture and State Property (1897-1898) to account for variation in local prices for horses, cattle, sheep and meat production (reported in table A14). Because livestock accounting in the Russian empire suffered from the under-registration problem, I follow Vainshtein (1960) and Markevich and Harrison (2011) to adjust original figures using correction coefficients as reported in table

A16. I do not apply these corrections to the livestock in the Great Duchy of Finland (Finland had a separated statistical agency and followed different practices of data collection than the rest of the empire). To estimate husbandry output in Finland, I use prices in neighbor provinces.

Table A17 summarizes Prokopovich (1918) coefficients, which he used for his estimates of income from husbandry in fifty European provinces of the Russian empire around 1900. I expand his approach to the whole empire. In particular, I assume that each seventh cow was slaughtered for meat annually, and the weight of the average carcass was 10.54 *pud*s. I use local beef prices from the Ministry of Agriculture and State Property (1897-1898) reported in table A14 to estimate monetary value of the produced beef. I employ Prokopovich (1918) coefficient of veal to beef ratio (0.23) to account for the production of veal.

Again, following Prokopovich (1918), I assume that the share of cash cows was 42 per cent. One cash cow produced on average 60 buckets, and one bucket of milk cost 0.41 roubles. I use relative (to the average) local prices of cows from the Ministry of Agriculture and State Property (1897-1898) reported in table A14 to account for potential regional variation in productivity of cash cows and milk prices.

Cows and horses also contributed to leather production. On average, annual revenue from skin procurement was 0.82 rubles per head of cattle or horse. To account for potential local variation in skin prices, I use relative (to the average) local prices of cows and horses from the Ministry of Agriculture and State Property (1897-1898) reported in table A14.

I follow Prokpovich (1918) to estimate income from breeding pigs. I assume that a half of all pigs were for slaughter, and the average weight of a pig was six *pud*s where three *pud*s were lard. I use local pork and lard prices from the Ministry of Agriculture and State Property (1897-1898) reported in table A14 to estimate monetary income from breeding pigs. The production of lard required barley; seven *pud*s of barley were needed on average for production of one *pud* of lard. I use local barley prices from the Ministry of Agriculture and State Property (1897-1898) reported in table A14 to estimate the monetary costs of the required barley inputs. In addition to lard and pork, each pig produced about 0.03125 *pud*s of stubble, which cost 41.6 rubles per *pud*.

According to Prokopovich (1918), three quarters of all sheep produced wool (the rest were lambs). Wool output per coarse-wooly sheep was 0.125 *pud*s per year and per thin-wooly sheep 0.2 *pud*s per year. I use local wool prices from the Ministry of Agriculture and State Property (1897-1898) reported in table A14 to estimate monetary income from wool production. In addition, 0.307 of all sheep were for slaughter; the average sheep gave about 1.13 *pud*s of lamb and 0.12 of lard. I use local lamb and lard prices from the Ministry of Agriculture and State Property (1897-1898) reported in table A14 to estimate monetary income from lamb and lard production. The sheep skin cost 1.83 rubles. Goats were like coarse-wooly sheep in terms of generating revenues.

Horses and other draft animals were important intermediates for production of crops, but they could generate direct revenues in the case of sales outside of agriculture. I follow Prokopovich (1918) and include this subsector into estimations while its output might be partially used as intermediate goods in other sectors, like trucking, for which I account in my GRPs estimates. In the robustness section, AIV of this appendix, I discuss how sensitive my results are to this choice. I present estimates without value added by this subsector in table A65. According to Prokopovich (1918), about 0.56 percent of all horses were sold to urban dwellers annually. I take horse prices from the Ministry of Agriculture and State Property (1897-1898) reported in table A14. I assume that camels, buffalo, mules, and donkeys were like horses in terms of their draft power and potential revenues.

Following Prokopovich (1918), I assume that value added in poultry was equal to zero because expenditures on fodder were similar to revenues.

With an exception of expenditures on barley needed for pigs to accumulate lard, I do not account for other expenditures on fodder for livestock. As Prokopovich (1918) noted, such an approach generates an upper bound of value added in husbandry and agriculture in general. There is no reliable information on expenditures on livestock fodder either at the province level or at the national level in the late 19th century. Nefedov (2011) cites Vilson's estimations for the mid 19th century and 1917 estimations by the Ministry of Food Provision and concludes that consumption of cereals by livestock substantially increased between the middle of the 19th century and 1917 due to transformation of meadows into arable lands. However, both the speed and geography of

this transformation is poorly documented. I use 1917 figures on consumption of cereals by cattle and other livestock (Table A18) and local oat prices from the Ministry of Agriculture and State Property (1897-1898) reported in table A14 to estimate an upper bound of expenditures on consumption of cereals. Deducting these expenditures from value added in husbandry, one could get a lower bound estimate of income generated in this sector. In the robustness section, AIV of this appendix, I discuss how sensitive my results are to the lack of such deduction. I present estimates with such deduction in table A65.

Table VA3 in the online appendix presents the results of my reconstruction of value added in meat production, milk production, leather production, horse breeding, pig farming, sheep and goats farming.

3. Revenues from sales of hay and straw to urban areas

Hay and straw were intermediate goods for husbandry unless they were sold to urban areas. I follow Prokopovich (1918) and include this subsector into estimations while its output might be partially used as intermediate goods in other sectors, like trucking, for which I account in my GRPs estimates. In the robustness section, AIV of this appendix, I discuss how sensitive my results are to this choice. I present estimates without value added by this subsector in table A65. Prokopovich (1918) estimated total revenues from sales of hay and straw in the 50 European provinces of the Russian empire as 24,476 thousand rubles (table A19). I assume that these sales were the same in 1897 and allocate the revenue between the 50 provinces using the number of urban citizens in each province (known from table A1) as weights. In addition, I assume that sales per urban citizen were the same in the rest of the empire and proxy total sales from the province-specific urban population. To account for the local variation in hay prices, I multiply the obtained figure for each single province by the local hay price relative to the average hay price in the empire. Hay prices are from the Ministry of Agriculture and State Property (1897-1898) reported in table A14.

Table VA3 in the online appendix presents the results of my reconstructions of revenues from sales of hay and straw to urban areas.

4. Sunflower, mustard, hop and rapeseed.

There is no detailed province-level information about the production of sunflower,

mustard, hop and rapeseed in the Russian empire. Because of that I take Prokopovich's (1918) approximations for the 50 European provinces for the year 1900 (table A19) and assume that the output was the same in 1897. Next, I allocate the output between the 50 provinces using the number of registered employees in the related subsector of agriculture in each province (known from table A2) as a weight; doing that, I introduce an additional requirement of province suitability for sunflowers. In addition, I assume that labor productivity was the same in the rest of the empire and proxy output from province-specific occupational data. To account for local variation in sunflower prices, I multiply the obtained figures by relative (to the average) local sunflower prices. Sunflower prices are from the Ministry of Agriculture and State Property (1897-1898) reported in table A14.

I assume that production of sunflowers, mustard, hop and rapeseed was negligible in the Finnish provinces because of local climate.

Table VA3 in the online appendix presents the results of my reconstruction of value added in the production of sunflower, mustard, hop and rapeseed.

5. Beekeeping and vegetable gardening

There is no detailed province-level information about the production of beekeeping and vegetable gardening in the Russian empire. Because of that I take Prokopovich's (1918) approximations for the 50 European provinces for the year 1900 (table A19) and assume that the output was the same in 1897. Next, I allocate the output between the 50 provinces using the number of registered employees in the related subsector of agriculture in each province (known from table A2) as a weight. In addition, I assume that labor productivity was the same in the rest of the empire and proxy output from province-specific occupational data.

I cannot apply the procedure described above to Finland because the 1897 census did not cover the Great Duchy of Finland. Instead, I approximate value added in beekeeping and gardening in the Finnish provinces by multiplication of value added in beekeeping and gardening per rural citizen in the empire by the rural population in the Finnish provinces.

Table VA3 in the online appendix presents the results of my reconstruction of value added in the production of gardening and beekeeping.

6. Fruits and grape

Table A20 presents data on lands in gardens and vineyards as well as revenues per *desyatina* (equal to 1.09 hectare) and total revenues from annual provincial governors' reports. Data are missed for many provinces. I use the 1897 census data (table 2) to estimate revenues per registered employee in fruit and grape cultivation in the provinces, for which data on revenues are available. I use these figures and 1897 occupation data to proxy revenues from fruit and grape production in the other provinces of the empire. I approximate missing provincial revenues per registered employee with their regional averages (e.g. I use average revenues per registered employee in the West Caucasus region to approximate revenues per registered employee in Chernomorks province). If provincial revenues per registered employee are unknown for all provinces in the region, I approximate them with the average revenue per registered employee in the neighboring region. I assume that labor was the only input in this subsector and approximates value added by revenues.

I cannot apply the procedure described above to the Finnish provinces because the 1897 census did not cover the Great Duchy of Finland. I assume that production of fruit and grape was negligible in Finland because of the local climate.

Table VA3 in the online appendix presents the results of my reconstruction of value added in fruit and grape cultivation.

7. Fishing

The Ministry of Finance (1902) and annual governors reports provide data on output in this subsector for a number of provinces (table A21). There are many provinces with missing values. I use the 1897 census data (table 2) to estimate average output per registered employee in fishing for the provinces where data on output are available. I approximate missing provincial outputs per registered employee with their regional averages (e.g. I use average revenues per registered employee in the North region to approximate revenues per registered employee in Vologda province). If provincial outputs per registered employee are unknown for all provinces in the region, I approximate them with the average output per registered employee in the neighboring region. I use these figures and 1897 occupation data to proxy output in fishing in other provinces of the empire.

I cannot apply the procedure described above to the Finnish provinces because the 1897 census did not cover the Great Duchy of Finland. Instead, I use the estimate of value added in fishing and hunting in Finland in 1897 by Hjerpee (1989) and use rural population figures as weights to allocate this value added between Finnish provinces. Using this procedure, I cannot distinguish value added in fishing from value added in hunting in Finnish provinces.

Table VA4 in the online appendix presents the results of my reconstruction of value added in fishing.

8. Hunting

Table A22 provides data on the number of hunters, earnings per hunter and their total earnings from Silantiev (1898) and annual governors' reports for various areas of the Russian empire. Data are fragmented and are not available for many provinces. I use 1897 census data (table 2) to estimate average earnings per registered employee in hunting in the provinces, for which data on output are available. I use this figure and 1897 occupation data to proxy earnings in hunting in the other provinces of the empire. I assume that intermediate inputs in hunting were negligible. I use earnings in hunting as a proxy for value added.

I cannot apply the procedure described above to the Finnish provinces because the 1897 census did not cover the Great Duchy of Finland. Instead, I use the estimate of value added in fishing and hunting in Finland in 1897 by Hjerpee (1989) and use rural population figures as weights to allocate this value added between Finnish provinces. Using this procedure, I cannot distinguish value added in fishing from value added in hunting in Finnish provinces.

Table VA4 in the online appendix presents the results of my reconstruction of value added in hunting.

9. Forestry

Table A23 presents data on areas in state and non-state forests in the empire form the Central Statistical Bureau (1900), the Ministry of Finance (1900a) and annual governors' reports. The Forest Department (1898) reported revenues from state forests by province (table A24). Given the extensive nature of the sector, I assume that profitability of state and non-state forests per area unit was the same in each province and approximate

revenues from non-state forests from their territory. Summing-up revenues from state and non-state forests, I get total revenues in forestry. I assume that intermediates inputs in forestry were negligible, and use revenues in forestry as a proxy for value added.

For Primorie province, where forest area is not known, I approximate value added in forestry by multiplication of the number of registered employees in forestry in the province in 1897 (from table A2) by the average output per registered employee in forestry.

For Finland, I use the estimate of value added in forestry by Hjerppe (1989) and use forestry output in Finnish provinces in kind from Central Statistical Bureau (1900) as weights to allocate the value added between Finnish provinces.

Table VA4 in the online appendix presents the results of my reconstructions of value added in forestry.

C. Services

1. Railroads

Table A26 reports railways revenues and expenditures by railway line in 1897 from the Ministry of Transportation (1899); there are no similar province-level figures. The only province level-figures on the length of railways are in table A25 (Ministry of Transportation 1899). Accordingly, I estimate value added in the sector of railways first at the national level (with an exception of Finland). For that, I sum up expenditures on salaries and wages and profits (profits are defined as total revenues minus total expenditures) from table A26. I could use either railways lengths or numbers of registered employees in this subsector as weights to get province-level figures. I prefer occupation data, assuming that value added should be closer linked with labor inputs rather than with the length of railways. I report the alternative estimates as well. The correlation between the two variables is 0.82 (significant at the one-percent level).

For the Finnish provinces, I estimate value added by Finnish railways (estimated as a sum of expenditures on labor and profits) and use total revenues generated by railway stations from particular provinces from table A27 as weights (I assign each station to a particular province and aggregate station revenues at the province-level).

Table VA5 in the online appendix presents the results of my reconstruction of value added in railways.

