European
Historical
Economics
Society

EHES Working Papers in Economic History | No. 10

Landownership Concentration and the Expansion of Education

Francesco Cinnirella,
Ifo Institute and CESifo, Munich

Erik Hornung,
Ifo Institute, Munich

October 2011

EHES Working Paper | No. 10 | October 2011

Landownership Concentration and the Expansion of Education

Francesco Cinnirella,
Ifo Institute and CESifo, Munich

Erik Hornung,
Ifo Institute, Munich

Abstract

This paper studies the effect of landownership concentration on school enrollment for nineteenth-century Prussia. Prussia is an interesting laboratory given its decentralized educational system and the presence of heterogeneous agricultural institutions. We find that landownership concentration, a proxy for the institution of serf labor, has a negative effect on schooling. This effect diminishes substantially in the second half of the century. Causality of this relationship is confirmed by introducing soil-texture to identify exogenous farm size variation. Panel estimates further rule out unobserved heterogeneity. We argue that serfdom hampered peasants' demand for education whereas the successive emancipation triggered a demand thereof.

JEL Codes: O43, Q15, I25, N33

Keywords: Land concentration, Institutions, Serfdom, Education, Prussian economic history

Acknowledgements:
Comments from Sascha O. Becker, Davide Cantoni, Greg Clark, Oded Galor, Mike Hoy, Florian Ploeckl, Ludger Woessmann and seminar participants at the 23rd Conference of the European Association of Labour Economists in Cyprus, the Annual Conference of the German Economic Association in Frankfurt/Main, the European Historical Economics Society Conference in Dublin, the Annual Congress of the European Economic Association in Oslo, the Annual Economic History Society Conference in Cambridge, the Berlin Colloquium at the Humboldt-Universität in Berlin, the FRESH meeting in Madrid, the University of Copenhagen, the Sound Economic History Workshop in Lund, and the Ifo Institute in Munich, are gratefully acknowledged. Financial support by the Pact for Research and Innovation of the Leibniz Association is also gratefully acknowledged.

1 Introduction

An expanding literature investigates the effects of different historical institutions on current economic outcomes (Engerman and Sokoloff, 1997; Acemoglu et al., 2001, 2002). Places which developed extractive institutions are generally characterized by the concentration of political power in the hands of a small elite and by a vast majority of population without effective rights or protection. Cross-country studies show that those regions tended to stagnate in terms of economic growth. On the contrary, areas with more democratic institutions took advantage of economic opportunities and were more conducive to economic growth (Engerman et al., 2002).

As formalized in the theoretical model of Galor et al. (2009), an unequal distribution of land among the population might delay the implementation of educational institutions and indirectly affect growth. The delay in the expansion of formal education is caused by large landowners who oppose educational reforms to reduce the mobility of the rural labor force. Similarly, Engerman and Sokoloff (1997) and Sokoloff and Engerman (2000) suggest that the elite in Latin America opposed democracy and mass investment in human capital because they were afraid of the poor majority gaining power.

In this paper, we explore the relationship between landownership concentration—a proxy for the extractive institution of serf labor—and the expansion of education in nineteenth-century Prussia. We construct a unique dataset spanning the whole nineteenth-century which includes, among other variables, primary school enrollment rates and the distribution of landownership by size. The literature generally measures landownership inequality with the Gini coefficient. We define landownership concentration as the share of large landholdings in a county. We argue that our variable is better suited to capture the "extent" of serf labor in Prussia.[1]

Cross-sectional estimates show a significant negative effect of landownership concentration on primary school enrollment rates. We show that this effect is robust to the introduction of several demand and supply factors. In particular, we find that this negative effect is not due to a lower supply of schools or teachers.

We estimate the effect of landownership concentration on education for five different years throughout the nineteenth-century. That allows to study how the effect changes over time. We find that the effect of landownership concentration decreases substantially in the second half of the nineteenth-century. We suggest that (i) the condition of limited freedom that characterized the peasantry under serfdom in the

[1]See section 3 for a detailed account of the distribution of power in Prussia. However, we will provide also alternative estimates of the relationship between landownership inequality and education using the Gini coefficient.

first half of the nineteenth-century and (*ii*) the successive peasants' emancipation can explain the changes in the effect of landownership concentration on education throughout the century.

In fact, peasants in eastern Europe had to provide serf labor to their landlord for a few days a week (corvée), they fell under the landlord's jurisdiction, and above all they were not allowed to abandon the village without the landlord's permission (Bowman, 1980). We argue that all these restrictions hindered peasants' *demand* for education. It was only with the complete emancipation of the peasantry in the second half of the nineteenth-century that the *demand* for education increased also in the eastern areas of Prussia, catching up with the high levels of education of the western regions.

In order to establish whether the negative effect of landownership concentration on education is causal, we adopt an instrumental variables approach. Similar to Easterly (2007) who uses land suitability for particular crops to identify exogenous variation in inequality, we use soil characteristics at the county level to identify exogenous variation in farm size and thus in landownership concentration (see also Frankema (2010)). In particular, we use county level data reporting the composition of the soil in terms of clay, loam and sand. The instrument is based on an empirical regularity extensively investigated in agricultural economics according to which there is a systematic negative correlation between soil quality and farm size (Bhalla and Roy, 1988; Bhalla, 1988; Benjamin, 1995). Regions with higher soil quality historically experienced a stronger demand for land which determined a more accentuated land fragmentation, whereas the opposite happened to the less fertile areas of north-eastern Europe (Boserup, 1965). The assumption behind our identification strategy is that soil texture has an effect on education only through farm size. To ensure that the exclusion restriction is not violated, we introduce also controls for land productivity in different crops. Instrumental variables estimates suggest that the negative effect of land concentration on education is indeed causal.

We are also able to organize our data as a panel. That allows to estimate models with county fixed effects which can rule out unobserved heterogeneity. Panel estimates confirm the results obtained through the cross-sectional analysis: Within-county variation in landownership concentration is negatively associated with changes in enrollment rates.

Prussia provides an interesting case since it is representative of the different agricultural institutions that characterized eastern and western parts of Europe. Historically, western parts of Europe were comparatively more densely populated.

Yet, during the Middle Ages, the Black Death and other plagues depopulated these regions. Thus serf labor became more valuable to the lords and peasants achieved more bargaining power resulting in more freedom and the progressive abolishment of the manorial system. On the other hand, eastern European regions, including Prussian regions east of the river Elbe, were more sparsely populated. Here, depopulation (also due to the Thirty Years' War) and a larger land-labor ratio led to opposite results. The lords tried to control peasants' mobility by confiscating their land and bonding them to the manorial demesnes (Brenner, 1976; Domar, 1970). In addition, the increasing demand for grain from western Europe during the seventeenth-century accentuated the use of serf labor in the East (Rosenberg, 1944, p. 233).

Due to these developments the East became characterized by extensive agriculture (grain crops) which led to an increasing concentration of large landholdings, while large manorial farms disappeared in the West leading to a comparatively equal distribution of land. Still, many regions were characterized by transitional stages. Thus Prussia, presenting such heterogeneity, makes a perfect laboratory to study the effect of different institutions on education. Our variable for landownership concentration aims to capture such variation in local institutions.

In section 2 we review the related literature; in section 3 we provide the historical background emphasizing the Prussian reforms in agriculture and education; in section 4 we describe the data; in section 5 we introduce the model and present OLS estimates; in section 6 we address the issue of causality and show instrumental variable estimates; in section 7 we present robustness checks, including estimates of panel models; finally, in section 8 we conclude.

2 Related literature

The literature on the long-run economic consequences of inequality on human capital is vast. Galor and Zeira (1993) show that in the presence of credit market imperfections, income distribution has a long-lasting effect on investment in human capital. Other scholars stress the redistributive channel. In particular, Alesina and Rodrik (1994) and Persson and Tabellini (1994) hypothesize that in a more egalitarian society taxation of physical capital and of human capital is lower, enhancing economic growth.[2]

[2]For an exhaustive review of the literature see Galor (2011).

Similarly, Ramcharan (2010) studies the effect of economic inequality on redistributive policy. The author studies the relationship between land inequality (measured by the Gini coefficient) and education expenditure using census data at the county level in the U.S. for the period 1890-1930. The paper shows that greater inequality is strongly associated with less redistribution and therefore with less expenditure in education. Interesting for our case, the author shows that using a measure for land concentration similar to what we use, the effect on per capita education expenditure is about four times as large as the impact estimated using the Gini coefficient (Ramcharan, 2010, p. 739). Ramcharan (2010) identifies exogenous variation in land inequality using geographic variables such as surface elevation, rainfall, and crop choice.[3]

Galor et al. (2009) investigate the negative relationship between inequality in landownership and the accumulation of human capital. Their theoretical model shows that, due to a low degree of complementarity between human capital and the agricultural sector, large landowners opposed the implementation of human-capital promoting institutions such as public schooling.[4] They test the prediction of their model using variation in the distribution of landownership and educational expenditure across states and over time in the U.S. for the period 1900-1940. Consistently with their theoretical predictions, they find that higher land inequality has a negative effect on education expenditures. In order to ensure that the effect is causal, they instrument landownership inequality through the interaction between nationwide changes in the relative price of agricultural crops that are associated with economies of scale and variation in climatic characteristics across states.

Go and Lindert (2010) explain differences in enrollment rates between the North and the South analyzing US counties in 1850. They point to local governments having more autonomy and the population having more equally distributed political voice in the North. Among other things, they test if extending the voting power to lower income groups raises the taxes paid for schooling and thus the enrollment rates in primary schools. Similarly, Vollrath (2010) finds that landownership inequality, measured by the Gini coefficient, predicts taxes for local school funding at the US county level in 1890.

Other studies that attempt to establish the effect of political or economic inequality on education rest primarily on cross-country differences. Mariscal and Sokoloff

[3]For the long-run effects of property rights institutions on health and educational outcome you are referred to Banerjee and Iyer (2005). Dell (2010), instead, analyzes the persistent effect on human capital of coercive labor institutions.

[4]In a similar fashion, Galor and Moav (2006) suggest that the complementarity in production between physical capital and human capital created an incentive for the capitalists to support the provision of public education.

(2000) show how the extension of the franchise in Latin America increased schooling by estimating the correlation between schooling and literacy rates with inequality in political power. The relationship between the extension of the franchise and schooling has also been analyzed by Acemoglu and Robinson (2000). Gallego (2010) explores the role of historical variables and political institutions to explain differences in schooling in former colonies. He argues that the degree of democratization positively affects primary education, whereas decentralization of political power is more related to differences in higher levels of schooling, such as secondary and higher education.

Easterly (2007) tests the hypothesis of Engerman and Sokoloff (1997) using agricultural endowments as an instrument for inequality. He finds that inequality, measured with the Gini coefficient and as the share of income accruing to the top quintile, has a significant negative causal effect on per capita income (in 2002), institutions, and secondary school enrollment rates for the period 1998-2002 (Easterly, 2007, p. 766, Table 4).

Our paper contributes to the literature by establishing the causal effect of the concentration of large landholdings on primary school enrollment rate at the county level for a single country. This is, to our knowledge, the first study that explores this relationship for a country in Europe. Most importantly, our dataset allows to study how this relationship evolves throughout the nineteenth-century, a period characterized by important institutional changes.

3 Historical background

3.1 Landownership and land reform

Prussia's agricultural systems were basically demarcated by the river Elbe. Agriculture in the eastern areas was dominated by large landholdings and the *Gutsherrschaft* system. The noble landowners (also known as Junkers) managed large demesnes, while they leased the rest of their land to extract rents in form of servile duties and fees. On the contrary, in western areas the *Grundherrschaft* system, characterized by smaller landholdings, was more prevalent: noble landowners relied on cash rents and leased most of the land to the peasants. Yet, this division was not unambiguous and the characteristics varied strongly within these areas (Rosenberg, 1944). The amount of servile dues and services usually depended on the attributes of the contracted land and were the heaviest in regions with large noble demesnes (Melton, 1988, p 149). Most of the peasants were serfs not allowed to relocate without compensation payment and the permission of the lord. Becoming owner of the land they cultivated was rarely possible. In fact, the landowner had control over jurisdiction and policing and was patron over church and schools (Bowman, 1980, p. 787). Appointing school teachers was also one of his duties. In return, landowners had the obligation to protect their subordinates and to support them in case of accidents.