2. Water transport

There is little data on water transport either at the national or province level. The Ministry of Transportation (1898) reported total weight of goods transported by water in 1897 in the European part of the empire (without Finland). The earliest figure for the Asian part of the Empire is for 1904 (Central Statistical Agency 1906a). I approximate the 1897 value of goods transported by water in the Asian part of the empire by applying the average annual growth rate of the amount of goods transported by water in the Asian part of the empire in the period between 1904 and 1907, i.e. before mass migration to Siberia triggered by the Stolypin reform. I assume that growth rates were similar between the two periods: 1897 – 1904 and 1904 – 1907. Summing up the values of goods transported by water in the European and Asian parts of the empire, I get an approximation of the amount of goods transported by water in the whole empire (without Finland) in 1897 (table A28).

To approximate value added in water transport, I assume that value added per transported unit of weight was the same in railways and water transport, i.e. 0.0397893 ruble per *pud*. I estimate this coefficient by regressing value added in railways on the amount of transported goods by railways from table A26 (the constant term is suppressed). I do not include the measure on transported passengers into this regression because there are no data on passengers transported by water even for the European part of the empire in 1897. Note also that transportation of goods generated about three-quarters of all railroad revenue (Ministry of Transportation 1899).

I use the number of registered employees in water transport in a province from the 1897 imperial census (table A2) as a weight to distribute the estimated value added in waterways in the European and Asian parts of the empire (without Finland). For the Finnish provinces, I use the ratios of value added in waterways to value added in railways in Estlyandia province to approximate value added in waterways. I chose Estlyandia as it is a neighbor province (1) and similar to Finland's level of development (2).

Table VA5 in the online appendix presents the results of my reconstruction of value added in water transport.

3. Trucking

There are almost no data on non-railway and non-water transport activities for the late

Russian empire. I estimate value added in trucking by multiplying the number of registered employees in this sector from the imperial 1897 census (table A2) by the average annual income of a cabman (120 rubles).

I cannot apply the procedure described above to the Finnish provinces because the 1897 census did not cover the Great Duchy of Finland. For Finnish provinces, I use the ratio of value added in trucking to value added in railways in Estlyandia province to approximate value added in trucking. I chose Estlyandia as it is a neighbor province (1) and similar to Finland's level of development (2).

Table VA5 in the online appendix presents the results of my reconstructions of value added in trucking.

4. Communications: imperial post and telegraph, local post, telephone

For the imperial post and telegraph, I apply similar methodology as for the railroads. Table A29 reports total revenues, total expenditures and expenditures on wages and salaries by imperial post and telegraph (Department of Communication 1898). I sum up profits (total revenues minus total expenditures) with expenditures on labor to approximate value added in this sector at the national level (without Finland). I use data on registered employees in communication from the 1897 imperial census (table A2) as weights to allocate the national-level estimate between provinces.

For Finnish provinces, I use province-level figures on total revenues and expenditures on post and telegraph to estimate profits (Central Statistical Bureau, 1899). I approximate expenditures on wages and salaries in the Finnish provinces, multiplying the number of employees in post and telegraph in the Finnish provinces by expenditures on wages and salaries per employee in the rest of the empire (table A30).

Table A31 presents data on expenditures on local posts from the Ministry of Finance (1900a) and Central Statistical Bureau (1899). I assume that local posts were fueled primary by labor and approximate value added by expenditures on labor.

Table A32 reports city-level data on total revenue, total expenditures and expenditures on wages and salaries by city telephone companies. I estimate value added for city companies as the sum of profits (total revenues minus total expenditures) plus expenditures on labor and aggregate them by province.

Table VA5 in the online appendix presents the results of my reconstruction of

value added in imperial post and telegraph, local post and telephone.

5. Trade

I estimate value added in trade as 12 per cent of trade turnover (Prokopovich 1918), and follow methodology proposed by Strumilin (1924) to estimate trade turnover (table A38).

I use Strumilin's (1924) classification of trade firms: firms subject to pay the trade percentage tax (a); guild firms subject to pay the trade spreading tax (b); non-guild firms subject to pay the trade spreading tax (c); 'old' guild firms released from paying the trade spreading tax (d); newly established guild firms released from paying the trade spreading tax (e); 'old' and newly established non-guild firms released from the paying trade spreading tax (f); firms in small-retail trade (g); taverns, bars, pubs, beverage sales etc. (h) and the state alcohol monopoly (i).

Table A33 repots data on firms subject to pay the trade percentage tax (a), the number and revenues of those firms that actually paid the trade percentage tax from Ministry of Finance (1900b). I assume that the revenue per firm of those firms that were subject to pay the trade percentage tax but did not actually pay it were similar to the revenues of firms that paid the trade percentage tax. I use this assumption to estimate the total revenue of firms of this type. According to Strumilin (1924), the revenue-to-turnover ratio was 0.052 for this type of firms. I use this figure to estimate the trade turnover of this type of firms (table A38).

Table A34 reports data on guild firms subject to pay the trade spreading tax (b) from the Ministry of Finance (1900b). The table contains information on the numbers of firms from the 1st and 2nd guilds that actually paid the tax as well as their total turnover. In addition, this table reports the number of firms from the 1st and 2nd guilds which got tax exemption privileges. I adjust total turnover to account for turnover of firms with exemption privileges, assuming that turnover per firm of these firms was similar to turnover per firm of those firms which actually paid the tax. The original source includes not only trading firms in the category "guild firms subject to pay trade spreading tax," but also firms in banking, transportation and infrastructure. I deduct turnover of these firms (adjusted in the same way as described above) from total turnover of guild firms subject to pay the trade spreading tax to get turnover of only trading firms.

Table A35 reports data on trade non-guild firms subject to pay the trade spreading

tax (c) from the Ministry of Finance (1900b). Like in the case of guild firms, the table contains information on the number of non-guild firms that actually paid the tax and their turnover, as well as information on the number of non-guild firms with tax exemption privileges. I adjust turnover to account for the turnover of non-guild firms with exemption privileges, assuming that turnover per firm of these non-guild firms was similar to turnover per firm of those non-guild firms which actually paid the tax. I also deduct the turnover of non-guild firms in alcohol retailing for which I account separately in table A36.

The Ministry of Finance (1900b) did not provide information on the number and turnover of 'old' guild firms released from paying the trade spreading tax (d); newly established guild firms released from paying the trade spreading tax (e); 'old' and newly established non-guild firms released from paying the trade spreading tax (f); and firms in small retail trading (g). I follow Strumilin's (1924) corrections to account for trade turnover of these firms. According to Strumilin (1924), turnover of firms in categories (d) as well as in category (e) was between 2 and 3 percent of the trade turnover of guild firms in category (b); I use 2.5 per cent as the average. Strumilin (1924) argues that turnover of firms in category (f) was between 6 and 9 percent of non-guild firms in category (c); I use 7.5 percent coefficient as the average. Finally, according to Strumilin (1924), trade turnover of small retailing firms (except alcohol retailing) was about 20 percent of the total trade turnover of guild and non-guild firms in categories (b) and (c). Table A38 presents the results of my corrections.

Table A36 reports data on non-guild firms in alcohol retailing from the Ministry of Finance (1900b). Like in the case of other non-guild firms, the table contains information on the number of firms that actually paid the tax and their turnover as well as information on the number of non-guild firms in alcohol retailing with tax exemption privileges. I adjust turnover to account for the turnover of non-guild firms in alcohol retailing with exemption privileges, assuming that turnover per firm of these non-guild firms in alcohol retailing was similar to turnover per firm of non-guild firms in alcohol retailing which actually paid the tax.

Finally, table A37 reports data on turnover in the state alcohol monopoly in those provinces where the monopoly was introduced by 1897 (Ministry of Finance 1900a).

Table A38 sums up the results of my reconstruction of trade turnover of various groups of trade firms. I cannot reconstruct trade turnover by type of firms for the provinces with missing values in tables A34-A36: Akmola, Amur, Arkhangelsk, Dagestan, Enisej, Zabajkalsk, Zakaspijsk, Irkutsk, Kars, Olonetz, Primorie, Samarkand, Sakhalin, Semipalatinsk, Semirechie, Syrdariya, Tobolsk, Tomsk, Turgaj, Ural, Fergana, Yakutiya. Because of the lack of relevant data, I use a different methodology for these provinces. In particular, I multiply the number of registered employees in trade sector in these provinces from the 1897 imperial census (from table A2.) by the average trade turnover per registered employee in trade sector in other provinces of the empire.

Table A39 reports the number of traders per city in Finland. I aggregate them at the province level in table A40. Table A40 also reports information on the number of rural traders by province from the Central Statistical Bureau (1900). Summing up rural and urban traders, I get the total number of traders in each Finnish province. I take the estimate of value added in trade in Finland in 1897 by Hjerppe (1989), and use the number of traders from the Central Statistical Bureau (1900) as weights to allocate the value added between Finnish provinces.

Table VA6 in the online appendix presents the results of my reconstruction of value added in trade.

6. Banking

I use Salomatina reconstruction (2014) of profits and expenditures on labor in this subsector (table A41) to estimate value added. For Finnish provinces, table A42 and A43 report statistics on the revenues and expenditures of Finnish banks from the Central Statistical Bureau (1899). I reconstruct value added by Finnish commercial banks, multiplying their profits by the value added-to-profit ratio of commercial banks in the Russian empire (from table A41). I reconstruct value added by Finnish savings banks by multiplying their expenditures by the average value added-to-expenditures ratio from table A42.

Table VA6 in the online appendix presents the results of my reconstruction of value added in banking.

7. Public goods: public education, health, national defense, public administration, public infrastructure

I approximate value added in these subsectors by expenditures on public goods. I sum up expenditures by the imperial government, provincial and district authorities, as well as by cities and local communities.

Table A47 presents data on imperial government expenditures by ministry at the province level from the Ministry of State Control (1898). Ministries, however, could spend their budgets on various activities. Thus, the Ministry of Education as well as the Ministry of Transportation spent on public schools. Unfortunately, official statistics did not report imperial expenditures by types of expenditures at the province level. However, Yasnopolskij (1891, 1897) reconstructed province-level imperial expenditures by type for one year, 1887, using the archives of the Ministry of State Control (table A44). He also reported province-specific expenditures by ministry for that year (table A45). From these figures, I estimate ministry-to-type-of-expenditures transformation coefficients, regressing expenditures by type on expenditures by ministry (table A46). I suppress constant terms in these regressions and exclude Saint Petersburg province as an outlier. I apply estimated transformation coefficients to 1897 province-specific expenditures by ministry (table A47) to reconstruct province-specific imperial expenditures by type (table A48). For Saint Petersburg province, I assume that shares of expenditures by type did not change between 1887 and 1897. Yasnopolskij (1891, 1897) argued that there were few changes in the composition of imperial spending over time, which justifies such an approach.

Table A49 presents province-level and district-level public expenditures in provinces with *zemstvo*s from the Ministry of Finance (1900). I aggregate them at the province level in table A50. Table A51 presents province-level public expenditures in provinces without *zemstvo*s from the Ministry of Finance (1900). Together, table A50 and A51 report public expenditures made by province and district authorities in all provinces of the Russian empire except the Great Duchy of Finland.

Table A57 reports the budget of Finland in 1897 from the Central Statistical Bureau (1899). Unfortunately, expenditures are not specified by province. I assume that expenditure on defense, administration and public infrastructure were proportional to population except for expenditures of specific types, like the expenditures on the Senate that was located in Helsinki. I assume that expenditures on education were proportional

to the number of students (table A58) with an exception of expenditures on the university, which was located in Helsinki. I use the geography of hospitals in Finland in 1898 to approximate province-level expenditures on health. Table A60 reports the results of my estimations for Finland.

In table A52, I reproduce public expenditures made by city administrations aggregated by provinces from the Ministry of Finance (1902). Table A53 reports the expenditures of Finnish cities in 1897 from the Central Statistical Bureau (1899). I aggregate them by province in table A54.

Expenditures of local communities are known only for Finnish provinces in 1897 (Central Statistical Bureau, 1899; see also table A56) and for the 50 provinces in the European part of the empire in 1894 (Central Statistical Agency, 1898; see also table A55). I assume that expenditures by local communities in the 50 provinces were the same in 1894 and 1897. I proxy expenditures in the non-European non-Finnish provinces by estimating average expenditures per rural capita in the 50 European provinces and multiply that number by the rural population in each corresponding province.

I assign each type of expenditures from table A48 (imperial government), A50 (*zemstvo* provinces), A51 (non-*zemstvo* provinces), A52 (cities), A54 (Finnish cities), A55 (communes), A56 (Finnish communes), A57 (All Finland budget), A60 (Finnish provinces) to one of five groups – administration, defense, education, health and public infrastructure – and sum them up by province and type of expenditures. Table VA7 presents the results.

8. Housing: urban and rural

There are few systematic data on housing at the province level. For urban housing, there are only data on city expenditures on public buildings (table A52). To approximate value added in urban housing, I sum up expenditures on public buildings with my approximation of private expenditures on urban housing. I estimate the later using urban population figures from table A1 relative (to the average) price of renting from table A62 and Gregory's (1982) national level estimate (without Finland) of expenditures on urban renting in 1897 (table A60). I multiply the urban share in a province by provincial relative price of renting and allocate the national-level estimate between provinces using the products of these multiplications as weights.