In the eastern parts of Prussia, serfdom was abolished for the first time towards the end of the eighteenth-century on the royal domains. Eventually, the "Oktoberedikt" of 1807 made all peasants free people starting from November 11 (St. Martin's Day) 1810. In 1811 the reformer Hardenberg passed a regulation act on landownership according to which all peasants gained the right to own landed property of any type. Yet, the amount of compensation for land and foregone servile duties was not clarified for some time. Thus the reform did not gain momentum, also due to the war on Napoleon, until a discharge act in 1821.

According to the act, peasants with the right to inherit the lease could compensate by conveying one third of the land to the landlord while peasants on normal lease had to convey up to half of the land in the eastern parts of Prussia. Because of this practice and due to the dissolution of the common lands, average farm size of the large landholders rose strongly until the mid of the nineteenth-century (Schiller, 2003).[5] Exempted from the right to redeem the land were farmers on small parcels who remained obliged to demesne services (this limitation reduced the eligible peasants by roughly two thirds). In this way, the Junkers were still able to draw from a large labor force (Harnisch, 1984). It was only with the Commutation Law in 1850 that all peasants ultimately gained complete legal emancipation and were able to redeem small parcels (Pierenkemper and Tilly, 2004; Bowman, 1980).

During the Napoleonic occupation between 1807 and 1814, the western parts, which belonged to Prussia after 1815, introduced the Napoleonic Code which basically meant the abolishment of feudal privileges. Serfdom was therefore abolished *de jure*, in the Rhineland in 1805 and in Westphalia in 1808. Redemption of land was based on the payment of 15 times the annual value of the servile duties in Rhineland, and 25 times the annual value in Westphalia. In fact, Westphalian nobles blocked many of the changes introduced by the Napoleonic Code. While the Code remained effective after 1815 in the areas left of the river Rhine, all other areas (re-)established Prussian law (*Allgemeines Landrecht*).[6]

[5] Many small farmers had to give up their land completely or where bought up by the large landowners.

[6] For more details about the regional adoption of reforms you are referred to Acemoglu et al. (2011) and the corresponding online appendix.

3.2 The political power of the landowners

The political power on each level of Prussian administration was possessed by large landowners, most of them noble and conservative.[7] At the local level, the ordinary landowning patron was responsible for the villages in his domains. As mentioned above, the appointment of teachers was a responsibility of patrons and, according to Neugebauer (1992, p.684), they were free to select teachers until 1900.[8]

At the county level, which had been introduced by the administrative reform of 1815, policing and administration was executed by the county chief (*Landrat*). He was nominated by the owners of all knight estates, appointed by the King and had to be owner of a knight estate located within county borders. Therefore the *Landrat* represented the interests of the large landowners in the county assembly.[9]

Furthermore, each owner of a knight estate was also entitled to vote in the district and the provincial assembly (*Provinziallandtag*), both introduced in 1824. Half of the delegates represented the owners of knight estates called the First Estate, one third represented the civic landowners, and one sixth of the votes were held by peasants with large farms. These assemblies were advisors regarding property, personal and tax laws and had legislative power at the municipal level.

Despite the progressive reforms of the Stein-Hardenberg era, many aspects of the political power of landowners remained untouched throughout the nineteenth-century. Finally, in 1872, a new order disposed the power of knight estates over rural parishes and police and rearranged the composition of the county councils. From that point onwards, large landowners were represented by only two fifth of the delegates; free rural parishes held the same share of votes and towns accounted for one fifth of the delegates.

3.3 Education reform

It was not before 1806 that mass education became an increasingly relevant issue in Prussia. In fact, the King's previous edicts for compulsory education were attempts to persuade local lords to provide education at their own expenses (Lindert, 2003, 2004). The war against France triggered a "decade of feverish activity" (Schleunes, 1979, p. 317) during which the expansion of schooling rushed. Wilhelm von Humboldt, though head of cultural and educational affairs for only one year, is thought of having laid down a vision with essential impact on the reform. He favored a general

[7] See Webb (1982) on how large landlords used their political power to introduce protective tariffs in Prussia.

[8] Note that teachers were only selected by noble landowners; the payment of teachers fell on the entire community.

[9] According to Wagner (2001) the position of a *Landrat* was often filled by landowners with rather modest fortunes, who depended on the additional salary.

education over specialization and advocated a unified schooling system which would have educated everybody equally.[10]

Lindert (2004) reinterprets Prussian history of schooling as a bottom-up rather than a top-down story. He claims that schooling had never been of special national interest and was left to the local leaders. In fact, the expansion of elementary education was pushed back by conservative feudal interest. Already in 1763, when Prussia first introduced compulsory schooling laws, the country squire's opposition to mass education became specific and obvious. Especially among the Junkers of Pomerania, themselves often illiterate, local opposition against education prevailed. More educated peasants were said to run off to the cities or at least to use their abilities to shirk.[11] Literacy was seen as a potential weapon which might be used by peasants to appeal to royal courts in order to reduce their servile duties (Melton, 1988).

These duties varied regionally and ranged from at least two days a week to daily work. In parts of Pomerania, for example, peasant families had to provide two servants, a boy, and four horses to work on the lord's demesne on a daily basis. These duties remained partially effective even after the agricultural reform. In 1848 peasants claimed that they still had no personal freedom, since it did not matter if they had to fulfill the duties as serf-tenants or as owners of the land (Carsten, 1988, p 98). Duties effectively concerned the children at school-age through the institution of menial service (*Gesindezwang*) which forced them to work as servants in the Junkers' household. In the provinces of Pomerania and Silesia, Junkers drafted the most promising and capable children of the peasantry to work as fulltime servants (Carsten, 1988, p 63). Additionally, children were also burdened with the corvée (here: *Scharrwerk*) which, as the parish priest of Dombrowken (Eastprussia) complained in 1773, was so much work that children were kept from attending school even in the winter (Carsten, 1988, p 65).

In the countryside education remained closely connected to the church, where teachers were often clergymen, most of the teachings were of religious sort, and school inspection was enforced by the church. Though introduced already in 1794, state supervision of elementary schooling remained dead letter. An edict from 1812 ruled to establish school boards, chaired by the patron, composed of the parish priest

[10] Humboldt thought that: Every man obviously is only a good craftsman, merchant, soldier or businessman if he in general, and without regard to his profession, is a decent and enlightened person and citizen according to his class. If school instruction is giving him what is necessary to achieve this, he will accomplish the abilities of his profession with ease and will keep the freedom, like it often happens in life, to change from one to another (own translation from von Humboldt (1964, p. 218)).

[11] For example, in 1763 another law prohibited the teaching of Latin in rural schools, since *Landräte* repeatedly claimed that the most stubborn and unrelenting peasants were exactly those who had studied Latin.

and several heads of a family. In 1826 another decree suggested that patrons should preferably appoint certified teachers.

Similarly to Lindert, Becker and Woessmann (2009) interpret the growth of schooling in Prussia as a demand side expansion. While the population urged to become more literate, a strong force to prevent them was the landowner. Reforms thus gained more impact on free country parishes where school patronage was less conservative. Also, secular school inspection became more usual in towns. In both rural and urban areas secularization and the increase of public influence on schooling were a long-winded process lasting throughout the 19th century.

3.4 School financing

The composition of sources for school funds in Prussia was not legally fixed and it differed by region or province. Large parts of the expenses were covered by school assets like estates, entitlements or capital rents. Many of these assets had been granted when the commons were divided during the agrarian reform (1811) or had been given as donations. Assets were managed by so-called schooling societies, bodies of the municipality which received additional taxes from the heads of each household (excluding noble landowners) proportional to their wealth and income, independent of their religious denomination and number of children. In those cases where society funds were not sufficient to support schools and teachers, tuition fees were charged. Important for our case, the noble estate owners (*Gutsbesitzer*) were obliged to support families which could not afford to pay for schooling. Otherwise they were exempted from all taxes and therefore from any financial support for the schools (Kuhlemann, 1991). In addition, very poor school districts received financial support from the King (Königliches Statistisches Bureau in Berlin, 1889, pp. 58-59). It was only in 1888 that tuition fees were abolished and public funding became fixed. The exemption of noble estate owners from school financing and school taxes which had been legitimated in the "Allgemeines Landrecht" from 1794 was finally abolished in 1906.

The Prussian Statistical Office provides detailed information about public primary school financing at the province level for 1861, 1864, 1867, 1871, 1878, and 1886. In Table 1 we show the total amount of school funds per capita at the province level. Similarly to Lindert (2003, 2004), we find that Eastern provinces had relatively lower school funds compared to Western provinces. For example, primary school funds amounted to about 1.4 Marks per capita in East- and Westprussia in

1861,[12] whereas it amounted to 1.8 Marks in the Western province of Rhineland. However, it is important to note that this difference may reflect differentials in the cost of living. In fact, if we deflate school funds per capita with rye prices in 1861, funds in the provinces of East- and Westprussia are 10 and 14% higher than in Rhineland, respectively.[13]

School funds can be divided into three broad categories: local taxes and endowments from school societies, tuition fees, and state funds. In Table 2 we show the extent of state expenditures on public primary schooling at the province level. The data show that, throughout the period considered, poor Eastern provinces such as Eastprussia, Westprussia, and Poznan benefited to a larger extent from State contributions than the Western provinces of Westphalia and Rhineland. Also the province of Brandenburg received a relatively high contribution from the State.

On the contrary, school funds in Eastern provinces relied to a much lesser extent on tuition fees. In 1861, in the provinces of Rhineland and Westphalia tuition fees accounted for 24 and 27% of the total school funds, respectively; instead, in East- and Westprussia tuition fees accounted for only 11 and 13% of the total school funds, respectively.

Thus, the aggregate data show that school funds in Eastern areas were relatively lower in nominal terms with respect to the richer and more urban regions of the West, though in real terms the gap might have been nonexistent. Consistent with this finding, we observe no appreciable difference between East and West when looking at the number of schools and teachers per child at school-age. In Figure 1 and 2 we plot the number of schools and of teachers per 100 children (6-14) for 25 districts by year (1816, 1849, 1864, 1886, 1896). To facilitate the interpretation, the districts are sorted from East to West. One can immediately observe that, if anything, Eastern districts had more schools per child at school-age compared to Western districts. We obtain a very similar picture if we look at the number of teachers per child at school-age.

Therefore, the lower (nominal) level of financial support in the rural eastern areas arguably did neither impact the availability of schools nor that of teachers.[14] Supply of education was not restrained in the Eastern regions. As we will demonstrate in the rest of the paper, it was a lack of demand for education due to extractive

[12] Note that both provinces of East- and Westprussia belong to the East-Elbe part of the Kingdom of Prussia.
[13] The gap between school funds in East and West Prussia also decreases if we deflate by wheat prices in 1861.
[14] Of course we cannot exclude that it affected their quality.

institutions which explains the delay in educational attainment of the agricultural Eastern regions of Prussia to a large extent.[15]

4 Data

4.1 Prussian census data

In order to test whether the concentration of large landownership, a proxy for the institution of serf labor, delayed the expansion of education in Prussia, we collected county level census data. Our data comprehend five points in time (1816, 1849, 1864, 1886, 1896) spanning the entire nineteenth-century and come from various sources published on behalf of the Royal Prussian Statistical Bureau in Berlin.[16] The censuses contain a wealth of information, including data on education, landholdings by farm size, population, religion, and occupations. Given the inconsistency of definitions for some important variables, we decided to base our main analysis on cross-sectional comparisons. This approach allows to study how the effect of landownership concentration varied during the nineteenth-century. However, we are able to test our hypothesis organizing the data as a panel, accounting for unobserved heterogeneity.

Prussian county level data are challenging because of several changes in the administrative boundaries. The number of counties in Prussia increased during the nineteenth-century due to the fact that some counties were split into two or more counties for administrative reasons (mostly population growth) or due to acquisition of new territory. To achieve better comparability and to ensure that our results are not driven by changes in administrative boundaries, we aggregate our variables to a common set of boundaries. In particular, we aggregate census data to match the borders of 1849. This was not possible for some variables in the 1816 cross-section because, in some cases, counties were first grouped and then newly divided. Thus we decided to analyze the 1816 cross-section on the basis of its original structure.

For our dependent variable, we collected schooling data (enrollment rates) for the years 1816, 1849, 1864, 1886 and 1896. Primary school enrollment rates represent school attendance of the 6 to 14 years-old.[17]

[15]See Chaudhary (2009) about the role of private and public spending on education in India.

[16]See Appendix A.2 for more details about data sources.

[17]Consistent with the definition of mandatory schooling at the time, we consider both elementary schools (*Elementarschulen*) and middle schools (*Mittelschulen*) as primary schools. In a few cases enrollment rates exceed 100%. This could be due to children commuting from neighboring counties or because of children above age 14 enrolled in school.

Data on landownership by size is available for the years 1816, 1849, 1858, 1882 and 1895. The Prussian censuses counted the number of landholdings per county and classified them into size bins.[18] The first full census in 1816 classified landholdings in 3 groups: properties or leasehold estates up to 15 Prussian Morgen (henceforth PM), from 15 to 300 PM and more than 300 PM.[19] This categorization reflects the contemporary agricultural structure of farming. Farmers with less than 15 PM usually required some additional form of income. Landholdings between 15 and 300 PM were generally large enough for the subsistence of a family, whereas farms with more than 300 PM were usually cultivated by laborers and serfs while the owner was not expected to perform any manual work (Harnisch, 1984). The censuses of 1849 and 1858 extended the classification of landholdings from 3 to 5 bins. From 1882 onwards, the census considered only the arable land which was classified into 6 bins, increasing to 7 in the 1895 census.

Unfortunately, the total area of all holdings by bin-category was not published until 1882, so the calculation of conventional measures of inequality is not straightforward. Also, since political power was to a large extent proportional to the size of the possessed land (Eddie, 2008, p. 86), the share of largest landholdings seems to be a variable better suited to capture that dimension of political inequality.

Our measure of landownership concentration is the number of landholdings larger than 300 PM (circa 75 ha.) divided by the total number of holdings per county. To allow inter-temporal comparisons, we define large holdings as those which exceed 300 PM in the years 1816, 1849 and 1858 and 100 hectares (equal to 392 PM) in 1882 and 1895.[20] Our indicator is therefore bounded between zero and one: the larger the indicator, the higher is the share — the stronger the concentration — of large landowners in the county. In this way, we aim to capture those large estates which heavily relied on serf labor, and where peasants' freedom and mobility was severely limited especially in the first half of the nineteenth-century.

However, in order to compare our index of concentration with a more common measure of inequality such as the Gini coefficient, we need to compute the average size of farms for each bin-category. For the years 1849 and 1858 we use district level data on the average size of a farm by category from 1858. This information is not

[18] It is important to note that the original census data contains both types of tenure: landownership and landholding. Throughout the paper we use these terms interchangeably. Also note that the number of estates might not be equal to the number of landowners/-holders since they might own/hold more estates at the same time. Similarly, we are aware that both the King and the Church owned vast amounts of land. Unfortunately, the data do not allow to identify the landowner. See Eddie (2008) for a discussion of farm statistics and property statistics.

[19] One PM is equal to circa 0.25 ha.

[20] This change in size is due to redefinition of bin sizes by the Prussian Statistical Bureau when reforming measurement from PM to hectare.

available for 1816, therefore we assume the average area of a farm within each bin to be the mean of each category. For the (open) top category with farms of more than 300 PM, we use a value of 3000 PM. We choose this value in order to get an overall inequality in 1816 consistent with the trend we observe for the successive years. In fact, in the period 1849-1896 the Gini coefficient increases from 0.70 to about 0.76 (see Table 3). With the value of 3000 PM for the top open category, the Gini coefficient for 1816 is about 0.65.[21]

Our covariates aim to control for different aspects of the demand and supply of education. We include the share of Protestant population as Protestants are expected to have a higher demand for education, the share of urban population, the share of population employed in the industrial sector,[22] the share of population employed in agriculture, and population density.[23] The age structure of the county might also influence the demand for education.[24] In our case this factor is proxied by the child dependency ratio, calculated as the share of the young population (0 to 14/18) over the working population (15/19 to 65/70). Additionally we control for the share of the population whose first language is not German.[25] This variable controls for differences between the former Polish parts of Prussia (Poznan, Silesia and West-Prussia) and Germany where the demand for education might have been lower for linguistic and cultural reasons. This variable is inversely correlated with the share of Protestants as the Polish population was predominantly Catholic. Finally, we also control for the supply of schools using school density, defined as the number of schools per square kilometer.

Quite importantly, inheritance laws differ across Prussia and that might have an impact on our analysis. In fact, inheritance laws follow a geographical pattern: The north-eastern parts of Prussia are dominated by non partible inheritance (*Anerbenrecht*), while the south-western parts are characterized by partible inheritance (*Realteilung*). This clearly leads to a different average size of landownership in the south-western counties. In our empirical analysis, we shall take this institutional difference into account with a binary variable which takes on value one in counties with partible inheritance.

[21]Note that our results do not change qualitatively when we use the lower values 300, 600, 1000, or 2000 for the open category.

[22]Given the low level of industrialization in Prussia in the first decades of the nineteenth-century, for 1816 we calculate the number of looms over the total population as an indicator of non-agricultural occupation.

[23]Unfortunately the 1816 census provides only information on the laborers in agriculture, whereas the 1849 census provides information on the total population in agriculture, including family members.

[24]About the relationship between fertility and education in pre-demographic transition Prussia see Becker et al. (2010, 2012).

[25]Unfortunately the earliest available census data about spoken languages is from 1890.

4.2 Descriptive statistics

Table 3 provides descriptive statistics of our variables for each year separately. The first thing to note is the relatively high levels of enrollment rates. Already in 1816, about 60 percent of children aged 6-14 were enrolled in primary schools (see also Becker and Woessmann (2010)). Subsequently, we can observe an increase in average enrollment rates over time, rising from 60% in 1816 to 94% in 1895. Interestingly, the enrollment rate in 1864 is lower compared to the previous census in 1849.[26]

The variable *large landholdings* indicates the share of the largest holdings in the county. In 1816, 1.7% of the holdings belonged to this category (>300 *PM*). Due to problems of consistency of the definitions, we can only compare the values of landownership concentration for 1816, 1849 and 1864 *directly*. The descriptive statistics show that the share of large landholdings increased during the first half of the nineteenth-century, reaching a share of 2.5% in 1858.[27] This evidence is consistent with the accounts of the agrarian reform in Prussia which stress how the landed elite benefited most from the reforms (Schiller, 2003).

As mentioned, a direct comparison with 1882 and 1895 is not straightforward since these censuses only account for the size of arable land. Furthermore, the unit of measurement in these censuses is the hectare and thus the top bin-category changes to 100 hectare.[28] This explains why our concentration index decreases to 0.8% in 1882 and to 1% in 1895.

The progress of industrialization is shown by the increase of urbanization which rises from 24% to 30% and industrial employment which increases from 1% to 12%. These variables provide large variations both across counties and over time, whereas variables such as religious denomination, language, and inheritance system are rather time invariant.

Figure 3 and 4 show the geographic distribution of enrollment rates and the concentration of landownership in 1849, respectively. Enrollment rates are higher in the central part of Prussia, whereas tend to be lower in the eastern areas, especially in the province of Posen. Becker and Woessmann (2009) suggest that, due to the influence of the Protestant Reformation, literacy as a by-product of Protestantism tended to spread almost concentrically around Wittenberg. The geographic distribution of the concentration of landownership is almost the reverse of Figure 3, as areas with lower enrollment rates are now characterized by a relatively high

[26] This peculiarity has already been found by Lindert (2004, p. 91, Tab. 5.1).
[27] The year 1864 in the table headings refers to the date for which data on enrollment rates are available.
[28] 100 hectares are equal to 392 PM.

concentration of large holdings. Western areas are strongly characterized by small holdings, though also western counties present a certain degree of variation.

5 OLS results

In this section we test the hypothesis of a negative association between the concentration of large landowners and school enrollment in Prussia. We estimate this relationship for five different cross-sections in the nineteenth-century to understand if and to what extent the landed elite lost influence in opposing the expansion of mass education.

We estimate a standard OLS model where the enrollment rate *edu* is a function of the concentration of landownership *land*, plus a vector of covariates X for a county i in a given year t:

$$edu_i^t = \alpha_1 + \beta_1 land_i^t + X_i^t \gamma_1 + \varepsilon_i^t. \tag{1}$$

β_1 is the coefficient of interest. The vector X contains the share of Protestants in the county, the share of population living in urban centers, the share of the population working in the industrial sector, the share of the population in agriculture, population density, and the child dependency ratio. In addition we control for school density, for inheritance laws using a binary variable which takes on value one for partible inheritance, and the share of the population whose first language is not German.

The results of the OLS estimates are presented in Table 4. We consistently find a significant negative association between the share of large landholdings and the enrollment rate. This effect, very large in 1816, seems to fade throughout the nineteenth-century. As already mentioned, it is possible to compare the magnitudes of the coefficients of the first three years (1816, 1849, and 1864) as the variables are defined similarly and have the same unit of measurement. In fact, the coefficient for 1816 is statistically larger than the coefficient for 1849, whereas the coefficients for 1849 and 1864 are not statistically different.

Comparing the coefficients over the entire period might lead to problematic conclusions. Since the dependent variable is theoretically bounded between zero and one, the steady increase in enrollment rates over time leads to decreasing variation. This might lead to a decrease in the coefficients over time due to a shrinking variation of the dependent variable. Therefore, we provide the interpretation of the effects also in terms of standard deviations.

In fact, the estimated effects are quite large. In 1816, an increase in the share of large estates by one standard deviation is associated with a decrease of the enrollment rate by about 0.34 standard deviations. In 1849 the effect is about 0.31 standard deviations, whereas it is 0.29 standard deviations for 1864. Finally, in the last two years 1886 and 1896, an increase by one standard deviation in landownership concentration implies a decrease of the county enrollment rate by 0.09 standard deviations.

As already shown in Becker and Woessmann (2009), there is a positive relationship between Protestantism and educational attainment which, however, loses significance in the last two cross sections. Another interesting result stems from the industrial variable. We observe that the coefficient becomes positive from 1849 and highly significant from 1886. This result is consistent with the study of Becker et al. (2011) which argues that, differently than in Britain, education was causal for industrialization in Prussia.[29]

Cultural differences in valuing education are captured by the 'language' variable. Counties with a higher share of non-German speaking population, mostly of Slavic origin, show significant lower levels of enrollment rates. Given that we control for the share of Protestants, the estimated effect captures cultural values that lie outside the religious sphere.

In our case it is important to note the significant positive relationship between school density and enrollment rates throughout the whole period. This result indicates that the supply of schools was not the mechanism through which large landowners opposed the spread of mass education. If that were the case, the coefficient β_1 attached to large landholdings should decrease when controlling for school density. We actually find that the effect of large landholdings slightly increases when controlling for school density.

Clearly school density is to a large extent endogenous to enrollment rates. However, it is reasonable to argue that the number of schools might have responded with some delay to an increase in the demand for education. This seems to be the case when we observe the trend of schools per child over time: Despite a steady increase in school enrollment, the number of schools per 100 children aged 6-14 de-

[29]This is especially true for non-textile sectors. The authors also go a step forward showing that the effect of education on industrialization is causal. Note that different from Becker et al. (2011) we do not exclude hand driven looms in the 1849 cross-section.

creased substantially, at least until 1886.[30] Therefore, it seems justifiable to use school density as a covariate in our model.[31]

In sum, OLS estimates show that landownership concentration has a strong negative association with enrollment rates in nineteenth-century Prussia. Results suggest that the effect of landownership concentration on education diminished throughout the century. The complete emancipation of the peasantry in the second half of the nineteenth-century is likely to have triggered a higher demand for education.

6 Establishing causality

6.1 The causal effect of landownership concentration

The effects estimated by OLS are not causal as they might be affected by omitted variable bias or reverse causality. A variable which is correlated with both landownership and enrollment rates would bias our results. In addition, reverse causality might also be an issue. Peasants with a higher level of education might have been comparatively more able to appeal to the King in order to obtain better tenants conditions. Peasants with a higher level of education might also increase the productivity of the land they work. In a fixed-rent regime, this could imply that higher education peasants might have been more able to redeem their land. Finally, higher educated peasants might have had a higher incentive to sell their (small) estate to a large landowner in order to reap the benefits of their education in other trades. All these mechanisms would bias previous OLS estimates.