For Finland, I take the estimate of value added in the sector of rural and urban housing in 1897 by Hjerppe (1989) and use population as weights to allocate the value added between Finnish provinces.

Table A60 reports annual expenditures on rural housing per capita, according to various surveys (Commission on Peasant Welfare 1903). I approximate value added in rural housing by multiplying the rural population (from table A1) by the average expenditures on housing per capita.

Table VA8 in the online appendix presents the results of my reconstruction of value added in urban and rural housing.

9. Servants

Table A61 presents the mean annual earnings of male and female servants in urban settlements in the Russian empire in 1904 (Central Statistical Agency 1906b). I use the total earnings of servants in a province (estimated by multiplying annual earnings in 1904 by the total number of registered servants from table A2 and adjusted for an increase in wages between 1897 and 1904) to approximate value added by servants.

Table VA8 in the online appendix presents the results of my reconstruction of value added by servants.

AII. 1897 real GRPs reconstruction methodology

There are no local price indexes for the provinces of the Russian empire either for the 19th or the early 20th century.[2] Similar, information on local prices is scare. The only volume which reports local prices and has comprehensive geographical coverage for the considered time period is from the Central Statistical Agency (1906b). This volume provides information on prices in the cities of the empire in 1904, the closest year to my benchmark year of 1897. I assume that relative regional prices did not change much over seven years and use 1904 figures to construct relative regional price index, which I apply for my 1897 nominal estimates.

[2] The literature suggests only two series of prices indexes for Russia in the late 19th – early 20th centuries, namely, one for Saint Petersburg and one for Moscow (Kohn, 1926; Strumilin 1966).

Central Statistical Agency (1906b) contains data on six food products, namely, rye bread, wheat bread, meat (high quality and regular), and salt and sugar, for four seasons, as well as data on two non-food items (kerosene and accommodation rent) and the price of unskilled labor (separately, female and male). I use the relative prices of these goods and services in provincial capitals (relative to the corresponding averages across all provinces) to build on the relative provincial prices indexes. I use relative food prices (with equal weights) to construct the relative food price index. I view the prices of unskilled labor as proxies for the prices of services. I use them along with the prices of non-food goods (again with equal weights) to construct the relative non-food price index. Given the high share of food items in the expenditures of Russian households at that time, I follow Mironov (2010) and apply 0.7 and 0.3 weights for food and non-food price indexes to construct the unified relative price index. There are no prices for Sakhalin and Finland in the original statistical volume. I use price indexes in neighbor provinces. Table A62 presents original prices for individual items and my estimates of indexes.

I use the obtained price index to switch from nominal to real gross regional product. Table A63 presents the results.

AIII. 1897 regional labor productivity reconstruction methodology

The 1897 census suffers from an under registration problem in regard to gainfully employed workers. In family firms, the census counted only the heads of households as gainfully employed. All other family members were registered as dependents irrespective of whether they actually contributed to the family businesses or not (Gozulov 1936 pp. 205-206). Because of peasant family farms, the problem was worse in agriculture and less pronounced in other sectors. The under-enumeration of employment in the census makes it problematic to use the number of registered workers as a denominator to estimate labor productivity.

At the same time, various household surveys conducted by *zemstvo* unanimously show that between 99 and 100 percent of all able-bodied members of working age in the households of peasants and urban dwellers participated in the labor force (Svavitskij and Svavitskij, 1926). The share of rentiers was negligible (Troijnitskij 1905). The share of members unable to work was about one percent of the working age population. Because

of all that, to estimate labor productivity I divide real gross regional product by number of working age citizens. I use the working age definition elaborated by Russian statistical congress in 1887, namely, males between 16 and 59, and females between 16 and 54. This is also the definition which the *zemstvo* surveys used (Svavitskij and Svavitskij, 1926).[3] Table A63 presents the number of rural and urban citizens in the working age by gender and province. Table VA9 reports my estimates of labor productivity.

AIV. Robustness of GRPs estimates

In this section, I discuss how robust my estimates are if I relax assumptions, which push my estimates up (1) and on which there are either debates in the literature or in regard to which the literature does not give clear recommendations (2). Table B1 in the online appendix reports summary statistics for the alternative sets of estimates, and table A65 reports the alternative estimates for each single province. Below, I illustrate the magnitudes of potential biases, discussing the difference between baseline and alternative estimates of income at the national level.

If I apply no upward correction coefficients to the output of cereals (which would most likely be wrong, according to many contemporary statisticians as discussed in Wheatcroft and Davies 1994), total national income drops from the original 9,701 mln rubles (the baseline estimate for the whole empire including Finland) to 9,388 mln rubles; these are 82.5 and 80.25 rubles in per capita terms, respectively.[4] If I employ the 10 percent correction originally suggested by Ivantsov (1915), the difference is smaller, namely 82.5 vs. 81.4 rubbles per capita.

Potential double-counting in the construction sector and public infrastructure expenditures also cannot produce a large measurement error because these sectors only partially intersect and were relatively small in the late Russian empire, namely, 2.1 and 1.25 per cent of national income respectively. If I exclude public infrastructure

[3] For Finland, I use Central Statistical Bureau (1903). Due to the lack of regional decomposition figures, I assume that the age structure was the same across the Finnish provinces.

[4] In the baseline reconstruction, I apply 19 upward corrections to grain output everywhere in the empire outside Finland (or to about 20 percent of total output). Accordingly, the lack of such correction transforms into more than a three percent drop in total output.

completely (that most likely would be wrong), nominal GRP per capita in an average province drops from 82.5 to 81.5 rubles.

Another potential source of double counting are revenues from sales of horses to cities, and revenues from sales of hay and straw to cities, which might appear as intermediate goods in trucking services. But these two subsectors were relatively small and generated less than one percent of value added. If I exclude these sectors, nominal GRP per capita in an average province drops from 82.5 to 82.44 or 82.21 rubles, respectively.

If instead of subsector-specific ratios of value added to gross industrial production in 1913 in large industry from the 1918 industrial census (Central Statistical Agency 1924) I use 1900 ratios (Varzar 1903), the change results in only a 3.5 ruble drop in GRP per capita, from 82.5 to 79 rubles in per capita terms. In the baseline estimates I prefer 1913 to 1900 ratios because the later are of problematic quality. Varzar (1903) himself noted that a number of industrial firms reported actual costs of production rather than market value of produced goods. This might explain lower ratios of value added to gross industrial production in Varzar (1903) than in Central Statistical Agency (1924).

Potential double counting of cereals used as fodder for livestock is another source of potential upward bias in my estimates. Unfortunately, there is no solid micro data for the late-19th century to adjust for this problem. Nefedov (2011) cites Vilson's estimations for the mid-19th century and a 1917 estimation by the Ministry of Food Provision and concludes that consumption of cereals by livestock substantially increased due to the transformation of meadows into arable lands. However, both the speed and geography of this transformation is poorly documented. Because of that, I follow Prokopovich (1918) and do not apply any corrections for grain consumed by livestock in my baseline estimates. If one uses 1917 norms of cereals consumption by livestock, this would result in a 6.5 rubles drop in GRP per capita at the national level. Such a drop would slash more than half the value added in animal husbandry (about 13 rubles per capita, according to the baseline estimate), which looks unrealistic.

AV. Data description and sources for variables used in the regression analysis

This section lists sources for variables which I use in the regression analysis in

section 5 of the paper and describes them in full details. Table B2 in the online appendix provides summary statistics of these variables.

I start with two dependent variables – *labor productivity and real income per capita*. They are constructed as discussed in section AII and AIII of this appendix.

I construct *sea dummy and distance to Moscow variable* using an historical map with administrative borders of Russian provinces (constructed as a part of *ristat.org* project by Kessler and Markevich 2019, forthcoming).

Share of employment in mining is estimated from data on occupations in table A2 in the online appendix and estimates of population in the working age from table A63 in the online appendix. For Finland, data on employment in mining is from Central Statistical Bureau (1900).

Share of serfs is estimated as the number of serfs in a province in 1858 divided by the regional population. For the 50 European provinces as well as for the Caucasian, Siberian and steppe provinces, the total population and number of serfs are from Troijnitskij (1861). I sum up all categories of serfs to get their total number (namely, private serfs, serfs-servants, serfs belonging to institutions, industrial serfs, serfs under "conditional law"). I use district-level data to account for changes in the administrative borders of the provinces between 1858 and 1897. For Polish provinces, I assign zeros because Polish peasants were formerly free since the abolition of serfdom by Napoleon in this region (Kostushko 1962, Milutin 1863). For Finnish provinces, I assign zeros as well because there was no serfdom in Finland (Hejrpe 1989). I assume zero serfs in the Central Asia provinces, which the Russian empire conquered in the late 1860s and 1870s, i.e. after the abolition of serfdom in European Russia, and where was no serfdom. Similar, I assume zero serfs in Sakhalin provinces in the Far East and Kars and Chernomorsk provinces in the Caucasus region, which became part of the empire after the abolition of serfdom.

Number of deposits fields known by 1920 is from USGS Major mineral deposits of the world dataset and USGS Mineral operations outside the United States dataset (https://mrdata.usgs.gov/major-deposits/). They are obtained by deducting fields which were discovered between 1920 and 2005 from the fields known by 2005, excluding those which were not in operation any more (1), and adding fields which had

explicitly been in development at some point before 1920 but were discontinued by 2005 (2). To get per unit of area figures, I divide the number of deposits fields by province area derived from historical map with administrative borders of Russian provinces (constructed as a part of *ristat.org* project by Kessler and Markevich 2019, forthcoming).

I extend data on the *share of former monasterial serfs* in the 50 European provinces from Markevich and Zhuravskaya (2018) to the rest of the empire using Beskrovnii et al. (1972). Beskrovnii et al. (1972) reports information about the number of (former) monasterial serfs in 1796 and 1814. I follow Markevich and Zhuravskaya (2018) and take an average of the shares of monasterial serfs for these two periods.

The number of monasteries by the moment of the nationalization is from Zverinskij (1890-1897).

Share of large estates is estimated as the number of serfs in estates with more than 21 serfs over the total number of estates in the province from Troijnitskij (1861).

Data on *shares of state lands and commune lands* are from Central Statistical Agency (1911). For the 50 European provinces, the Central Statistical Agency (1911) reports the results of a 1905 land survey. Data on the 10 Polish provinces are from 1907. There are no land data for other provinces.

Data on *shares of in-migrants and out-migrants, share of urban population, literacy* are from the 1897 population census (Trojnitskij 1905, Trojnitskij 1903-1905). There are no data on migration and literacy for the Finnish provinces, because the 1897 population census did not cover the Great Duchy of Finland. Urban population figures for Finland are from Central Statistical Bureau (1900). *Population density, urban population density and rural population density* are estimated by taking population figures and dividing them by the area of the corresponding province. The areas of the provinces are derived from a historical map with the administrative borders of the Russian provinces (constructed as a part of *ristat.org* project by Kessler and Markevich 2019, forthcoming).

To get *market potential* I calculate internal market potential and foreign market potential. I calculate internal market potential as the inverse distance-weighted sum of nominal incomes of other regions of the Russian empire (known from the present study). I calculate foreign markets as the inverse distance-weighted sum of GDPs of foreign countries (taken from the Maddison dataset). To construct *inverse distances,* I use a

historical map with the administrative borders of the Russian provinces (constructed as a part of *ristat.org* project by Kessler and Markevich 2019, forthcoming).

Land quality is an average suitability of land for rye, wheat or barley. Suitability of land for particular cereal is a province-level, area-weighted mean cereal suitability measure calculated off the pixel-level raster data on cereal suitability under low-input, rain-fed conditions taken from FAO GAEZ.

Krugman specialization indexes for agriculture, industry and services are constructed from the estimates of value added in the corresponding subsectors. For each province, I first construct a set of its pairwise specialization metrics reflecting the difference between this province and each province from the set of other provinces in terms of the composition of value added in the corresponding sector (industry, agriculture, services), and then take the average over these pairwise metrics for the corresponding sector. To be precise, I calculate pairwise metric of similarity between provinces j and k as $M_{jk} = \sum_{i=1}^{n} | \frac{VA_{ij}}{VA_j} - \frac{VA_{ik}}{VA_k} |$ where VA_{ij} is the value added in subsector i=1,..,n for province j, and VA_j is the value added in the corresponding sector for province j and similarly for province k (if the pairwise metric is equal to zero, this means that two provinces are completely the same in terms of composition of value added; if the pairwise metric is equal to two, the provinces are completely different). At the second step, I estimate the index for province j as $I_J = \frac{\sum_{k \neq j} M_{jk}}{n-1}$.

Mean industrial firm revenue in 1894 is from Gregg (2019). The original source is the Ministry of Finance, Department of Trade and Manufacturing (1897). There are no data on 18 out of 97 provinces of the empire.

Ruggedness is constructed following replication files from Nunn & Puga (2012). As in their paper, it is surface area-weighted measure of ruggedness.

Length of rivers is the cumulative length in kilometers of all the rivers within a province's boundaries. Vector layer with rivers is from the Open Street Map (https://www.openstreetmap.org).

Mean temperature in 1901 – 1910 is the average monthly temperature in the decade between 1901 and 1910. Monthly figures are daily average temperature (in degrees Celsius) from Climatic Research Unit TS 4.02 data

(see https://crudata.uea.ac.uk/cru/data/hrg/).