In order to overcome such problems, we adopt an instrumental variable approach. Similar to Easterly (2007), we identify the causal effect of land concentration on education using exogenous variation in farm size due to differences in the geological composition of the soil (soil texture). In fact, a 1866 census assessed the composition of the soil at the county level and classified it into three categories: (i) loam and & clay, (ii) sandy-loam and loamy-sand, (iii) and sand. Terrain of the first category (loam & clay) tend to be relatively more fertile; sandy areas, instead, are the least fertile, whereas fertility of counties with a prevalence of the second category are somewhere in between. In Figure 5 we display the relative dominance of each category. Darker areas are those under the first category, therefore with a higher soil quality; areas with brighter colors are dominated by soil of lower quality.

[30] Schools per child (6-14) were 1.025 in 1816, 0.817 in 1849, 0.716 in 1864, 0.646 in 1886, and 0.754 in 1896.

[31] We decided to use school density instead of school per child (6-14) because we believe that the former is better suited to capture the cost associated with reaching the next school. It is important to note that throughout the paper our results are almost identical if we control for school per child.

There are a few mechanisms which corroborate our instrumental variable strategy. In fact, there is a vast literature in agricultural economics which finds a systematic negative correlation between soil quality and farm size (Bhalla and Roy, 1988; Bhalla, 1988; Benjamin, 1995). Regions which exhibit relatively lower quality of soil, therefore with lower marginal value of land, experience a lower demand for land and are thus characterized by higher average farm sizes (Barrett et al., 2010; Bhalla and Roy, 1988; Bhalla, 1988). This is indeed the pattern followed by north-eastern European regions during the Middle Ages where the local lords, in order to attract more agricultural workers and encourage migration, granted the peasants with higher levels of freedom compared to the manorial system in the West (Rosenberg, 1943). It was only after the Black Death epidemics in the 14th century that serfdom developed in eastern Europe, establishing the roots of landownership inequality that we observe also in the nineteenth-century.

Eastwood et al. (2010, p. 3355) answer the question why the so called colonial land grab, which predominantly installs large landholdings in underpopulated areas, is not adjusted over time when labor mobility grows, with life-style advantages for large farmers, local extra-economic status and power, and ethnic barriers limiting incentives to turn to more efficient small scale farming. In fact, by 1427, large landowners started to restrict freedom of movement of peasants and agricultural workers and increasingly exploited obligatory labor services (Rosenberg, 1944, p. 231-232).

On the other hand, regions with relatively higher soil quality experience a stronger demand for land which determines a more accentuated land fragmentation and more secure property rights. According to Boserup (1965); Binswanger and McIntire (1987); Binswanger and Rosenzweig (1986), increasing population pressure results in increasing intensification of land use and in a growing pressure for security of land tenures (Eastwood et al., 2010). In fact, we find that in the regions of Rheinland and parts of Saxony and Silesia, where the soil is dominated by loam and clay (i.e. higher quality), farm size is small on average, leading to a low concentration of large landholdings (compare Figure 4 and 5). In addition, soil quality might also influence crop choice which, in turn, due to economies of scale, might affect the final distribution of land (Vollrath, 2009). The advantage of our instrument with respect to crop choice is that soil quality is a "true" exogenous variable, whereas crop choice is still a choice variable, although heavily dependant on the type of terrain.[32]

[32] In this fashion, Easterly (2007) uses the suitability of crops such as sugar and wheat instead of actual crop production to identify exogenous variation in inequality.

Thus, our first stage is expressed by the following equation:

$$land_i^t = \alpha_2 + SOIL_i \beta_2 + X_i^t \gamma_2 + \eta_i^t \tag{2}$$

where *land* is the variable for landownership concentration, *SOIL* is a vector of variables describing soil texture, and X is the vector of covariates of equation 1. The exclusion restriction demands that soil texture has no direct effect on enrollment rates. In order to ensure that the exclusion restriction is not violated, we additionally control for yields per hectare of the most common crops (wheat, rye, barley, oats and potatoes).

In principle, the vector *SOIL* should contain the three variables that characterize the geological composition of soil, namely the different shares of loam & clay, sandy-loam and loamy-sand, or sand. Yet, the soil variables are highly correlated with each other: for instance, the correlation between the share of loam & clay and the share of sand is -0.7. Such a high correlation might influence the first stage regression, weakening the performance of our instruments. Therefore, we use a principal component analysis (PCA)[33] in order to recover two components that are uncorrelated by construction.[34] The first component explains 66.4% of the total variation of the three soil variables, whereas the second component explains an additional 32.9% of the variation. Thus, these two components add to more than 99% of the variation in soil texture.[35]

6.2 Instrumental variable results

First stage estimates of equation 2 are presented in the upper panel of Table 5. The two components are significantly correlated with landownership concentration. The first component is negative for loam & clay and picks up variation where landownership is less concentrated, leading to a positive correlation. The second component is positive for sandy-loam and loamy-sand and picks up variation in regions where landownership is highly concentrated. The power of the instrument is summarized by the first stage F-statistics, which, with the exception of 1816, are very close to or above the standard threshold value of 10.

[33] Principal component analysis (PCA) is a statistical technique used to reduce the number of variables in an analysis describing a series of uncorrelated linear combinations of the variables that contains most of the variance. See Appendix A.1 for more details.

[34] It is important to note that using the three soil variables as instruments provides virtually the same results, though in that case the F-statistics of the first stage indicate a weak performance of the instruments.

[35] For a similar approach see Galor et al. (2009). In their paper, they use a geographic element and a relative price element to identify exogenous variation in land distribution. The geographic element is constituted by variables such as temperature, rainfall and heating. Given the high correlation between these three variables, they also resort to a principal component analysis.

Second stage estimates are presented in the lower panel of Table 5. The results confirm the negative effect of landownership concentration on education and imply that the effect is indeed causal. Also in this case, there seems to be a fading effect of landownership concentration over time, although the inconsistency of the definitions across the censuses does not allow to be definitive on the matter. Yet, this seems to imply that the abolishment of serfdom and the emancipation of the peasantry triggered a stronger demand for education. The opposition of the large landowners towards the spread of mass education faded more or less steadily through the nineteenth-century, reaching a lower but stable level in the last part of the century.

Our IV estimates imply that OLS estimates are downward biased. This results is consistent with the findings of Easterly (2007). Instrumenting inequality using the the ratio between the suitability of wheat to sugar crops, he finds that IV estimates are about three times larger compared to OLS estimates (Easterly, 2007, p.766, Table 4). A similar ratio between IV and OLS estimates is found by Ramcharan (2010).

In terms of magnitude, we find that if the share of large estates in the county increases by one standard deviation, the enrollment rate would decrease by about 0.93 standard deviations in 1816, 0.85 standard deviations in 1849, 0.57 standard deviations in 1864, 0.28 standard deviations in 1886, and 0.19 standard deviations in 1896.[36] It implies that if the share of large landholdings in 1849 had doubled (for instance from the average of 2.4 to 4.8%), the enrollment rate in primary schools would have decreased by almost nine percentage points. A similar increase at the end of the century (from the average of 0.8 to 1.6%) would have resulted in only one percentage point lower enrollment rates.

7 Robustness checks

7.1 Agricultural productivity

The exclusion restriction is violated if soil texture has a direct effect on enrollment rates. This might be the case if, for instance, soil fertility led to variation in employment of child-labor in agriculture and thus drove children out of school. In order to tackle this problem, we additionally include productivity measures of the most important crops in our analysis. However, detailed county-level information on yields

[36]These findings do not depend on the threshold of 300 PM chosen to define large landholdings. For instance, for 1849 and 1864 it is possible to use the bin-category over 600 PM, and for 1896 the bin-category over 200 ha. Results are qualitatively similar when using these categories.

is only available for the years 1886 and 1896. Therefore, we decided to combine the 1886 yields data with the earlier cross-sections.[37]

Estimates presented in Table 6 show that controlling for yields per acre leads to similar results as in Table 5.[38] While the coefficient for large landholdings in 1849 and 1864 show a significant negative effect, the point estimates in 1886 and 1896 are smaller and insignificant. That suggests, again, that the effect of landownership concentration tended towards zero at the end of the nineteenth-century.

7.2 Urban vs rural

It could be argued that the relationship we estimate should exclude cities, since urban and rural landownership (and education) might be very different. Unfortunately we can only separate urban and rural data for the 1849 census. In that case, using only rural enrollment and landownership data, leaving all other variables unchanged, we find the same qualitative results (Table 7, column 6). In order to tackle this issue further, we can exclude those counties which consist only of a city.[39] The estimates, presented in Table 7, are consistent with the results discussed so far. As expected, with the only exception of 1864, the first-stage F-statistics suggest that the instruments are stronger when omitting city counties.

7.3 East vs West

Further concerns might arise due to the accentuated east-west gradient, most of all regarding landownership concentration. In fact, we can show that when excluding the two western provinces of Rheinland and Westphalia, results (both for OLS and IV) are qualitatively similar, confirming the findings discussed throughout the paper.

Analogously, one could argue about the existence of an east-west gradient regarding school financing. However, as discussed in more detail in Section 3.4, the differences in school funds in real terms were probably low. In fact, we observe that the supply of schools and teachers was not restrained in the East with respect to the richer regions of the West (see Figures 1 and 2). As we are trying to demonstrate, it was a lack of demand for education due to extractive institutions which explains to a large extent the delay in educational attainment of the agricultural Eastern regions of Prussia.

[37] We are aware that assuming constant productivity differences over time is a strong assumption. However, for completeness, we decided to show results for all cross-sections.

[38] For a small number of counties information about productivity is not reported in the censuses because the considered crops were not cultivated. That explains the smaller number of observations.

[39] The city-counties are Aachen, Berlin, Danzig, Halle, Frankfurt/Oder, Köln, Königsberg, Magdeburg, Münster, Potsdam.

7.4 The Gini coefficient

As explained in more detail in section 3, political power in nineteenth-century Prussia was associated with the size of land property. We argued that our measure of land concentration is better suited to capture the extent of labor exploitation than standard measures of inequality. In fact, our concentration variable can differ substantially from a standard measure of inequality such as the Gini coefficient. Let us assume an extreme case where a given county has 100 landholdings and assume that all the holdings belong to the largest category ($>$ 300 PM). In this case, the Gini coefficient would be equal to zero since land is equally distributed among the 100 large landholders. Our concentration variable, instead, would take the value of one since large landholdings represents 100% of the total number of holdings. In such an extreme case, we expect to find a particularly low level of enrollment rates as all (landless) workers are bound, through the serfdom system, to the few large landowners. We argue that this institutional setting, for a given level of supply of education, inhibited the demand for schooling.

In order to compare our measure of land concentration with standard inequality measures, we compute the Gini coefficient and estimate its effect on education.[40] Results are reported in the upper panel of Table 8. The Gini coefficient has no effect on school enrollment rates and the same result holds if we use alternative measures for inequality such as the Theil index.[41]

Yet, it is important to note that the Gini coefficient computed across all landholdings does not consider landless agricultural workers. We correct for this using a method proposed by Vollrath (2010, p. 8-9) which incorporates the total number of adult males.[42] In particular, the adjusted Gini is equal to $pG + (1 - p)$, where p is the ratio of landholdings to adult males and G is the standard Gini coefficient. The lower the number of farms with respect to the adult male population, the larger is the adjusted Gini with respect to the standard Gini. The descriptive statistics in Table 3 show that the adjusted Gini is systematically larger than the standard Gini.

In the lower panel of Table 8 we present OLS estimates when regressing enrollment rates on the adjusted Gini. As expected, the standard Gini underestimates the relationship between land inequality and enrollment rates. The results using the

[40]For 1886 and 1896, we merged the two smallest categories in order to have an homogenous bin size of 0-2 hectares. This seems to be particularly sensible for 1896 where the smallest bin size (0-0.5 hectares) is likely to include also small private gardens. In this way we dilute this category in the next largest one which we expect to include truly small landholders.