Mean precipitation in 1901 – 1910 is the average monthly precipitation for the decade between 1901 and 1910. Monthly figures (in millimeters) are from Climatic Research Unit TS 4.02 data (see https://crudata.uea.ac.uk/cru/data/hrg/).

Mean vapor pressure in 1901 – 1910 is the average monthly vapor pressure for the decade between 1901 and 1910. Monthly figures (in hectopascals) are from Climatic Research Unit TS 4.02 data (see https://crudata.uea.ac.uk/cru/data/hrg/).

Mean relative humidity in 1901 – 1910 is the average monthly relative humidity for the decade between 1901 and 1910. Monthly figures are calculated following Crafts and Wolf (2014, page 1112, footnote 3) using Lawrence's (2005) approximation for relative humidity from time-corresponding temperature and vapor pressure figures.

Number of frost days is the average number of days with temperature below zero in each month for the decade between 1901 and 1910. Monthly figures obtained from Climatic Research Unit TS 4.02 data (see https://crudata.uea.ac.uk/cru/data/hrg/).

Longitude and *latitude* are longitude and latitude of province's centroids estimated using an historical map with the administrative borders of the Russian provinces (constructed as a part of *ristat.org* project by Kessler and Markevich 2019, forthcoming).

Zemsto dummy equals one in the provinces with elected local governments (*zemstvo*s). *Oblast* dummy stands for characteristics of a province status (*oblast* vs. *guberniya*). Administrative status of the regions is from Andreevskij, Arsentiev, Petrushevskij (1890-1907).

Share of cities with 1892 city charters is from Ministry of Internal Affairs (1901). This share is not specified for Sakhalin because of lack of cities there.

Number of years in the Russian empire derived from Andreevskij, Arsentiev, Petrushevskij (1890-1907).

Share of orthodox is from the 1897 population census (Trojnitskij 1905). The Central Statistical Bureau (1903) is the source for the share of orthodox in Finnish provinces. Due to data limitation, I use 1900 figures rather than 1897 figures in the case of Finland.

References:

Andreevskij I.E., Arsentiev K.K., Petrushevskij F.F. (1890-1907). *Entsiklopedicheskij slovar Brokgauza i Efrona* [Brokgauz and Efron encyclopedia]. Saint Petersburg.

Beskrovnii L.G., Ya.E. Vodarskii, and V.M. Kabuzan. (1972). *Perepisi naseleniya Rossii: itogovie materiali podvornikh perepisei I revizii naseleniya Rossii, 1646 – 1858* [Russian Population Censuses: results of household censuses and tax censuses in Russia, 1646 - 1858]. Moscow: Institut istorii AN SSSR.

Blau A.A. (Ed.) (1900). *Svod dannikh o fabrichno-zavodskoj promishlennosti v Rossii za 1897 g.* [Dataset on large industry in Russia in 1897] Saint Petersburg.

Central statistical agency (1898a). *Mirskie dokhody i raskhody za 1894 v 50 guberniyakh Evropejskoj Rossii* [Incomes and expenditures of local communities in 1894 in 50 provinces of European Russia]. Saint Petersburg.

Central statistical agency (1898b). *Urozhaj za 1897* [Yields in 1897]. Vol. 1 and 2. Saint Petersburg.

Central statistical agency (1902). *Svod statisticheskikh svedenij po selskomu khozyastvu Rossii v kontse XIX veka* [A collection of statistical figures on agriculture in the late 19th century] Saint Peterburg.

Central statistical agency (1906a). *Ezhegodnik Rossij. 1905. God 2-j*. [Annual statistical book on Russia. 1905. Vol. 2], Saint Petersburg.

Central statistical agency (1906b). *Goroda Rossii v 1904 g.* [Russian cities in 1904]. Saint Petersburg.

Central statistical agency (1911). *Ezhegodnik Rossij. 1910. God 7-j.* [Annual statistical book on Russia. 1910. Vol. 7], Saint Petersburg.

Central statistical agency (1923). *Trudi TsSU* [Central statistical agency proceedings]. Vol. 18. *Sbornik statisticheskih svedenij po SSSR. 1918-1923* [Statistical volume on the USSR. 1918-1923]. Moscow.

Central statistical agency (1924). *Trudi TsSU.* [Central statistical agency proceedings]. Vol. 18. *Sbornik Statisticheskih svedenij po SSSR. 1918 - 1923 gg. Za pyat let raboty TsSU* [Works of the Central Statistical Agency of the Soviet Union. Vol. 18. Statistical volume on the USSR, 1918-1923.] Moscow.

Central Statistical Bureau. (1900). *Statisticheskij Ezhegodnik Finlyandii. God 20-ij.* [Annual statistical volume of Finland. Vol. 20.]. Helsinki.

Central Statistical Bureau. (1900). *Statisticheskij Ezhegodnik Finlyandii. God 21-ij.* [Annual statistical volume of Finland. Vol. 21.]. Helsinki.

Central Statistical Bureau. (1901). *Statisticheskij Ezhegodnik Finlyandii. God 22-ij.* [Annual statistical volume of Finland. Vol. 22.]. Helsinki.

Central Statistical Bureau. (1903). *Statisticheskij Ezhegodnik Finlyandii. Novaya seriya. God 1-ij.* [Annual statistical volume of Finland. New series. Vol. 1.]. Helsinki, 1903;

Commission on peasant welfare (1903). *Materialy Vysochajshe utverzhdennoj 16.11.1901 komissii po issledovaniu voprosa o dvizhenii s 1861 po 1900 blagosostoyaniya selskogo naseleniya srednezemledelcheskikh gubernij, srvanitelno s drugimi mestnostyami Evropejskoj Rossii* [Materials of the 16.11.1901 Govermental Commision] Vol. 1, Saint Petersburg, 1903, Pp. 216-219.

Crafts, Nicholas and Wolf, Nikolaus, (2014). "The Location of the UK Cotton Textiles Industry in 1838: A Quantitative Analysis," *The Journal of Economic History*,

74(4): 1103-1139.
Department of communication (1898). *Pochtovo-telegraphnaya statistika za 1897 god* [Statistics on post and telegraph for 1897]. Saint Petersburg.
Forest department (1898). *Otchet po lesnomu upravleniu ministerstva zemledeliya I gosudarstvennikh imushchestv za 1897 god* [1897 report of forest department]. Saint Petersburg.
Gregg, Amanda. (2019). "Factory Productivity and the Concession System of Incorporation in Late Imperial Russia," *Working Paper.*
Gozulov A.I. (1936). *Perepisi naseleniya SSSR i kapitalisticheskikh stran (opit istoriko-metodologicheskoj kharakteristiki proizvodsvta perepisej naseleniya)* [Population censuses in the USSR and capitalist countries]. Moscow: TsSUNKhU press, 1936.
Governor of Amur province. (1899). Obzor Amurskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Amur province in 1897. Appendix to the annual governor report to the emperor]. Blagoveshchensk.
Governor of Baku province. (1903). Obzor Bakinskoj oblasti in 1902. Prilozhenie k vsepoddanejshemu otchetu [Review of Baku province in 1902. Appendix to the annual governor report to the emperor]. Baku.
Governor of Chernomorsk province. (1898). Obzor Chernomorskoj gubernii in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Chernomorsk province in 1897. Appendix to the annual governor report to the emperor]. Novorossijsk.
Governor of Dagestan province. (1898). Obzor Dagestanskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Dagestan province in 1897. Appendix to the annual governor report to the emperor]. T-Kh-Shura.
Governor of Elizavetopol province. (1898). Obzor Elizavetpolskoj gubernii in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Elisavetopol province in 1897. Appendix to the annual governor report to the emperor]. Tiflis.
Governor of Erivan province. (1898). Obzor Erivanskoj gubernii in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Erivan province in 1897. Appendix to the annual governor report to the emperor]. Erivan.
Governor of Fergana province. (1898). Obzor Ferganskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Fergana province in 1897. Appendix to the annual governor report to the emperor]. Vladivostok.
Governor of Kutais province. (1898). Obzor Kutaiskoj gubernii in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Kutais province in 1897. Appendix to the annual governor report to the emperor Kutais.
Governor of Primorie province. (1899). Obzor Primoskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Primorie province in 1897. Appendix to the annual governor report to the emperor]. Vladivostok.
Governor of Sakhalin province. (1900). Obzor ostrova Sakhalin in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Sakhalin province in 1897. Appendix to the annual governor report to the emperor]. Saint Petersburg.
Governor of Samarkand province. (1898). Obzor Samarkandskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Samarkand province in 1897. Appendix to the annual governor report to the emperor]. Samarkand.

Governor of Syr-dariya province. (1897). Obzor Syr-dariinskoj gubernii in 1895. Prilozhenie k vsepoddanejshemu otchetu [Review of Syr-dariya province in 1895. Appendix to the annual governor report to the emperor]. Tashkent.
Governor of the Kars province. (1898). Obzor Karskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Kars province in 1897. Appendix to the annual governor report to the emperor]. Kars.
Governor of Tiflis province. (1898). Obzor Tiflisskoj gubernii in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Tiflis province in 1897. Appendix to the annual governor report to the emperor]. Tiflis.
Governor of Ural province. (1898). Obzor Uralskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Ural province in 1897. Appendix to the annual governor report to the emperor]. Uralsk.
Governor of Yakutsk province. (1898). Obzor Yakutskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Yakutsk province in 1897. Appendix to the annual governor report to the emperor]. Yakutsk.
Governor of Zabajkalie province. (1898). Obzor Zabajkalskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Zabajkalie province in 1897. Appendix to the annual governor report to the emperor]. Chita.
Governor of Zakaspijskaya province. (1898). Obzor Zakaspijskoj oblasti in 1897. Prilozhenie k vsepoddanejshemu otchetu [Review of Zakaspijskaya province in 1897. Appendix to the annual governor report to the emperor]. Ashkhabad.
Gregory Paul (1982). *Russian National Income, 1885-1913*. Cambridge: Cambridge University Press.
Groman V.V. (1912). *Obzor stroitelnoj deyatelnosti v Rossii* [Review of construction industry in Russia]. Saint Petersburg.
Hjerpee, Riitta. (1989) *The Finnish Economy, 1860-1985. Growth and Structural Change*. Helsinki: Government Printing Center.
Ivantsov D.N. (1915). *K kritike russkoj urozhajnoj statistiki (opit analiza nekotorikh ofitsialnikh i zemskikh tekushchikh dannikh)*. [Critical notes on the Russian grain yield statistics (An experience of analysis of Russian official and zemstvo data]. Petrograd.
Kessler, Gijs and Andrei Markevich (2019, forthcoming), *Electronic Repository of Russian Historical Statistics, 18th - 21st centuries*. Version 2., http://ristat.org/
Kohn, M.P. (1926). *Russkie indeksy tsen* [Russian price indexes]. Moscow: Ekonomichekaya zhizh.
Kostushko I.I. (1962). *Krestyanskay reforma 1864 v tzarstve Polskom* [1864 peasant land reform in the kingdom of Poland]. Moscow, 1962.
Lawrence, Mark G. "The Relationship between Relative Humidity and the Dewpoint Temperature in Moist Air: A Simple Conversion and Applications." *Bulletin of the American Meteorological Society* 86, no. 2 (2005): 225–33.
Markevich, Andrei and Mark Harrison (2011). 'Great War, Civil War, and Recovery: Russia's National Income, 1913 to 1928', *Journal of Economic History,* 71(3): 672-703.
Markevich, Andrei and Ekaterina Zhuravskaya (2018).). "Economic Effects of the Abolition of Serfdom: Evidence from the Russian Empire," *The American Economic Review.* 108(4-5): 1074-1117.

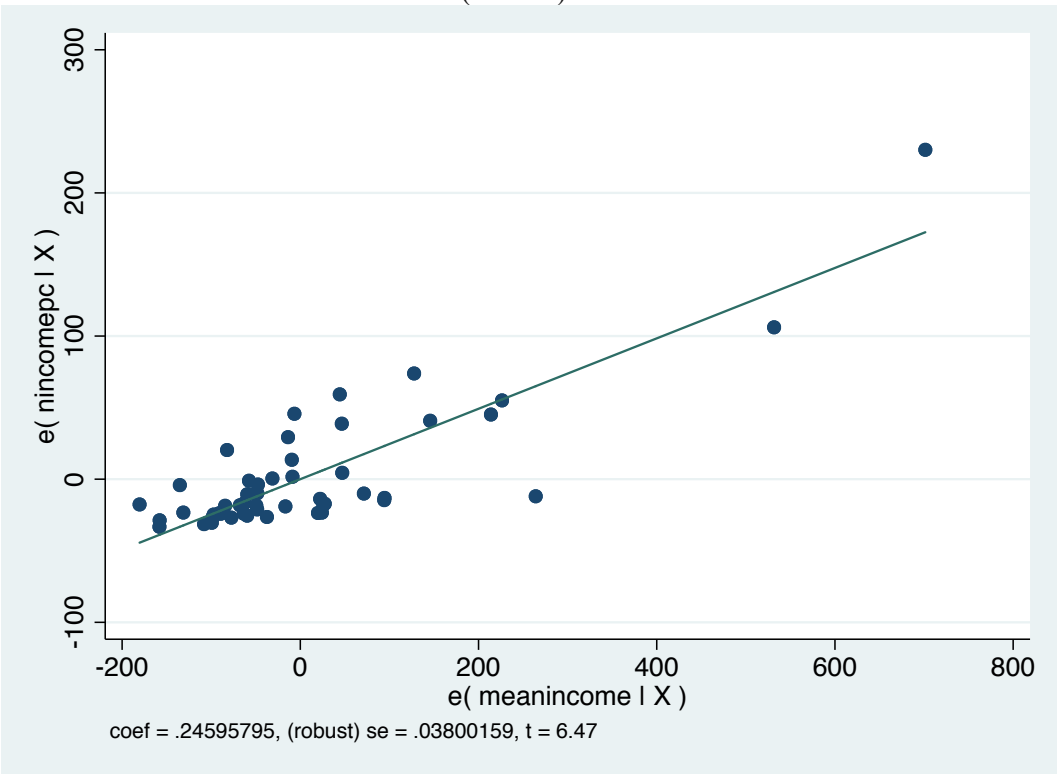
Milutin N.A. (1863) *Issledovaniya v Tzarstve Pol'skom* [Surveys in the kingdom of Poland]. Saint Peterburg.
Ministry of Agriculture and State Property (1897-1898). *1897 god v selskokhozyastvennom otnoshenii po otvetam, poluchennim ot khozyaev* [1897 agricultural year according to answers from farmers]. Vol. 1-6. Saint-Petersburg.
Ministry of Finance Department of Trade and Manufacturing (1897). *Fabrichno-zavodskaya promishlennost Rossii: Perechen' fabric i zavodov*. [Russian Factory Production: List of Factories and Plants]. Saint Petersburg.
Ministry of Internal Affairs (1900). *Otchet veterinarnogo upravleniya MVD za 1897* [1897 report of veterinary department of the Ministry of internal affairs]. Saint Petersburg.
Ministry of Internal Affairs (1901). *Spisok gorodskikh poselenij Rossijskoj imperii* [A list of urban settlements of the Russian empire]. Part 1. *Goroda* [Cities]. Saint Petersburg.
Ministry of Transportation. (1898). Statisticheskij sbornik Ministerstva Putej Soobshcheniya. Svedeniya o vodnikh putyakh za 1897 [Statistical volume of the Ministry of roads. Data on waterways in 1897]. Saint Petersburg.
Ministry of Transportation. (1899). Statisticheskij sbornik Ministerstva Putej Soobshcheniya. Svedeniya o zheleznikh dorogakh za 1897 [Statistical volume of the Ministry of roads. Data on railroads in 1897]. Saint Petersburg.
Ministry of Finance. (1898). *Svod tovarnikh tsen na glavnikh russkikh I inostrannikh rinkakh za 1897 god* [Review of prices at the main Russian and foreign markerts]. Saint Petersburg.
Ministry of Finance (1900a). *Ezhegodnik Ministerstva Finansov* [Annual volume of the Ministry of Finance]. Vol. 29. 1899 Saint Petersburg.
Ministry of Finance (1900b). *Statisticheskie rezultati protsentnogo i raskladochnogo sborov za 1896, 1897, 1898 godi po ischisleniyu, klassifikatsii i opredelenii oborotov i pribilej torgovikh i promishlennikh predpriyatij podlezhashchikh sim sboram* [Statistics on trade taxes for 1896, 1897 and 1898] (1900). Saint Petersburg.
Ministry of Finance (1902). *Ezhegodnik Ministerstva Finansov* [Annual volume of the Ministry of Finance]. Vol. 31. 1901. Saint Petersburg.
Ministry of State Control (1898). *Otchet gosudarstvennogo kontrolya po ispolneniyu gosudarstvennoj rospisi i finansovikh smet za 1897* [Report on state budget expenditures in 1897]. Saint Petersburg.
Mironov B.N. (1985). *Roznichnie tzeni za dva stoletiya (18-19 vv).* [Retail prices over two centuries, 18-19 Cc.] Leningrad: Nauka.
Mironov B.N. (2010). *Blagosostoyanie naseleniya I revolutsii v imperskoj Rossii*, Moscow.
Nefedov S.A. (2011). K diskussii ob urovne potrebleniya v poreformennoj i predrevolutsionnoj Rossii [To the discussion on the level of consumption in post-emancipation and pre-revolutionary Russia]. *Rossijskaya istoriya.* 2011(1): 73-86.
Nunn, Nathan and Diego Puga (2012). 'Ruggedness: The Blessing of Bad Geography in Africa', *Review of Economics and Statistics,* 94(1): 20-36.
Prokopovich S.N. Ed. (1918). *Opyt ischisleniya narodnogo dokhoda 50 gubernii Evropejskoj Rossii v 1900-1913* [Estimations of national income of 50 provinces of

European provinces in 1900-1913]. Moscow.
Salomatina S. (2014). Rekonstruktsiya dannikh v istoricheskoj statistike: otdeleniya kommercheskikh bankov Rossijskojimperii v 1897 g. [Data reconstruction of historical statistics: riginal branches of commercial banks in the Russian empire in 1897]. *Vestnik Permskogo universiteta. Seriya Istoriya.* [Journal of Perm university. History seria], 26(3): 87-98.
Silantiev A.A. (1898). *Obzor promislovikh okhot v Rossii* [A review of industrial hunting in Russia], Saint Petersburg.
Strumilin S.G. (1924). Tovarooborot v Rossii [Trade turnover in Russia]. In *Ekonomicheskoe obozrenie*. Vol. 19. Moscow.
Strumilin S.G. (1966). *Ocherki ekonomicheskoj istorii Rossii i SSSR* [Essays on economic history of Russia and the USSR]. Moscow: Nauka.
Svavitskij Z.M. and N.A. Svavitskij (1926) *Zemskie podvornie perepisi, 1880-1913. Pouezdnie itogi* [Zemstvo household surveys, 1880-1913. District-level results]. Moscow: TSU.
Trojnitskij N. A. (ed.). (1905). *Obshchij svod po imperii resultatov razrabotki dannikh pervoj vseobshchej perepisi naseleniya, proizvedennoj 28.01.1897* [The First Population 1897 Census of the Russian Empire. General results] Vols.1-2. Saint Peterburg.
Trojnitskij N. A. (ed.). *Pervaya Vseobshchaya perepis' naseleniya Rossijskoj imperii 1897 goda. (Gubernskie itogi)*. [The First Population 1897 Census of the Russian Empire (Province-lelel results)] Vols.1-89. Saint Peterburg, 1903-1905.
Trojnitskij N. A. (Ed.). (1905). *Raspredelenie naseleniya po vidam glavnokh zanyatij I vozrastnim gruppam po otdelnim territorialnim rajonam. Tablitza 20.* [The distribution of population by occupations, age and regions] Vol., 1-4. Saint Petersburg 1905.
Trojnitskij A.G. (1861). *Krepostnoe naselenie po 10 narodnoj perepisi* [Serf population according to the 10th tax census]. Saint Petersburg.
Vainshtein A.L. (1960). Iz istorii predrevolutsionnoj statistiki zhivotnovodstva [From history of pre-revolutionary statistics of husbandry] in *Ocherki po istorii statistiki SSSR*. Moscow.
Varzar V.E. (1903) *Spisok fabric I zavodov Evropejskoj Rossii* [Ministry of Finance of the Russian Empire. List of Factories and Plants of European Russia]. St. Petersburg.
Wheatcroft S.G. and R.W. Davies (1994) 'Agriculture', in *The Economic Transformation of the Soviet Union, 1913-1945*. Eds. R.W. Davies, Mark Harrison and S.G. Wheatcroft. New York.
Yasnopolskij N.P. (1891, 1897) *O geographisheskom raspredelenii gosudarstvennikh dokhodov i raskhodov Rossii* [On geography of state incomes and expenditures], Kiev. Vol. 1-3.
Zverinskij V.V. (1890-1897). *Material dlya istoriko-topographicheskogo issledovaniya o pravoslavnikh monastiryakh Rossijskoj imperii* [Materials for historical and geographical studies on orthodox monasteries in Russian empire]. Saint-Petersburg. Vol. 1-3.

Tables and figures.

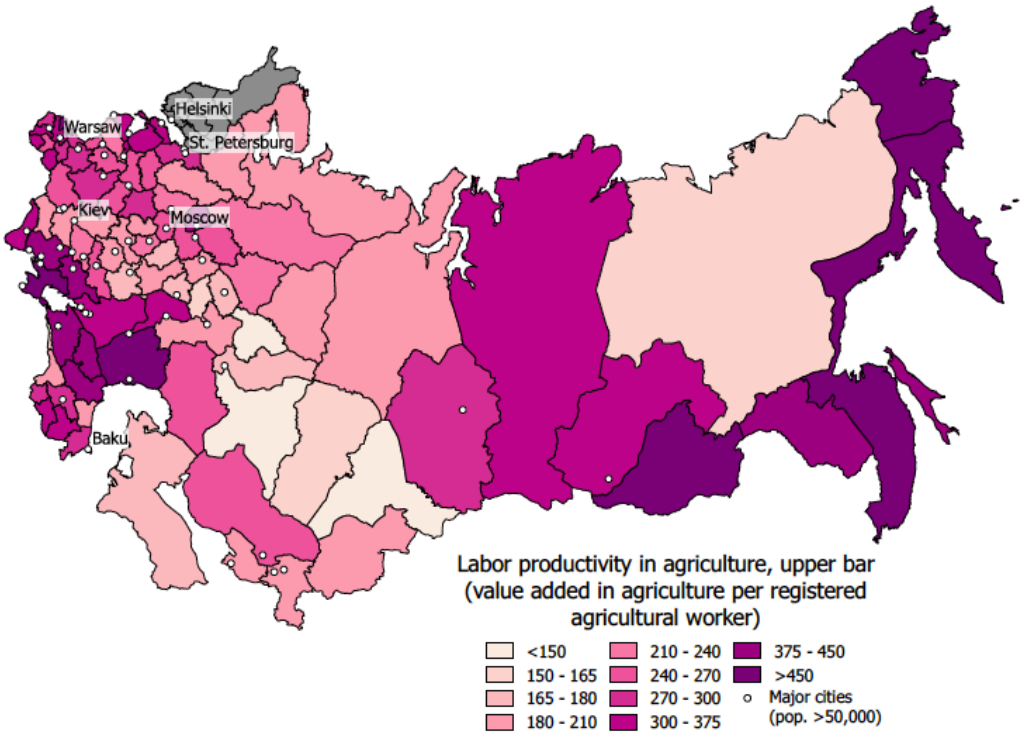
Consult separate .xls file for appendix tables A1-A65 and VA0-VA10.

Figure B1. Validity check: Markevich 1897 GRPs per capita reconstructed from the production size (Y-axis) vs. Lindert and Nafziger 1904 mean household income reconstructed from the income side (X-axis).



Source: 1897 per capita GRPs – the current study; 1904 household mean income – Lindert and Nafziger (2014).

Figure B2. Labor productivity in agriculture (upper bar) in the provinces of the Russian empire in 1897 (1897 rubles)



Source: the current study.

Figure 3. Conditional scatter plot for labor productivity in 1897 and sea dummy

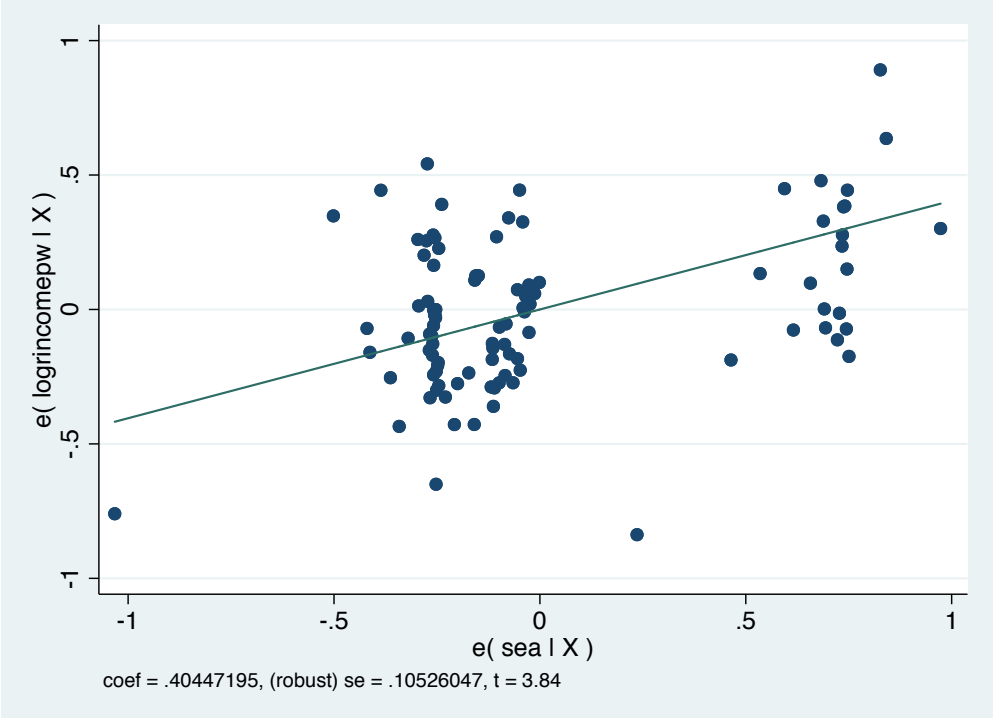


Figure 4. Conditional scatter plot for labor productivity in 1897 and the share of employment in mining