[41]Result are similar for different generalized entropy indices.

[42]One could argue that our concentration index should also be adjusted for landless agricultural workers. This does not apply to our case since our variable is intended to capture average farm size which, in turn, is a proxy for the institution of serfdom.

adjusted Gini show that, for the first half of the nineteenth-century, the relationship between land inequality and enrollment rates is negative and significant.

The soil texture information used to identify exogenous variation in land concentration appear to be a rather weak instrument for the standard and the adjusted Gini. Therefore instrumental variable estimates generate coefficients with a very large bias.[43] However, we can show that the pattern of the results is very similar to the results discussed above: after the effective abolition of serfdom and with the complete emancipation of peasantry, the relationship between land inequality and enrollment rates ceases to be significant.

7.5 Panel models

In this section, we construct a panel dataset which allows to address the issue of unobserved heterogeneity. Time-invariant initial differences across counties which are not fully accounted for might affect estimates of equation 1. Panel models with county fixed effects solve this problem. In fact, fixed effect estimates show how changes in landownership concentration affect changes in enrolment rates within counties. In order to maintain constant borders we aggregate the data to resemble the administrative structure in place in 1800. Thus, our estimates are based on a panel consisting of 280 counties i observed at five points in time t (1816, 1849, 1864, 1886, 1896):

$$edu_{it} = \beta_3 land_{it} + \alpha_i + \tau_t + X_{it}\gamma_3 + \nu_{it} \qquad (3)$$

where α_i and τ_t are county and time fixed effects, respectively.

As already mentioned in section 4, the census definition of landownership changed substantially in the last part of the century, including only arable land starting from 1882. Time fixed effects, beyond capturing average changes over time, will also take into account the change in definition for landownership.

Panel estimates are presented in Table 9. We show specifications using landownership concentration and the adjusted Gini alternatively. For comparison, in columns 1 and 4 we present pooled models with neither county nor time fixed effects. In the next columns we sequentially introduce county and time fixed effects.

In all specifications we find a significant negative effect of landownership concentration and the adjusted Gini on enrollment rates. Including time fixed effects does not change the results qualitatively. In fact, time fixed effects reinforce the coefficient of the adjusted Gini (column 6) whereas it reduces the effect of landownership

[43] Results not presented but available upon request.

concentration (column 3). The latter can be explained by the removal of the scale effect triggered by the change in the definition of landownership for the last two censuses.

In sum, the panel analysis confirms the findings of the previous sections. The magnitude of the coefficients estimated in the panel models is consistent with the ones estimated through OLS and IV. This implies that the results obtained in the cross-sectional analysis are not affected by unobserved heterogeneity at the county level.

8 Conclusion

In this paper, we show to what extent landownership concentration, a proxy for the extractive institution of serfdom, affected the spread of primary education in nineteenth-century Prussia. We argue that landowners hampered peasants' demand for education through the institution of serf labor, delaying the spread of mass primary education.

Indeed, using a unique county level dataset that covers the entire nineteenth-century, we find that counties with a higher share of large landholdings had significantly lower enrollment rates. This result holds when controlling for several demand and supply factors, including the availability of schools and ethnolinguistic fractionalization. We provide evidence that the negative effect of landownership concentration weakened over time, suggesting a diminishing influence of the landed elite towards the end of the nineteenth-century. This timing is consistent with our proposed mechanism of labor exploitation, since peasants gained their complete legal emancipation *de facto* only around 1850.

To overcome possible biases due to omitted variables and reverse causality we adopt an instrumental variable approach. We identify the causal effect of landownership concentration on education using exogenous variation in farm size due to differences in the geological composition of the soil (soil texture). Regions which exhibit relatively lower quality of soil, therefore with lower marginal value of land, experienced historically a lower demand for land and are thus characterized by higher average farm sizes. On the contrary, regions with relatively higher soil quality experienced a stronger demand for land which determined a more accentuated land fragmentation in the long run.

Instrumental variable estimates confirm the negative effect of landownership concentration on education and suggest that the effect is indeed causal. The estimates also confirm that the effect diminished over time, supporting the proposed mechanism of serf labor as explanation of the negative relationship between landownership concentration and education.

Finally, we construct a panel dataset which allows to rule out unobserved heterogeneity as a potential bias of our cross-sectional estimates. County and time fixed-effects estimates confirm the existence of a significant negative effect of landownership concentration on enrollment rates also within counties.

We suggest that the institution of serf labor, by limiting the freedom of peasant households, made any investment in formal education unprofitable, notwithstanding the presence of schools and teachers. Successively, the abolition of serf labor and the full emancipation of the peasantry triggered an increasing demand for primary education which allowed eastern areas to catch-up with areas characterized by higher levels of education.

References

Acemoglu, D., D. Cantoni, S. Johnson, and J. A. Robinson (2011): "The Consequences of Radical Reform: The French Revolution," *American Economic Review, forthcoming.*

Acemoglu, D., S. Johnson, and J. Robinson (2001): "The Colonial Origins of Comparative Development," *American Economic Review*, 91, 1369 – 1401.

——— (2002): "Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution," *Quarterly Journal of Economics*, 117, 1231 – 1294.

Acemoglu, D. and J. Robinson (2000): "Why Did the West Extend the Franchise? Democracy, Inequality and Growth in Historical Perspective," *Quarterly Journal of Economics*, 115, 1167 – 1199.

Alesina, A. and D. Rodrik (1994): "Distributive Politics and Economic Growth," *Quarterly Journal of Economics*, 108, 465 – 490.

Banerjee, A. and L. Iyer (2005): "History, Institutions, and Economic Performance: The Legacy of Colonial Land Tenure Systems in India," *American Economic Review*, 95, 1190 – 1213.

Barrett, C. B., M. F. Bellemare, and J. Y. Hou (2010): "Reconsidering Conventional Explanations of the Inverse Productivity-Size Relationship." *World Development*, 38, 88 – 97.

Becker, S. O., F. Cinnirella, and L. Woessmann (2010): "Education versus Fertility: Evidence from before the Demographic Transition," *Journal of Economic Growth*, 15, 177 – 204.

——— (2012): "The effect of investment in children's education on fertility in 1816 Prussia," *Cliometrica*, 6.

Becker, S. O., E. Hornung, and L. Woessmann (2011): "Education and Catch-up in the Industrial Revolution," *American Economic Journal: Macroeconomics*, 3, 92 – 126.

Becker, S. O. and L. Woessmann (2009): "Was Weber Wrong? A Human Capital Theory of Protestant Economic History," *Quarterly Journal of Economics*, 124, 531 – 596.

——— (2010): "The Effect of Protestantism on Education before the Industrialization: Evidence from 1816 Prussia." *Economics Letters*, 107, 224 – 228.

Benjamin, D. (1995): "Can Unobserved Land Quality Explain the Inverse Productivity Relationship?" *Journal of Development Economics*, 46, 51 – 84.

Bhalla, S. S. (1988): "Does Land Quality Matter?" *Journal of Development Economics*, 29, 45 – 62.

Bhalla, S. S. and P. Roy (1988): "Mis-Specification in Farm Productivity Analysis: The Role of Land Quality." *Oxford Economic Papers, New Series*, 40, 55 – 73.

Binswanger, H. P. and J. McIntire (1987): "Behavioral and Material Determinants of Production Relations in Land-Abundant Tropical Agriculture." *Economic Development and Cultural Change*, 36, 73 – 99.

Binswanger, H. P. and M. R. Rosenzweig (1986): "Behavioural and Material Determinants of Production Relations in Agriculture." *Journal of Development Studies*, 22, 503 – 539.

Boserup, E. (1965): *The Conditions of Agricultural Growth: The Economics of Agrarian Change under Population Pressure*, Aldine Pub. Co., New York.

Bowman, S. D. (1980): "Antebellum Planters and Vormärz Junkers in Comparative Perspective," *American Historical Review*, 85, 779 – 808.

Brenner, R. (1976): "Agrarian Class Structure and Economic Development in Pre-Industrial Europe," *Past & Present*, 70, 30 – 75.

Carsten, F. L. (1988): *Geschichte der preußischen Junker*, Suhrkamp, Frankfurt a. M.

Chaudhary, L. (2009): "Determinants of Primary Schooling in British India." *Journal of Economic History*, 69, 269 – 302.

Dell, M. (2010): "The Persistent Effects of Peru's Mining Mita." *Econometrica*, 78, 1863 – 1903.

Domar, E. D. (1970): "The Causes of Slavery or Serfdom: A Hypothesis," *Journal of Economic History*, 30, 18 – 32.

Easterly, W. (2007): "Inequality does Cause Underdevelopment: Insights From a new Instrument," *Journal of Development Economics*, 84, 755 – 776.

Eastwood, R., M. Lipton, and A. Newell (2010): "Farm size," in *Handbook of Agricultural Economics*, ed. by R. Evenson and P. Pingali, Elsevier, Amsterdam, vol. 4, 3323 – 3397.

Eddie, S. M. (2008): *Landownership in Eastern Germany before the Great War - A Quantitative Analysis*, Oxford University Press.

Engerman, S. L. and K. L. Sokoloff (1997): "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States." in *How Latin America fell behind: Essays on the economic histories of Brazil and Mexico, 1800-1914*, ed. by S. H. Haber, Stanford University Press, 260 – 304.

Engerman, S. L., K. L. Sokoloff, M. Urquiola, and D. Acemoglu (2002): "Factor Endowments, Inequality, and Paths of Development among New World Economies [with Comments]," *Economía*, 3, 41 – 109.

Frankema, E. (2010): "The Colonial Roots of Land Inequality: Geography, Factor Endowments, or Institutions?" *Economic History Review*, 63, 418 – 451.

Gallego, F. A. (2010): "Historical Origins of Schooling: The Role of Democracy and Political Decentralization," *Review of Economics and Statistics*, 92, 228–243.

Galor, O. (2011): "Inequality, Human Capital Formation and the Process of Development," Working Paper 17058, National Bureau of Economic Research.

Galor, O. and O. Moav (2006): "Das Human-Kapital: A Theory of the Demise of the Class Structure," *Review of Economic Studies*, 73, 85 – 117.

Galor, O., O. Moav, and D. Vollrath (2009): "Inequality in Land Ownership, the Emergence of Human Capital Promoting Institutions and the Great Divergence." *Review of Economic Studies*, 76, 143 – 179.

Galor, O. and J. Zeira (1993): "Income Distribution and Macroeconomics." *Review of Economic Studies*, 60, 35 – 52.

Go, S. and P. Lindert (2010): "The Uneven Rise of American Public Schools to 1850," *Journal of Economic History*, 70, 1 – 26.

Harnisch, H. (1984): *Kapitalistische Agrarreform und Industrielle Revolution*, Hermann Böhlaus Nachfolger, Weimar.

Königliches Statistisches Bureau in Berlin (1889): *Preussische Statistik*, vol. 101, Ernst Kuehn.

Kuhlemann, F.-M. (1991): "Schule, Hochschulen, Lehrer," in *Handbuch der deutschen Bildungsgeschichte*, ed. by C. Berg, Verlag C.H. Beck, München, vol. 4. 1870-1918: Von der Reichsgründung bis zum Ende des Ersten Weltkriegs, 170 – 370.

Lindert, P. H. (2003): "Voice and Growth: Was Churchill Right?" *Journal of Economic History*, 63, 315 – 350.

——— (2004): *Growing public: Social Spending and Economic Growth Since the Eighteenth Century*, Cambridge University Press.

Mariscal, E. and K. L. Sokoloff (2000): "Schooling, Suffrage, and Inequality in the Americas, 1800-1945," in *Political Institutions and Economic Growth in Latin America. Essay in Policy, History, and Political Economy*, ed. by S. H. Haber, Hoover Institution Press, Stanford, 159 – 218.

Melton, J. V. H. (1988): *Absolutism and the Eighteenth-Century Origins of Compulsory Schooling in Prussia and Austria*, Cambridge University Press.

Neugebauer, W. (1992): "Das Bildungswesen in Preußen seit der Mitte des 17. Jahrhunderts," in *Handbuch der preussischen Geschichte*, ed. by O. Büsch, Walter de Gruyter, Berlin, New York, vol. 2, 605 – 799.