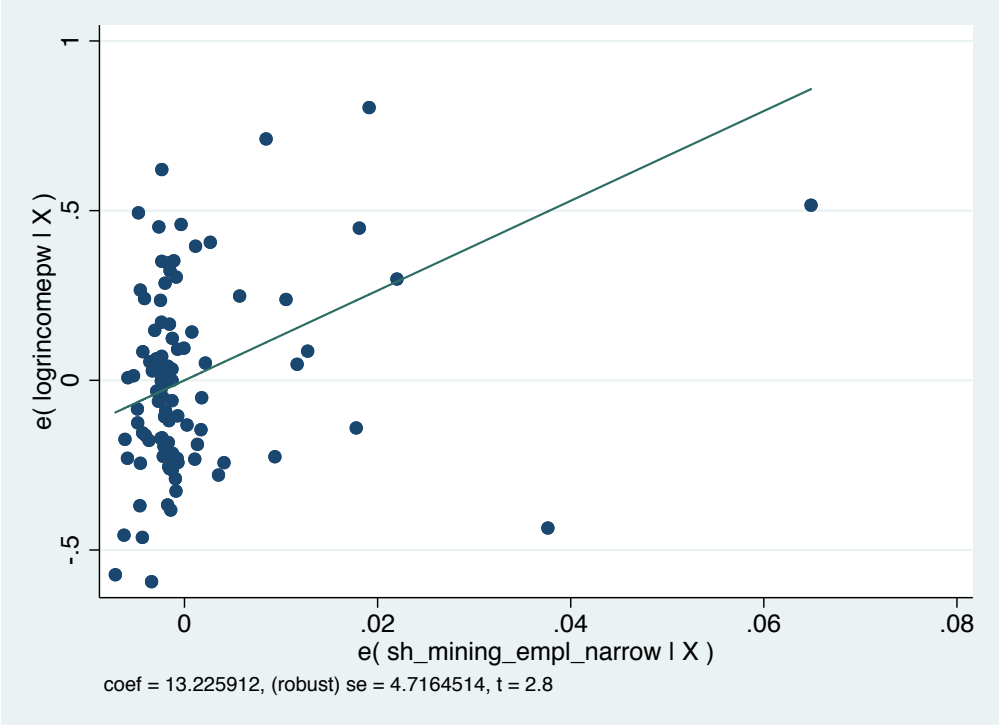


Figure 5. Conditional scatter plot for labor productivity in 1897 and the share of serfs in 1858

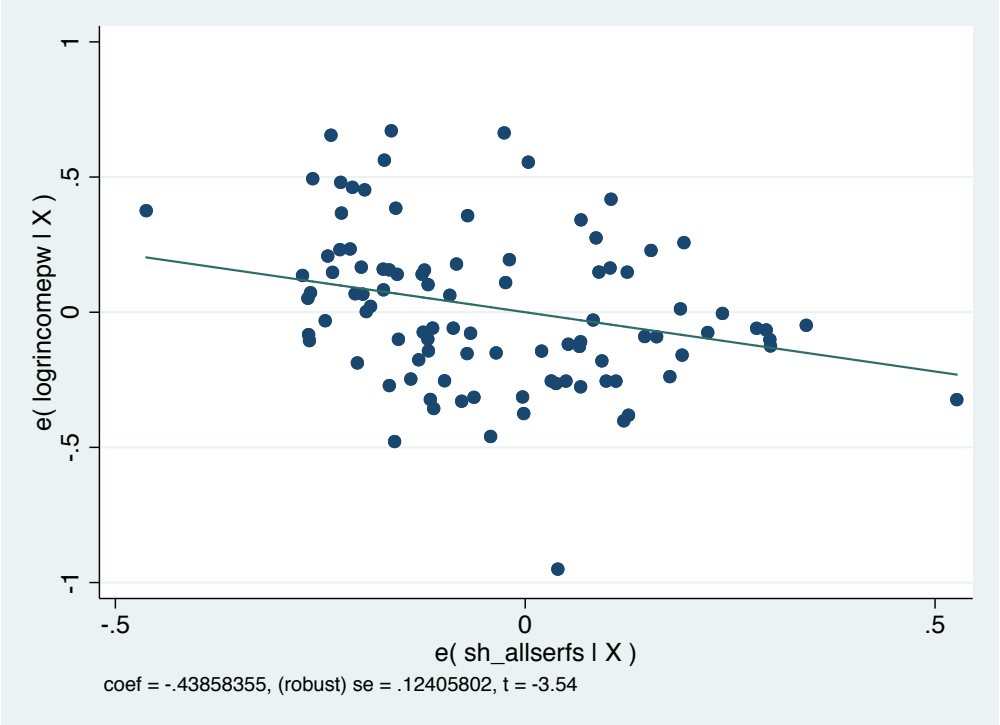


Table B1. Robustness. Summary statistics of alternative estimates of GRPs

Variable	Obs	Mean	Std. Dev.	Min	Max
Gross regional product per capita in roubles (nominal), baseline estimate	97	82.61	45.86	28.21	303.72
Gross regional product per capita in roubles (nominal), no upward cereals correction	97	80.35	45.89	27.51	302.85
Gross regional product per capita in roubles (nominal), no public infrastructure	97	81.59	45.24	27.82	297.11
Gross regional product per capita in roubles (nominal), no revenues from sales of horses	97	82.56	45.87	28.08	303.70
Gross regional product per capita in roubles (nominal), no revenues from sales of hay and	97	82.33	45.68	28.16	301.87
Gross regional product per capita in roubles (nominal), 1900 value added to output	97	79.15	43.30	28.13	289.18
Gross regional product per capita in roubles (nominal), correction for fodder consumption	97	76.05	45.87	18.70	301.98

Source: the current study.

Table B2. Summary statistics of variables used in the regression analysis

Variable	Obs	Mean	Std. Dev.	Min	Max
Sea dummy	97	0.237	0.428	0	1
Share of employment in mining	97	0.004	0.01	0.00001	0.071
Share of serfs in 1858	97	0.194	0.238	0	0.691
Log distance to Moscow	97	6.907	0.839	3.188	8.761
Number of deposit fields per sq km	97	0.000015	0.000025	0	0.00016
Share of (former) monasterial serfs	97	0.045	0.073	0	0.391
Dummy for the presence of monasteries in	97	0.546	0.5	0	1
Share of large estates in 1858	97	0.5	0.472	0	0.999
Share of state lands	60	0.159	0.177	0.0234	0.994
Share of commune lands	60	0.45	0.131	0.005	0.716
Share of in-migrants	89	0.053	0.064	0.002	0.371
Share of out-migrants	89	0.082	0.052	0.004	0.216
Share of urban populatin	97	0.131	0.099	0	0.673
Literacy	89	0.219	0.15	0.026	0.799
Market potentilal (log)	97	10.566	0.336	9.707	11.307
Sums of reverse distances (log)	97	-2.249	0.462	-3.727	-1.756
Population density (log)	97	2.597	1.627	-2.689	4.888
Urban density (log)	97	0.315	2.17	-11.238	4.062
Rural density (log)	97	2.447	1.6	-2.724	4.312
Land suitability for cereals (log)	97	3.339	2.571	-20.723	4.561
Specialization index for agriculture	97	0.194	0.068	0.135	0.584
Specialization index for industry	97	0.203	0.194	0.124	1.407
Specialization index for services	97	0.154	0.179	0.101	1.717
Mean industrial firm revenue (log)	79	10.941	0.765	8.248	12.491

Table B3. Labor productivity and geography. Robustness

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	1897 labor productivity (log)									
Ruggedness (log)	-0.02									0.13
	[0.044]									[0.091]
Log length of rivers		0.06								0.02
		[0.065]								[0.105]
Log temerature in 1901-1910 (mean)			0.16							-0.17
			[0.103]							[0.533]
Log precipitation in 1901-1910 (mean)				0.16						-0.029
				[0.128]						[0.327]
Log relative humidity in 1901-1910 (mean)					0.24					-0.74
					[0.206]					[0.811]
Log vapor preassure in 1901-1910 (mean)						0.33*				[0.924]
						[0.173]				[0.651]
Land suitability for cereals (log)							0.00			-0.017
							[0.015]			[0.0497]
Longitude (log)								-0.20**		[-0.148]
								[0.081]		[0.218]
Latitude (log)									-0.01	1.29
									[0.378]	[1.030]
Constant	4.92***	5.07***	4.46***	4.36***	3.92***	4.29***	4.95***	5.66***	5.00***	2.61
	[0.050]	[0.158]	[0.300]	[0.473]	[0.885]	[0.343]	[0.064]	[0.315]	[1.486]	[4.915]
Observations	97	97	97	97	97	97	97	97	97	97
R-squared	0.003	0.007	0.013	0.014	0.012	0.042	0.000	0.045	0.000	0.082

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table B4. Labor productivity and institutions. Robustness

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	1897 labour productivity (log)					
Zemstvo dummy	0.05					
	[0.084]					
Oblast dummy		-0.08				
		[0.097]				
Share of cities with 1892 city charters			-0.13			
			[0.090]			
Number of years in the Russian empire (log)				0.01		
				[0.050]		
Share of orthodox					-0.02	
					[0.100]	
Religious diversity						0.13
						[0.195]
Constant	4.92***	4.96***	5.05***	4.91***	4.96***	4.91***
	[0.040]	[0.048]	[0.080]	[0.250]	[0.067]	[0.066]
Observations	97	97	96	97	97	89
R-squared	0.006	0.006	0.017	0.000	0.000	0.005

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table B5. Determinants of labor productivity in the Russian Empire: an extension

VARIABLES	(1)	(2)	(3)	(4)
	1897 labor productivity (log)			
	OLS	OLS	OLS	IV, 2nd stage
Black sea dummy	0.44***			
	[0.108]			
Baltic sea dummy	0.74***			
	[0.198]			
North seas dummy	-0.07			
	[0.096]			
Pacific ocean dummy	0.65***			
	[0.104]			
Minimum distance to ports (log)		-0.18***		
		[0.053]		
Share of employment in mining relative to registered workers			6.64**	
			[2.976]	
Sea dummy				0.33***
				[0.125]
Share of employment in mining				16.56*
				[9.401]
Share of serfs in 1858				-1.15***
				[0.395]
Log distance to Moscow				-0.34***
				[0.074]
Constant	4.19***	5.40***	4.24***	7.40***
	[0.042]	[0.344]	[0.052]	[0.586]
Observations	97	97	89	97
R-squared	0.335	0.189	0.060	0.318

Panel B: First stages of 2SLS regressions Dependent var:	Share of employment in mining	Share of serfs in 1858
Number of deposit fields per sq km	125.97***	-742.42
	[20.100]	[677.000]
Share of (former) monasterial serfs	0.01	-0.51
	[0.007]	[0.386]
Dummy for the presence of monasteries in 1764	0.00	0.20***
	[0.002]	[0.055]
Sea dummy	0.00	-0.05
	[0.002]	[0.055]
Log distance to Moscow	0.00	-0.14**
	[0.001]	[0.053]
Observations	97	97
R-squared	0.311	0.487
F-stat	39.28	4.515

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table B6. Determinants of labor productivity in the Russian Empire: Geography and Institutions. Robustness with additional geographical controls

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	1897 labor productivity (log)					
	OLS	OLS	OLS	OLS	IV, 2nd stage	IV, 2nd stage
Sea dummy	0.46***			0.40***	0.40***	0.38***
	[0.113]			[0.120]	[0.121]	[0.120]
Share of employment in mining		19.24***		18.00***	18.52***	18.32***
		[5.759]		[4.905]	[6.031]	[4.600]
Share of serfs in 1858			-0.67***	-0.54***	-0.54***	-0.74**
			[0.159]	[0.138]	[0.132]	[0.334]
LC_2sls 95 percent confidence interval						[-1.42625,-.056281]
Log distance to Moscow			-0.18***	-0.22***	-0.10	-0.25***
			[0.055]	[0.041]	[0.066]	[0.053]
Ruggedness (log)	-0.01	0.06	0.06	-0.06	-0.06	-0.07
	[0.092]	[0.081]	[0.086]	[0.078]	[0.072]	[0.075]
Log length of rivers	0.06	-0.04	-0.03	-0.04	-0.04	-0.05
	[0.079]	[0.104]	[0.110]	[0.073]	[0.074]	[0.070]
Log temerature in 1901-1910 (mean)	-0.15	0.36	-0.45	-0.02	-0.01	-0.14
	[0.400]	[0.527]	[0.449]	[0.422]	[0.382]	[0.472]
Log precipitation in 1901-1910 (mean)	0.11	0.24	0.21	0.30	0.30	0.36*
	[0.213]	[0.260]	[0.280]	[0.213]	[0.216]	[0.199]
Log relative humidity in 1901-1910 (mean)	-0.69	-0.71	-0.84	-0.42	-0.41	-0.49
	[0.724]	[0.792]	[0.758]	[0.584]	[0.539]	[0.562]
Log vapor preassure in 1901-1910 (mean)	0.57	0.65	1.10**	0.19	0.18	0.28
	[0.603]	[0.618]	[0.550]	[0.520]	[0.476]	[0.533]
Land suitability for cereals (log)	0.01	-0.03	-0.00	0.01	0.01	0.02
	[0.034]	[0.044]	[0.041]	[0.033]	[0.030]	[0.036]
Latitude (log)	0.38	1.34	0.51	-0.49	-0.49	-0.63
	[0.962]	[0.933]	[0.951]	[0.767]	[0.710]	[0.820]
Constant	5.39*	-0.56	6.36	8.57***	8.52***	9.57***
	[3.202]	[3.757]	[3.897]	[2.887]	[2.758]	[3.548]
Observations	97	97	97	97	97	97
R-squared	0.312	0.193	0.226	0.534	0.534	0.524
Panel B: First stages of 2SLS regressions						
Dependent var:					Share of employment in mining	Share of serfs in 1858
Number of deposit fields per sq km					136.87***	
					[20.476]	
Share of (former) monasterial serfs						-0.14
						[0.311]
Dummy for the presence of monasteries in 1764						0.26***
						[0.066]
Sea dummy					0.00	-0.03
					[0.002]	[0.062]
Share of employment in mining						3.00
						[3.009]
Share of serfs in 1858					0.00	
					[0.004]	
Log distance to Moscow					0.00	-0.10
					[0.001]	[0.066]
Geographical controls					Yes	Yes
Observations					97	97
R-squared					0.506	0.583
F-stat					44.68	7.627

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. For weak instruments, LC_2sls 95 percent confidence interval is computed using the Stata command twostepweakiv *** p<0.01, ** p<0.05, * p<0.1.

Table B7. Labor productivity and the legacy of serfdom. Robustness with additional geographical controls

	(1)	(2)	(3)	(4)	(5)
VARIABLES	1897 labor productivity (log)				
Sea dummy	0.41***	0.21	0.28***	0.24***	0.23**
	[0.119]	[0.149]	[0.101]	[0.069]	[0.092]
Share of employment in mining	17.65***	25.48***	14.58***	15.53***	15.72***
	[4.739]	[5.754]	[5.155]	[4.120]	[5.001]
Share of serfs in 1858	-0.33	-0.44**	-0.29**	-0.22**	-0.19
	[0.219]	[0.177]	[0.133]	[0.101]	[0.132]
Log distance to Moscow	-0.23***	-0.30***	-0.19***	-0.12***	-0.12***
	[0.041]	[0.035]	[0.027]	[0.030]	[0.039]
Share of large estates in 1858	-0.17				
	[0.134]				
Share of land with state tenure		-0.80**			
		[0.380]			
Share of land with commune tenure		-1.06**			
		[0.454]			
Share of in-migrants			2.97***		
			[0.861]		
Share of out-migrants			-0.55		
			[0.615]		
Share of urban populatin				1.57***	
				[0.195]	
Literacy					1.41***
					[0.322]
Ruggedness (log)	-0.05	0.22	-0.01	0.06	0.01
	[0.080]	[0.153]	[0.079]	[0.060]	[0.071]
Log length of rivers	0.00	-0.07	0.05	0.01	0.05
	[0.082]	[0.181]	[0.094]	[0.059]	[0.097]
Log temerature in 1901-1910 (mean)	-0.19	0.72	-0.03	-0.07	-0.53
	[0.467]	[3.308]	[0.455]	[0.322]	[0.383]
Log precipitation in 1901-1910 (mean)	0.17	-0.85*	0.12	0.03	-0.06
	[0.245]	[0.472]	[0.226]	[0.165]	[0.224]
Log relative humidity in 1901-1910 (mean)	-0.48	-0.36	-0.96	-0.32	-0.40
	[0.598]	[3.719]	[0.649]	[0.565]	[0.541]
Log vapor preassure in 1901-1910 (mean)	0.33	1.66	0.57	0.40	0.45
	[0.556]	[3.462]	[0.519]	[0.437]	[0.430]
Land suitability for cereals (log)	0.04	-0.17	0.00	0.02	0.05*
	[0.039]	[0.185]	[0.035]	[0.025]	[0.029]
Latitude (log)	-0.41	4.31	0.34	0.22	-1.37
	[0.779]	[2.769]	[1.052]	[0.693]	[0.866]
Constant	9.34***	-9.52	7.54*	5.27**	13.36***
	[3.071]	[16.049]	[4.116]	[2.332]	[3.726]
Observations	97	60	89	97	89
R-squared	0.547	0.735	0.639	0.718	0.660

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table B8. Labor productivity and the new economic geography. Robustness with additional geographical controls

	(1)	(2)	(3)	(4)	(5)	(6)
	1897 labor productivity (log)					
VARIABLES	OLS	IV, 2nd stage	OLS	OLS	OLS	OLS
Sea dummy	0.40***	0.39***	0.42***	0.20**	0.28***	0.31***
	[0.119]	[0.113]	[0.124]	[0.080]	[0.073]	[0.112]
Share of employment in mining	18.06***	17.97***	17.93***	14.92***	7.80*	15.58***
	[4.775]	[4.644]	[4.773]	[4.559]	[4.220]	[4.729]
Share of serfs in 1858	-0.53***	-0.54***	-0.52***	-0.26**	-0.21	-0.44***
	[0.136]	[0.137]	[0.137]	[0.110]	[0.143]	[0.150]
Log distance to Moscow	-0.23***	-0.22***	-0.21***	-0.12***	-0.04	-0.20***
	[0.042]	[0.039]	[0.058]	[0.030]	[0.048]	[0.032]
Market potentilal (log)	0.13	-0.06	-0.01	-0.09		
	[0.213]	[0.232]	[0.355]	[0.187]		
Density (log)			0.05			
			[0.082]			
Urban density (log)				0.26***		
				[0.035]		
Rural density (log)				-0.29***		
				[0.059]		
Specialization index for agriculture					0.35	
					[0.483]	
Specialization index for industry					1.02***	
					[0.240]	
Specialization index for services					-0.30	
					[0.185]	
Mean industrial firm revenue (log)						0.14***
						[0.043]
Ruggedness (log)	-0.03	-0.07	-0.05	0.07	0.02	-0.02
	[0.102]	[0.095]	[0.117]	[0.070]	[0.062]	[0.101]
Log length of rivers	-0.04	-0.04	-0.07	0.06	-0.01	0.15
	[0.072]	[0.070]	[0.089]	[0.071]	[0.066]	[0.103]
Log temerature in 1901-1910 (mean)	0.02	-0.03	0.09	-0.03	-0.17	-0.31
	[0.417]	[0.408]	[0.461]	[0.348]	[0.363]	[0.589]
Log precipitation in 1901-1910 (mean)	0.21	0.34	0.23	0.06	0.11	-0.01
	[0.271]	[0.285]	[0.277]	[0.199]	[0.169]	[0.250]
Log relative humidity in 1901-1910 (mean)	-0.27	-0.49	-0.01	-0.50	-0.27	-0.58
	[0.606]	[0.635]	[0.677]	[0.736]	[0.538]	[0.709]
Log vapor preassure in 1901-1910 (mean)	0.05	0.27	-0.15	0.60	0.50	0.80
	[0.524]	[0.562]	[0.631]	[0.584]	[0.470]	[0.631]
Land suitability for cereals (log)	0.01	0.01	0.00	0.02	0.02	-0.16*
	[0.033]	[0.031]	[0.044]	[0.027]	[0.029]	[0.095]
Latitude (log)	-0.50	-0.49	-0.66	0.64	-0.08	-0.65
	[0.748]	[0.726]	[0.741]	[0.790]	[0.669]	[1.177]
Constant	7.15**	9.27**	7.94**	5.74*	5.38**	10.08*
	[3.571]	[3.896]	[3.781]	[3.013]	[2.351]	[5.253]
Observations	97	97	97	97	97	79
R-squared	0.537	0.531	0.541	0.730	0.713	0.643
Panel B: First stages of 2SLS regressions						
Dependent var:		Market potentilal (log)				
Sums of reverse distances		0.72***				
		[0.087]				
Sea dummy		-0.02				
		[0.035]				
Share of employment in mining		0.91				
		[1.784]				
Share of serfs in 1858		-0.09				
		[0.091]				
Log distance to Moscow		0.10**				
		[0.042]				
Geographical controls		Yes				
Observations		97				
R-squared		0.889				
F-stat		68.07				

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table B9. Determinants of labor productivity in the Russian Empire. Estimations standard errors corrected for spatial correlation

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	1897 labor productivity (log)					
	OLS	OLS	OLS	OLS	IV, 2nd stage	IV, 2nd stage
Sea dummy	0.45***			0.40***	0.23**	0.23***
	[0.089]			[0.091]	[0.102]	[0.086]
Share of employment in mining		13.93**		13.23***	20.00**	9.45***
		[6.118]		[3.785]	[8.048]	[3.418]
Share of serfs in 1858			-0.59***	-0.44***	-0.59***	-1.08***
			[0.206]	[0.151]	[0.160]	[0.400]
Log distance to Moscow			-0.18***	-0.22***	-0.28***	-0.33***
			[0.044]	[0.029]	[0.068]	[0.087]
Constant	4.87***	4.91***	6.34***	6.41***	6.89***	7.34***
	[0.030]	[0.040]	[0.343]	[0.208]	[0.471]	[0.665]
Observations	97	97	97	97	97	97
R-squared	0.262	0.080	0.152	0.467	0.213	0.233
Panel B: First stages of 2SLS regressions						
Dependent var:					Share of employment in mining	Share of serfs in 1858
Number of deposit fields per sq km					108.36***	
					[37.845]	
Share of (former) monasterial serfs						-0.16
						[0.297]
Dummy for the presence of monasteries in 1764						0.18***
						[0.049]
Sea dummy					0.00	-0.06
					[0.002]	[0.046]
Share of employment in mining						1.15
						[1.177]
Share of serfs in 1858					0.01	
					[0.005]	
Log distance to Moscow					0.01*	-0.15***
					[0.003]	[0.025]
Observations					97	97
R-squared					0.251	0.571
F-stat					8.20	13.41

Notes: Columns (1) to (4) are estimated using the Stata command *x_ols*, and (5) and (6) using Stata command *acreg*. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table B10. Labor productivity (alternative estimates) and the legacy of serfdom. Estimations standard errors corrected for spatial correlation

	(1)	(2)	(3)	(4)	(5)
VARIABLES	1897 labor productivity (log)				
Sea dummy	0.42***	0.41***	0.31***	0.25***	0.30***
	[0.090]	[0.080]	[0.082]	[0.062]	[0.082]
Share of employment in mining	14.06***	19.92***	10.54***	13.59***	13.28***
	[4.009]	[4.220]	[3.597]	[3.204]	[4.173]
Share of serfs in 1858	-0.20	-0.63***	-0.27	-0.17	-0.28
	[0.183]	[0.118]	[0.184]	[0.126]	[0.177]
Log distance to Moscow	-0.23***	-0.26***	-0.21***	-0.11***	-0.11***
	[0.025]	[0.027]	[0.031]	[0.023]	[0.043]
Share of large estates in 1858	-0.21**				
	[0.102]				
Share of land with state tenure		-0.69***			
		[0.160]			
Share of land with commune tenure		-1.06***			
		[0.182]			
Share of in-migrants			2.52**		
			[0.989]		
Share of out-migrants			-0.84		
			[0.710]		
Share of urban populatin				1.58***	
				[0.186]	
Literacy					0.88***
					[0.298]
Constant	6.60***	7.34***	6.32***	5.42***	5.50***
	[0.177]	[0.203]	[0.291]	[0.186]	[0.374]
Observations	97	60	89	97	89
R-squared	0.496	0.671	0.571	0.684	0.554

Notes: Columns (1) to (5) are estimated using the Stata command *x_ols*. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table B11. Labor productivity (alternative estimates) and the new economic geography. Estimations standard errors corrected for spatial correlation