Persson, T. and G. Tabellini (1994): "Is Inequality Harmful for Growth? Theory and Evidence," *American Economic Review*, 84, 600 – 621.

Pierenkemper, T. and R. Tilly (2004): *The German Economy during the Nineteenth Century*, Berghahn Books.

Ramcharan, R. (2010): "Inequality and Redistribution: Evidence from U.S. Counties and States, 1890-1930," *Review of Economics and Statistics*, 92, 729 – 744.

Rosenberg, H. (1943): "The Rise of the Junkers in Brandenburg-Prussia, 1410-1653: Part 1," *American Historical Review*, 49, 1 – 22.

——— (1944): "The Rise of the Junkers in Brandenburg-Prussia, 1410-1653: Part 2," *American Historical Review*, 49, 228 – 242.

Schiller, R. (2003): *Vom Rittergut zum Grossgrundbesitz: Ökonomische und soziale Transformationsprozesse der ländlichen Eliten in Brandenburg im 19. Jahrhundert*, Akademie Verlag, Berlin.

Schleunes, K. A. (1979): "Enlightenment, Reform, Reaction: The Schooling Revolution in Prussia," *Central European History*, 12, 315 – 342.

Sokoloff, K. L. and S. L. Engerman (2000): "History Lessons: Institutions, Factors Endowments, and Paths of Development in the New World," *Journal of Economic Perspectives*, 14, 217 – 232.

Vollrath, D. (2009): "Geographic Elements of Farm Size Inequality in the United States," mimeo, University of Houston.

——— (2010): "School Funding and Inequality in the United States, 1890," mimeo, University of Houston.

von Humboldt, W. (1964): "Bericht der Sektion des Kultus und Unterrichts an den König, Dezember 1809," in *Werke in fünf Bänden*, ed. by A. Flitner and K. Giel, vol. Band 4: Schriften zur Politik und zum Bildungswesen.

Wagner, P. (2001): "Landräte, Gutsbesitzer, Dorfschulzen - Zum Wandel der "Basisposten" preußischer Staatlichkeit in Ostelbien im 19. Jahrhundert," in *Preußens Weg in die politische Moderne*, ed. by B. Holtz and H. Spenkuch, Akademie Verlag, Berlin, 249 – 283.

Webb, S. B. (1982): "Agricultural Protection in Wilhelminian Germany: Forging an Empire with Pork and Rye," *The Journal of Economic History*, 42, 309 – 326.

A Appendix

A.1 Soil data and principal component analysis

The instrument soil texture is developed using data from a 1866 classification of soil. As a means to unify land taxes across provinces, the Prussian government released a law to determine soil, location, terrain, climate, humidity, accessibility, credit and all other aspects affecting the rate of return to each farm. These information were collected between 1861 and 1864 in each county by an assessment commission of 4 to 10 members, half of which were elected by the local assembly and the other half chosen by the central government. Of the total 2494 members of the commission, 701 were owners of knight estates and 1181 were other landowners, while the rest were public servants and landless farmers. Land surveying was conducted by approximately 3300 trained technical measurers. The local classification of soil was conducted by the 3 members of the commission, including the commissar and one technical measurer. The soil was either drilled or dug, classified and sketched into maps. Soil samples were taken in order to compare them to other soils in the same classification. The samples and their exact location had to be attached to the survey protocol and were send to the district commission. The district commission visited each county and compared the samples. The classification into tax classes was then made available to the municipalities, to estate owners and to county assemblies for objection. The average surveyed area was 568 Prussian Morgen per day, summing to total of 188,587 days of work. Given that our data provides county level averages of soil texture, possible biases due to systematic misreporting for purposes of tax evasion should be minimized. The data reports the total area, the area of loam and clay soils, the area of sandy loam and loamy sand soils, and the area of sand soils.

In order to display the relative dominance of each of these arable soil categories we calculate their shares in total soil. Since the soil shares are highly correlated at the county level, we can not use the original values as instruments without artificially inflating the significance of the first stage regressions. Thus we use principal component analysis (PCA) in order to eliminate correlation among the variables. We choose the first two components of the PCA which explain 99.3 percent of the variation in the soil variables as our instruments. The first component has an eigenvalue of 1.97 and its loadings are -0.71 for loam and clay, 0.49 for sandy loam and loam sand, and 0.51 for sand. The second component has and eigenvalue of 1.01 and its loadings are 0 for loam and clay, 0.72 for sandy loam and loam sand, and -0.7 for sand. The source of the soil data is Meitzen (1868, vol. 4).

A.2 Data sources

We have compiled a county-level database covering all Prussian counties over the nineteenth-century. The data were collected in several censuses by the Royal Prussian Statistical Bureau and obtained from several sources. We combine these data to prepare five cross section using censuses from the years 1816, 1849, 1864, 1886 and 1896 or, if not available, as close as possible to these years. We accounted for changes in the administrative boundaries of counties by adjusting all sources to the 1849 county borders. An exclusion is made for the 1816 cross section since it was not possible to calculate landownership data according to the administrative borders of 1849. Demographers consider nineteenth-century county-level data from Prussia as a unique source of highest-quality data for analysis at a micro-regional level (Galloway et al., 1994). Few of the data used for the 1886 and 1896 cross section were taken from Galloway (2007).

1816 cross-section

The Prussian Statistical Office, founded in 1805, started to publish detailed data at the county and municipality level in 1825. The data contain information from censuses in 1816, 1819, and 1821. The 1816 census provides information on education, landownership, population, demography, and religion. The 1819 census provides data on some occupations and means of production.

The 1816 census classifies landownership in 3 groups: properties or leasehold estates with up to 15 Prussian Morgen (PM), 15 to 300 PM and more than 300 PM. From this we construct landownership concentration as the share of estates with more than 300 PM. The county-level data also report on public elementary schools (Öffentliche Elementarschulen), the only school type equally available in rural areas and towns at the time. In addition, the 1816 census reports school data for the 172 medium and large towns in Prussia, which provide additional information on types of schools available only in towns: private elementary schools (Privat-Elementarschulen), public middle schools for boys or girls (Öffentliche Buerger- und Mittelschulen für Söhne oder Töchter), and private middle schools for boys or girls (Private Bürger- und Mittelschulen für Söhne oder Töchter). Children at recommended school age (6 to 14 years) could either attend elementary schools or middle schools, which had a broader curriculum as well as more grades. To capture all children at recommended school age, county and town enrollment data are aggregated to compute enrollment. The

1816 and 1819 censuses are also used to compute most of the other control variables for the 1816 cross-section.

(i) The share of Protestants is defined as the number of Protestants in 1816 divided by the total population in 1816.

(ii) The share of urban population is defined as the number of people living in cities which held city rights in 1816 divided by the total population in 1816.

(iii) The share of the population in the industrial sector is defined as the number of looms in 1819 over the total population in 1816 (since Prussia was not industrialized in 1816 we choose this as an indicator of proto-industrial occupations).

(iv) The share of the population in agriculture is defined as the number of servants in agriculture in 1819 over the total population in 1816 (no data is available for the total number of people in agriculture for this period).

(v) The child dependency ratio is defined as the ratio of the population younger than 15 years to the population between 15 and 60 years in 1816.

(vi) The population density is defined as the number of people per square kilometer of land area in the county in 1816.

(vii) The school density is defined as the number of schools per square kilometer of land area in the county in 1816.

The source of the 1816 and 1819 census data is Mützell (1823-1825, vol. 5-6).

1849 cross-section

The 1849 census classifies landownership in 5 groups: possessions (*Besitzungen*) with up to 5 PM, 5 to 30 PM, 30 to 300 PM, 300 to 600 PM and more than 600 PM. From this we construct landownership concentration as the share of estates with more than 300 PM. The census also reports education information on public elementary schools (Öffentliche Elementarschulen) and public middle schools for boys or girls (Öffentliche Mittelschulen für Söhne oder Töchter). To calculate enrolment rates we divide the number of attending students in 1849 by the number of children at school age between 6 and 14 in 1849.

The definition of the control variables (i), (ii), (v), (vi), and (vii) is similar to the 1816 census only using data from the 1849 census.

(iii) The share of the population in the industrial sector is defined as the number people working as craftsmen or in factories in 1849 over the total population in 1849.

(iv) The share of the population in agriculture is defined as the number of people subsisting from agriculture in 1849 over the total population in 1849 (The enumerator includes all family members, servants and farm-laborers).

The source of the 1849 census data is Statistisches Bureau zu Berlin (1851-1855, vol. 1-6b).

1864 cross-section

Unfortunately the 1864 census does not provide information on landownership. Thus we access landownership data from the agricultural census in 1858 which classifies landownership in 5 groups: possessions (*Besitzungen*) with up to 5 PM, 5 to 30 PM, 30 to 300 PM, 300 to 600 PM and more than 600 PM. From this we construct landownership concentration as the share of estates with more than 300 PM (Meitzen, 1868).

The 1864 census reports education information on public elementary schools (Öffentliche Elementarschulen), private elementary schools (Privat-Elementarschulen), public middle schools for boys or girls (Öffentliche Mittelschulen für Söhne oder Töchter), and private middle schools for boys or girls (Private Mittelschulen für Söhne oder Töchter). To calculate enrolment rates we divide the number of attending students in 1864 by the number of children at school age between 6 and 14 in 1864.

The definition of the control variables (i), (ii), (vi), and (vii) is similar to the 1816 census only using data from the 1864 census.

Unfortunately the 1864 census is missing occupational information for the city counties (*Stadtkreise*). We thus access data from the 1867 census to construct the occupation controls.

(iii) The share of the population in the industrial sector is defined as the number of people employed in mining and industry (including construction) in 1867 over the total population in 1867 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 16b).

(iv) The share of the population in agriculture is defined as the number of people employed in agriculture, forestry and hunting, and fishing in 1867 over the total population in 1867 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 16b).

(v) The child dependency ratio is defined as the ratio of the population younger than 15 years to the population between 15 and 65 years in 1864.

The source of the 1864 census data is Königliches Statistisches Bureau in Berlin (1861-1934, vol. 10).

1886 cross-section

We access landownership data from the occupation census in 1882 which classifies farms in 6 groups: farms with arable land up to 1 hectare, 1 to 2 ha, 2 to 10 ha, 10 to 50 ha, 50 to 100 ha, and more than 100 ha. From this we construct landownership concentration as the share of farms with more than 100 ha arable land (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 76c).

The education data is accessed from the 1886 census using information on public primary schools (Öffentliche Volksschulen). To calculate enrolment rates we divide the number of attending students in 1886 by the number of children at mandatory school age in 1886 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 101).

(i) The share of Protestants is defined as the number of Protestants in 1880 divided by the total population in 1880 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 66).

(ii) The share of urban population is defined as the number of people living in cities which held city rights in 1880 divided by the total population in 1880 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 61).

(iii) The share of the population in the industrial sector is defined as the number employed in mining and industry in 1882 over the total population in 1882 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 76b).

(iv) The share of the population in agriculture is defined as the number of people employed in agriculture and animal husbandry in 1882 over the total population in 1882 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 76b).

(v) The child dependency ratio is defined as the ratio of the population younger than 19 years to the population between 19 and 70 years in 1882 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 76b).

(vi) The population density is defined as the number of people per square kilometer of land area in the county in 1882 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 76b).

(vii) The school density is defined as the number of schools per square kilometer of land area in the county in 1886 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 101).

The agricultural yields data were chosen to stem from the same years as the education data in 1886. They are defined as the yields of winter wheat, winter rye, summer barley, oats and potatoes per hectare (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 92).

1896 cross-section

We access landownership data from the occupation census in 1895 which classifies farms in 7 groups: farms with arable land up to 0.5 hectare, 0.5 to 2 ha, 2 to 5 ha, 5 to 20 ha, 20 to 100 ha, more than 100 ha, and more than 200 ha. From this we construct landownership concentration as the share of farms with more than 100 ha arable land (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 142b).

The education data is accessed from the 1896 census using information on public primary schools (Öffentliche Volksschulen). To calculate enrolment rates we divide the number of attending students in 1896 by the number of children at mandatory school age in 1896 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 151b).

(i) The share of Protestants is defined as the number of Protestants in 1895 divided by the total population in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 148a).

(ii) The share of urban population is defined as the number of people living in cities which held city rights in 1895 divided by the total population in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 143).

(iii) The share of the population in the industrial sector is defined as the number employed in mining and industry in 1895 over the total population in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 142a).

(iv) The share of the population in agriculture is defined as the number of people employed in agriculture and animal husbandry in 1895 over the total population in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 142a).

(v) The child dependency ratio is defined as the ratio of the population younger than 19 years to the population between 19 and 70 years in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 143).

(vi) The population density is defined as the number of people per square kilometer of land area in the county in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 148a).

(vii) The school density is defined as the number of schools per square kilometer of land area in the county in 1896 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 151b).

The agricultural yields data were chosen to stem from the same years as the education data in 1896. They are defined as the yields of winter wheat, winter rye, summer barley, oats and potatoes per hectare (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 147).

All cross-sections

Two control variables are common to all cross sections since data were unavailable for any other point in time.

(viii) The inheritance dummy is coded using county-level maps of historical inheritance laws from ca. 1900. The dummy takes the value one for counties that predominantly practiced partible inheritance (*Naturalteilung*) and takes the value zero for counties that predominantly practiced non partible inheritance (*Anerbenrecht*) (Sering, 1897-1905).

(ix) First language not German (1890) is defined as the share of people whose mother tongue is not German in 1890 over the total population in 1890 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 121a).

References

Galloway, P. R. (2007): "Galloway Prussia Database 1861 to 1914," www.patrickgalloway.com.

Galloway, P. R., E. A. Hammel, and R. D. Lee (1994): "Fertility Decline in Prussia, 1875-1910: A Pooled Cross-Section Time Series Analysis," *Population Studies*, 48, 135 – 158.

Königliches Statistisches Bureau in Berlin (1861-1934): *Preussische Statistik*, vol. 1-305, Berlin: Statistisches Bureau zu Berlin.

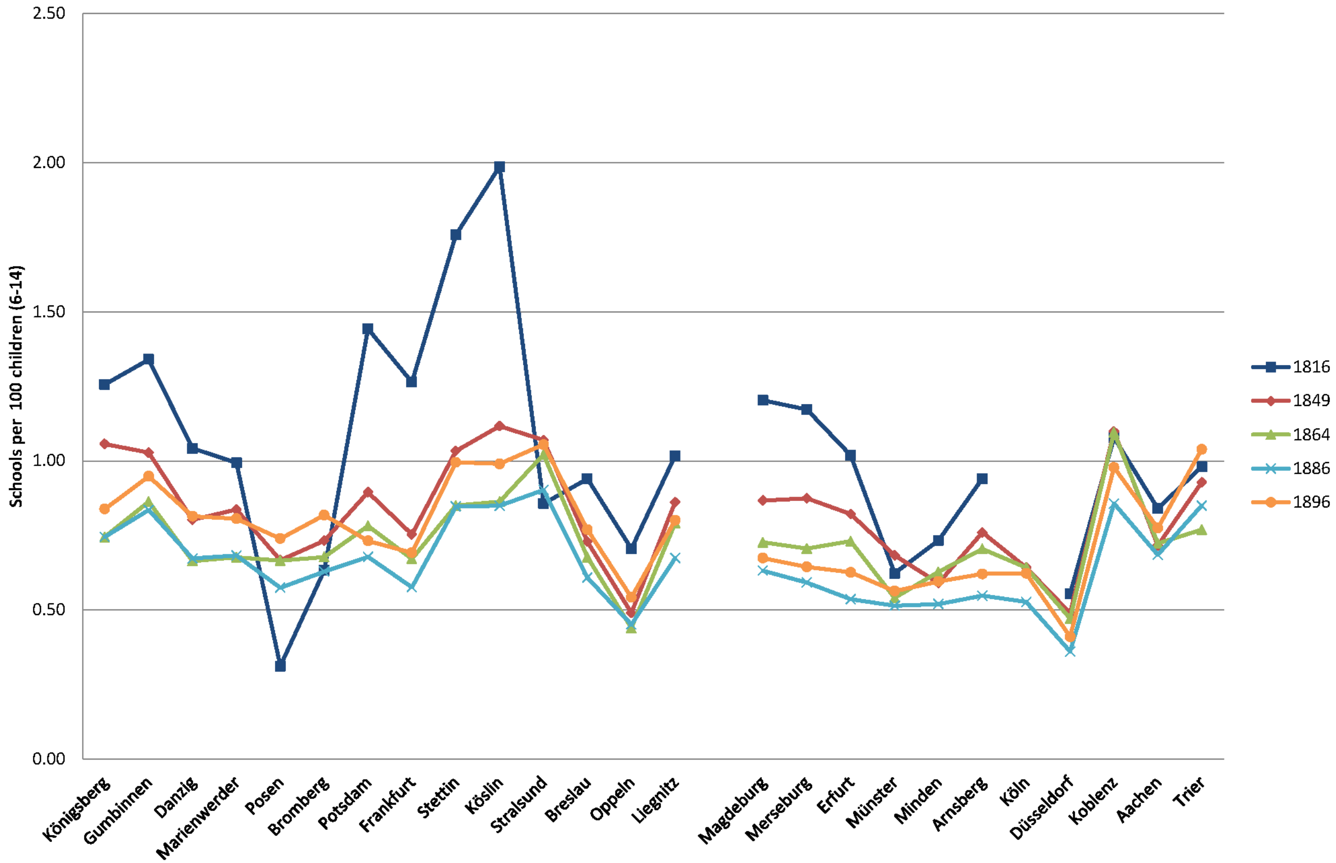
Meitzen, A. (1868): *Der Boden und die landwirthschaftlichen Verhältnisse des Preussischen Staates*, vol. 1-4, Verlag von Paul Parey, Berlin.

Mützell, A. A. (1823-1825): *Neues Topographisch-Statistisch-Geographisches Wörterbuch des Preussischen Staats*, vol. 5 & 6, Karl August Kümmel, Halle.

Sering, M. (1897-1905): *Die Vererbung des ländlichen Grundbesitzes im Königreich Preussen*, vol. 1-14, Paul Parey, Berlin.

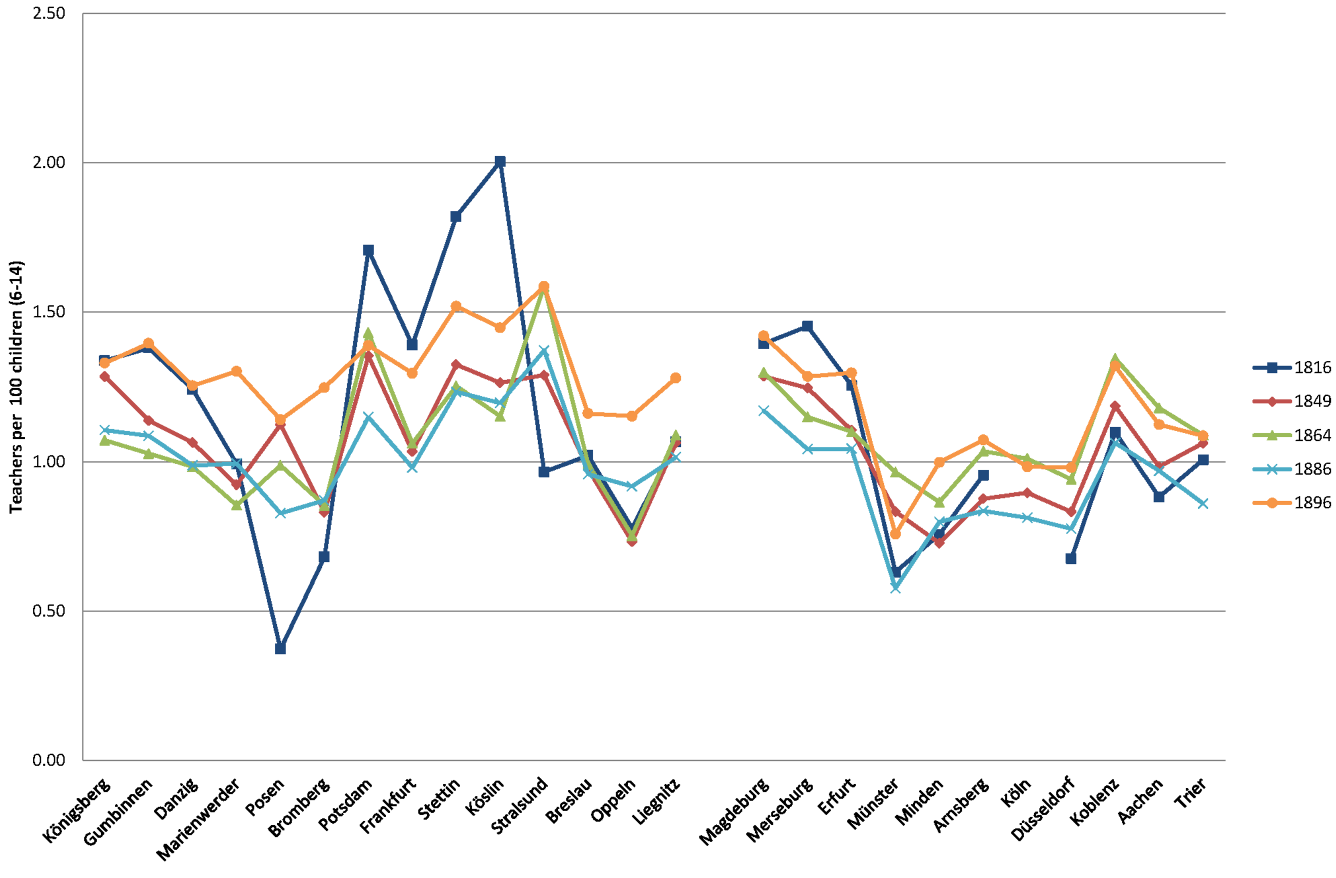
Statistisches Bureau zu Berlin (1851-1855): *Tabellen und amtliche Nachrichten über den Preussischen Staat für das Jahr 1849*, vol. 1-6b, Berlin: Statistisches Bureau zu Berlin.

Figure 1: Schools per 100 children, 1816-96



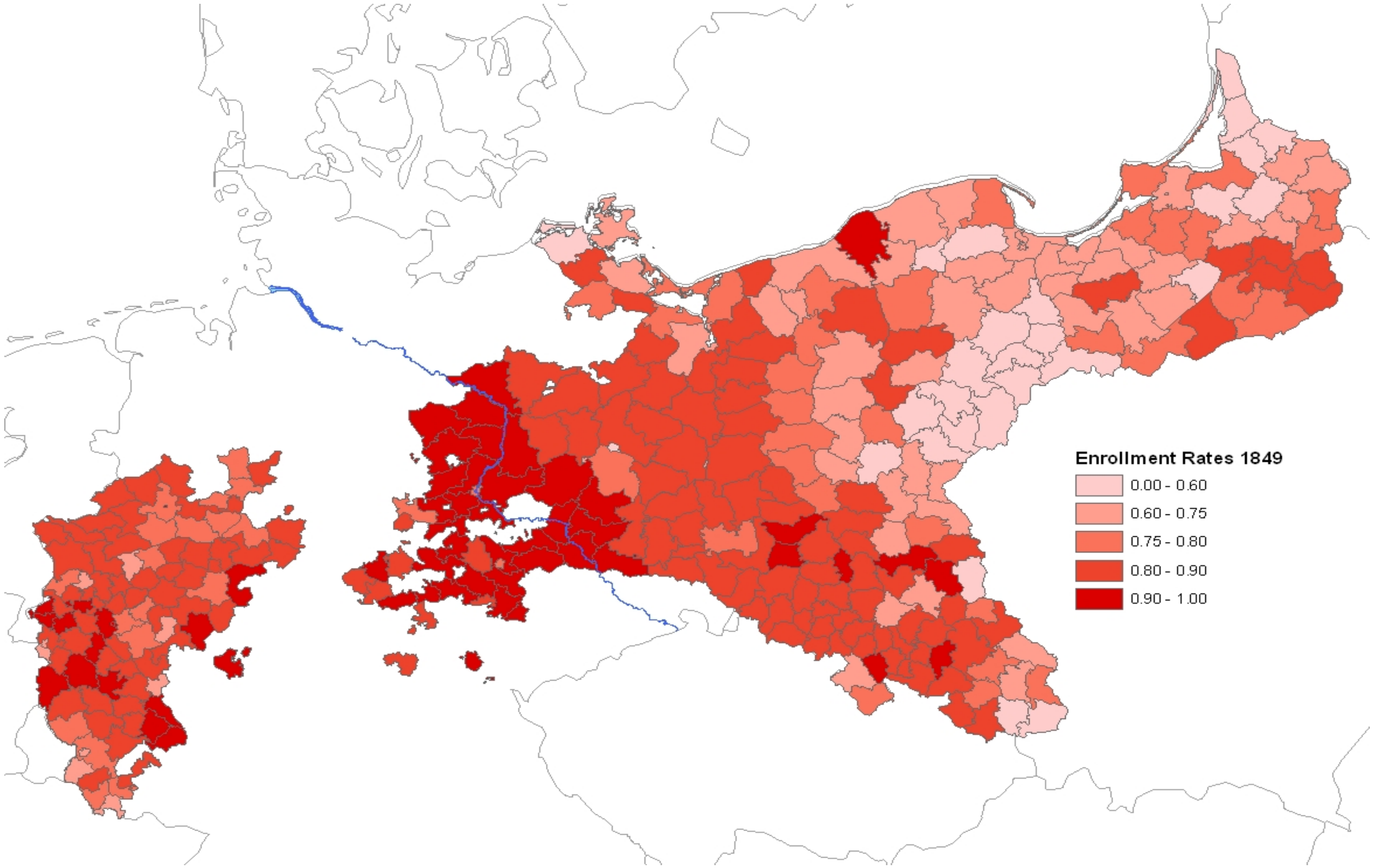
Note: Number of schools per 100 children at school-age (6-14) at the district level.
Districts are ordered approximately from east to west, the gap roughly representing the river Elbe.
Source: See Appendix A.2 for data sources and details.