	(1)	(2)	(3)	(4)	(5)	(6)
	1897 labor productivity (log)					
VARIABLES	OLS	IV, 2nd stage	OLS	OLS	OLS	OLS
Sea dummy	0.39***	0.28***	0.42***	0.25***	0.31***	0.37***
	[0.087]	[0.073]	[0.078]	[0.052]	[0.075]	[0.092]
Share of employment in mining	14.80***	9.27***	15.33***	15.17***	5.07*	13.11***
	[3.898]	[3.324]	[3.511]	[2.897]	[3.049]	[2.889]
Share of serfs in 1858	-0.48***	-0.46***	-0.48***	-0.24**	-0.12	-0.45***
	[0.123]	[0.099]	[0.125]	[0.097]	[0.210]	[0.152]
Log distance to Moscow	-0.20***	-0.16***	-0.17***	-0.12***	-0.04	-0.17***
	[0.037]	[0.057]	[0.043]	[0.019]	[0.043]	[0.021]
Market potentilal (log)	0.21**	0.23*	-0.03	-0.08		
	[0.086]	[0.121]	[0.170]	[0.080]		
Density (log)			0.08*			
			[0.048]			
Urban density (log)				0.25***		
				[0.027]		
Rural density (log)				-0.23***		
				[0.050]		
Land suitability for cereals (log)				0.02**		
				[0.009]		
Specialization index for agriculture					0.64*	
					[0.373]	
Specialization index for industry					1.10***	
					[0.181]	
Specialization index for services					-0.42***	
					[0.142]	
Mean industrial firm revenue (log)						0.15***
						[0.038]
Constant	4.11***	3.59**	6.17***	6.95***	4.85***	4.44***
	[1.017]	[1.520]	[1.683]	[0.800]	[0.372]	[0.470]
Observations	97	97	97	97	97	79
R-squared	0.491	0.353	0.515	0.712	0.672	0.564
Panel B: First stages of 2SLS regressions						
Dependent var:		Market potentilal (log)				
Sums of reverse distances		0.82***				
		[0.117]				
Sea dummy		-0.01				
		[0.047]				
Share of employment in mining		3.50**				
		[1.567]				
Share of serfs in 1858		-0.00				
		[0.132]				
Log distance to Moscow		0.10**				
		[0.047]				
Observations		97				
R-squared		0.947				
F-stat		48.92				

Notes: Columns (1), (3) to (6) are estimated using the Stata command *x_ols*, and (2) using Stata command *acreg*. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table B12. Determinants of labor productivity in the Russian Empire: Geography and Institutions. Alternative estimates of labor productivity

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	1897 labor productivity (log)					
	OLS	OLS	OLS	OLS	IV, 2nd stage	IV, 2nd stage
Sea dummy	0.48***			0.44***	0.41***	0.37***
	[0.112]			[0.115]	[0.134]	[0.132]
Share of employment in mining		14.10**		13.38**	23.46***	14.33***
		[6.366]		[5.184]	[7.838]	[5.113]
Share of serfs in 1858			-0.65***	-0.48***	-0.49***	-1.26***
			[0.148]	[0.126]	[0.121]	[0.453]
LC_2sls 95 percent confidence interval						[-2.82053,-.496487]
Log distance to Moscow			-0.21***	-0.24***	-0.26***	-0.36***
			[0.044]	[0.038]	[0.034]	[0.083]
Constant	4.66***	4.71***	6.30***	6.35***	6.48***	7.41***
	[0.040]	[0.048]	[0.336]	[0.274]	[0.243]	[0.667]
Observations	97	97	97	97	97	97
R-squared	0.262	0.070	0.159	0.467	0.435	0.316
Panel B: First stages of 2SLS regressions						
Dependent var:					Share of employment in mining	Share of serfs in 1858
Number of deposit fields per sq km					124.53***	
					[19.291]	
Share of (former) monasterial serfs						-0.46
						[0.367]
Dummy for the presence of monasteries in 1764						0.20***
						[0.057]
Sea dummy					0.00	-0.05
					[0.002]	[0.055]
Share of employment in mining						0.00
						[2.382]
Share of serfs in 1858					0.00	
					[0.003]	
Log distance to Moscow					0.00	-0.14**
					[0.001]	[0.056]
Observations					97	97
R-squared					0.302	0.481
F-stat					41.67	5.964

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. For weak instruments, LC_2sls 95 percent confidence interval is computed using the Stata command twostepweakiv *** p<0.01, ** p<0.05, * p<0.1.

Table B13. Labor productivity (alternative estimates) and the legacy of serfdom

	(1)	(2)	(3)	(4)	(5)
VARIABLES	1897 labor productivity (log)				
Sea dummy	-0.26***	-0.28***	-0.23***	-0.12***	-0.13***
	[0.034]	[0.034]	[0.031]	[0.029]	[0.046]
Share of employment in mining	14.24***	20.79***	10.72**	13.78***	13.44***
	[5.255]	[4.956]	[5.238]	[4.228]	[4.944]
Share of serfs in 1858	-0.23	-0.69***	-0.30**	-0.19**	-0.31**
	[0.160]	[0.158]	[0.115]	[0.086]	[0.128]
Log distance to Moscow	0.45***	0.44***	0.34***	0.27***	0.33***
	[0.117]	[0.118]	[0.107]	[0.072]	[0.102]
Share of large estates in 1858	-0.22**				
	[0.106]				
Share of land with state tenure		-1.12***			
		[0.352]			
Share of land with commune tenure		-0.73***			
		[0.213]			
Share of in-migrants			2.51**		
			[0.991]		
Share of out-migrants			-0.80		
			[0.675]		
Share of urban populatin				1.72***	
				[0.218]	
Literacy					0.90***
					[0.326]
Constant	6.55***	7.31***	6.24***	5.28***	5.42***
	[0.252]	[0.373]	[0.261]	[0.214]	[0.374]
Observations	97	60	89	97	89
R-squared	0.494	0.682	0.550	0.686	0.539

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table B14. Labor productivity (alternative estimates) and the new economic geography

	(1)	(2)	(3)	(4)	(5)	(6)
			1897 labor productivity (log)			
VARIABLES	OLS	IV, 2nd stage	OLS	OLS	OLS	OLS
Sea dummy	0.43***	0.43***	0.46***	0.29***	0.35***	0.38***
	[0.112]	[0.108]	[0.107]	[0.069]	[0.080]	[0.105]
Share of employment in mining	15.19***	15.45***	15.92***	15.41***	4.49	13.43***
	[4.803]	[4.607]	[4.187]	[3.920]	[3.353]	[4.741]
Share of serfs in 1858	-0.53***	-0.54***	-0.53***	-0.27***	-0.13	-0.49***
	[0.124]	[0.120]	[0.132]	[0.096]	[0.127]	[0.128]
Log distance to Moscow	-0.22***	-0.22***	-0.18***	-0.12***	-0.04	-0.18***
	[0.046]	[0.046]	[0.051]	[0.024]	[0.044]	[0.037]
Market potentilal (log)	0.24**	0.27**	-0.09	-0.16		
	[0.102]	[0.119]	[0.166]	[0.117]		
Density (log)			0.11**			
			[0.044]			
Urban density (log)				0.27***		
				[0.031]		
Rural density (log)				-0.21***		
				[0.048]		
Land suitability for cereals (log)				0.01		
				[0.014]		
Specialization index for agriculture					0.61	
					[0.396]	
Specialization index for industry					1.20***	
					[0.238]	
Specialization index for services					-0.47**	
					[0.190]	
Mean industrial firm revenue (log)						0.17***
						[0.051]
Constant	3.72***	3.34**	6.57***	7.60***	4.67***	4.13***
	[1.172]	[1.363]	[1.683]	[1.168]	[0.320]	[0.679]
Observations	97	97	97	97	97	79
R-squared	0.495	0.494	0.534	0.720	0.671	0.562
Panel B: First stages of 2SLS regressions						
Dependent var:		Market potentilal (log)				
Sums of reverse distances		0.88***				
		[0.063]				
Sea dummy		-0.01				
		[0.032]				
Share of employment in mining		1.79				
		[2.201]				
Share of serfs in 1858		-0.02				
		[0.086]				
Log distance to Moscow		0.10***				
		[0.037]				
Observations		97				
R-squared		0.836				
F-stat		195.6				

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table B15 Determinants of labor productivity (levels) in the Russian Empire: Geography and Institutions

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	1897 labor productivity					
	OLS	OLS	OLS	OLS	IV, 2nd stage	IV, 2nd stage
Sea dummy	78.91***			72.11***	65.15**	63.07**
	[22.184]			[22.981]	[26.441]	[25.928]
Share of employment in mining		2,365.56*		2,315.12**	4,298.82***	2,429.10**
		[1,215.511]		[981.051]	[1,515.182]	[947.788]
Share of serfs in 1858			-110.58***	-82.74***	-85.82***	-176.00**
			[26.011]	[22.820]	[21.457]	[74.054]
LC_2sls 95 percent confidence interval						[-402.855,-36.9623]
Log distance to Moscow			-36.15***	-41.93***	-46.04***	-56.90***
			[8.966]	[7.970]	[7.034]	[12.412]
Constant	135.80***	143.60***	422.18***	433.16***	457.09***	560.14***
	[6.510]	[8.213]	[68.061]	[58.512]	[50.402]	[101.691]
Observations	97	97	97	97	97	97
R-squared	0.255	0.072	0.175	0.485	0.439	0.405
Panel B: First stages of 2SLS regressions						
Dependent var:					Share of employment in mining	Share of serfs in 1858
Number of deposit fields per sq km					124.53***	
					[19.291]	
Share of (former) monasterial serfs						-0.46
						[0.367]
Dummy for the presence of monasteries in 1764						0.20***
						[0.057]
Sea dummy					0.00	-0.05
					[0.002]	[0.055]
Share of employment in mining						0.00
						[2.382]
Share of serfs in 1858					0.00	
					[0.003]	
Log distance to Moscow					0.00	-0.14**
					[0.001]	[0.056]
Observations					97	97
R-squared					0.302	0.481
F-stat					41.67	5.964

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. For weak instruments, LC_2sls 95 percent confidence interval is computed using the Stata command twostepweakiv *** p<0.01, ** p<0.05, * p<0.1.

Table B16. Labor productivity (levels) and the legacy of serfdom

	(1)	(2)	(3)	(4)	(5)
VARIABLES	1897 labor productivity				
Sea dummy	73.54***	76.49***	53.36***	41.11***	53.25***
	[23.091]	[25.368]	[19.121]	[13.266]	[18.141]
Share of employment in mining	2,426.89**	3,520.48***	1,780.43*	2,387.62***	2,326.68**
	[990.048]	[1,090.447]	[974.491]	[797.784]	[952.556]
Share of serfs in 1858	-50.09*	-117.60***	-49.86***	-29.91**	-54.30**
	[27.846]	[29.940]	[18.073]	[12.052]	[22.005]
Log distance to Moscow	-44.18***	-51.34***	-41.57***	-20.32***	-22.43**
	[7.537]	[7.288]	[5.024]	[4.178]	[9.263]
Share of large estates in 1858	-28.03				
	[17.650]				
Share of land with state tenure		-123.79***			
		[37.353]			
Share of land with commune tenure		-191.91***			
		[65.695]			
Share of in-migrants			496.95**		
			[192.324]		
Share of out-migrants			-179.58*		
			[90.462]		
Share of urban populatin				314.56***	
				[38.636]	
Literacy					168.19**
					[67.100]
Constant	458.55***	608.62***	421.35***	236.62***	261.61***
	[56.163]	[76.606]	[42.758]	[31.031]	[75.696]
Observations	97	60	89	97	89
R-squared	0.501	0.679	0.619	0.752	0.588

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table B17. Labor productivity (levels) and the new economic geography

	(1)	(2)	(3)	(4)	(5)	(6)
			1897 labor productivity			
VARIABLES	OLS	IV, 2nd stage	OLS	OLS	OLS	OLS
Sea dummy	70.34***	70.39***	75.28***	43.47***	51.49***	66.22***
	[22.813]	[22.014]	[22.773]	[13.831]	[13.049]	[20.161]
Share of employment in mining	2,554.30***	2,547.35***	2,654.32***	2,599.19***	820.97	2,255.25**
	[940.356]	[918.061]	[875.087]	[841.082]	[604.423]	[942.037]
Share of serfs in 1858	-89.42***	-89.22***	-89.49***	-43.44**	-19.45	-86.57***
	[22.991]	[21.869]	[23.958]	[16.795]	[18.754]	[22.397]
Log distance to Moscow	-39.51***	-39.58***	-34.27***	-23.79***	-5.87	-32.78***
	[9.288]	[9.137]	[9.906]	[3.826]	[6.388]	[7.273]
Market potentilal without own potential (log)	31.59*	30.68	-12.91	-20.93		
	[18.054]	[20.921]	[26.615]	[17.054]		
Density (log)			15.24**			
			[6.980]			
Urban density (log)				46.75***		
				[7.011]		
Rural density (log)				-44.34***		
				[9.430]		
Land suitability for cereals (log)				3.80**		
				[1.814]		
Specialization index for agriculture					90.05	
					[59.983]	
Specialization index for industry					202.04***	
					[39.798]	
Specialization index for services					-52.98	
					[33.312]	
Mean industrial firm revenue (log)						27.15***
						[9.208]
Constant	83.23	93.39	472.04*	601.45***	125.34***	75.68
	[215.522]	[249.634]	[281.852]	[175.633]	[46.147]	[119.619]
Observations	97	97	97	97	97	79
R-squared	0.502	0.502	0.528	0.748	0.760	0.588
Panel B: First stages of 2SLS regressions						
Dependent var:		Market potentilal (log)				
Sums of reverse distances		0.88***				
		[0.063]				
Sea dummy		-0.01				
		[0.032]				
Share of employment in mining		1.79				
		[2.201]				
Share of serfs in 1858		-0.02				
		[0.086]				
Log distance to Moscow		0.10***				
		[0.037]				
Observations		97				
R-squared		0.836				
F-stat		195.6				

Notes: Observations are weighted by provincial population in the working age. Robust standard errors are in brackets. *** p<0.01, ** p<0.05, * p<0.1.