Figure 2: Teachers per 100 children, 1816-96



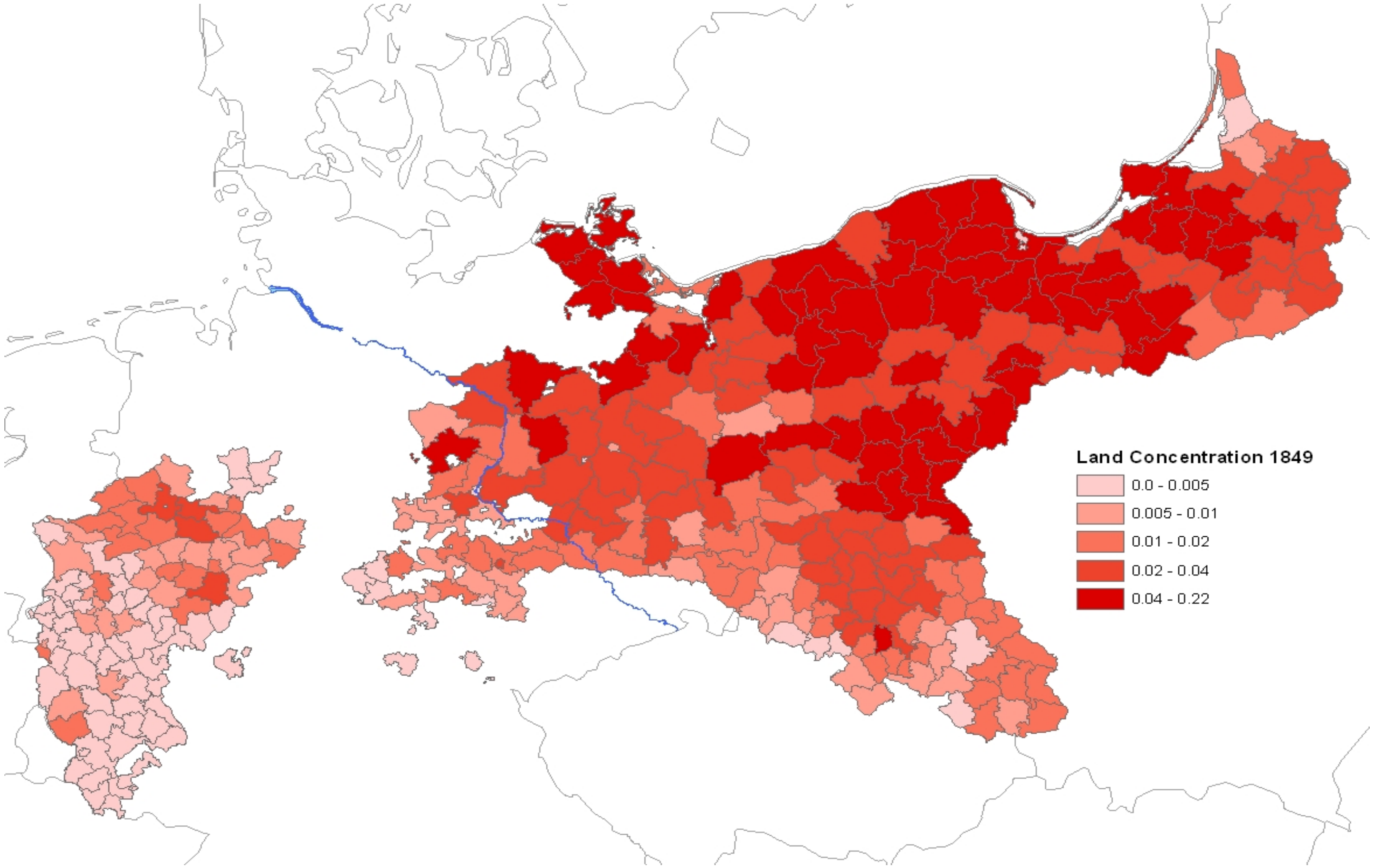
Note: Number of teachers per 100 children at school-age (6-14) at the district level.
Districts are approximately ordered from east to west, the gap roughly representing the river Elbe.
Source: See Appendix A.2 for data sources and details.

Figure 3: Enrollment rates in 1849



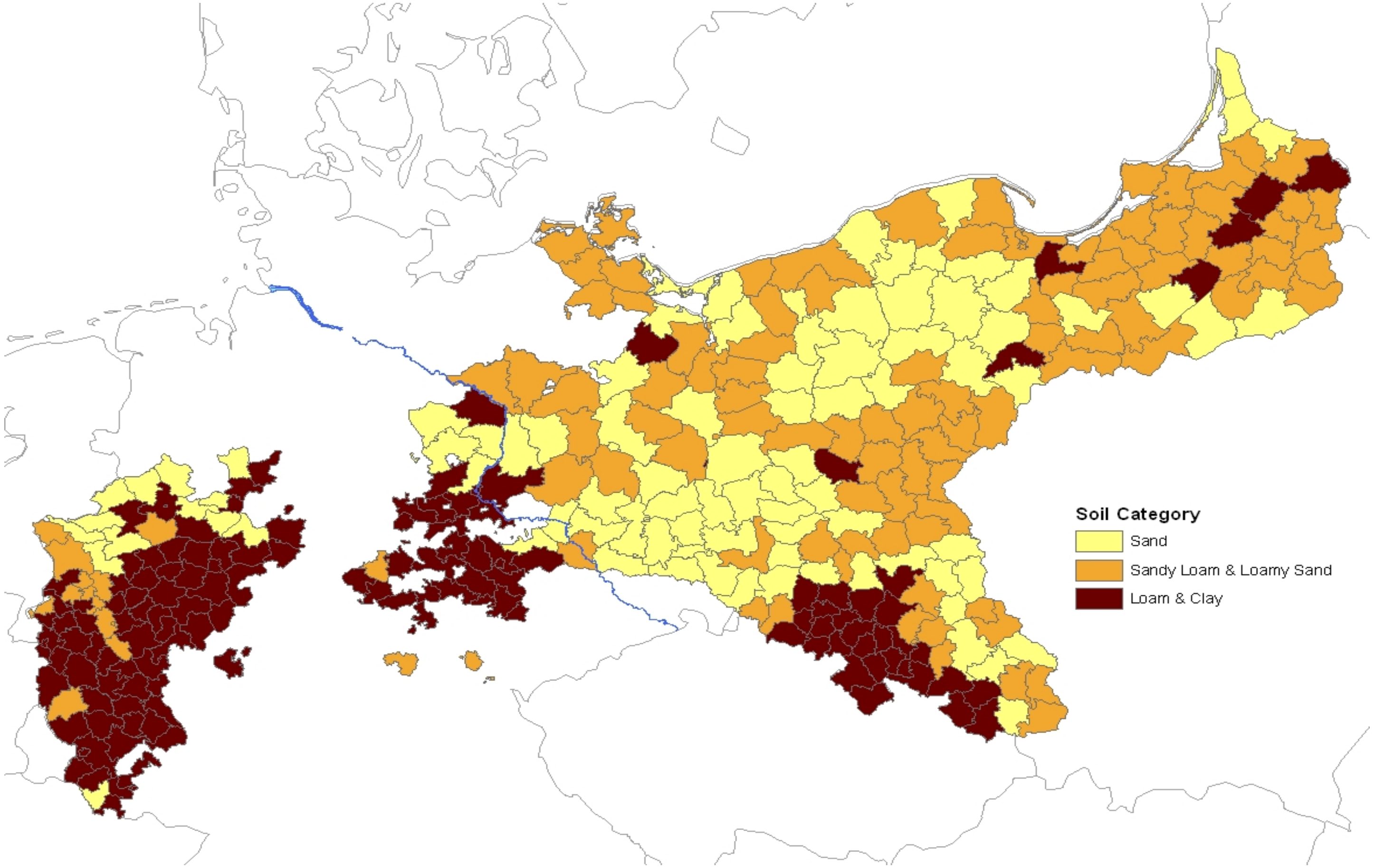
Note: The enrollment rates are constructed as the number of pupils enrolled in primary and middle schools over the number of children aged 6-14.
Source: Own illustration; see main text for details.

Figure 4: Concentration of landownership 1849



Note: The map shows the county structure of Prussia 1849.
The concentration of landownership, measured as the percentage of landowners with more than 300 PM, is roughly classified into quintiles.
Source: Own illustration; see main text for details.

Figure 5: Our Instrument - Soil Texture



Note: The map shows the relative dominance of one of the three soil categories in total soil.
Source: Own illustration; see main text for details.

Table 1: Total expenditure for public primary schools, 1861-86

	1861	1864	1867	1871	1878	1886
Eastprussia	1.414	1.469	1.723	1.781	3.053	3.637
Westprussia	1.458	1.603	1.798	1.915	3.124	3.673
Poznan	1.190	1.253	1.399	1.695	2.687	3.508
Silesia	1.276	1.337	1.517	1.716	2.792	3.488
Pomerania	1.670	1.727	1.945	2.305	3.668	4.533
Brandenburg	1.963	1.953	2.109	2.468	3.515	4.220
Saxony	2.079	2.124	2.233	2.536	3.727	4.678
Westphalia	1.559	1.597	1.812	2.110	3.840	4.742
Rhineland	1.883	1.950	2.124	2.603	4.664	4.991

Note: Total expenditure for public primary schools in German Marks per capita at the province level.
Source: Own calculations according to Königliches Statistisches Bureau in Berlin (1889).

Table 2: State expenditures for public primary schools, 1861-86

	1861	1864	1867	1871	1878	1886
Eastprussia	0.079	0.077	0.095	0.142	0.636	0.742
Westprussia	0.103	0.096	0.111	0.169	0.499	0.667
Poznan	0.085	0.080	0.103	0.141	0.638	0.798
Silesia	0.035	0.038	0.050	0.074	0.325	0.424
Pomerania	0.063	0.052	0.060	0.119	0.807	0.908
Brandenburg	0.111	0.083	0.106	0.129	0.461	0.501
Saxony	0.090	0.068	0.079	0.109	0.356	0.378
Westphalia	0.058	0.044	0.069	0.096	0.392	0.376
Rhineland	0.039	0.041	0.058	0.087	0.397	0.410

Note: State expenditures for public primary schools in German Marks per capita at the province level.
Source: Own calculations according to Königliches Statistisches Bureau in Berlin (1889).

Table 3: Descriptive statistics

VARIABLES	(1) 1816	(2) 1849	(3) 1864	(4) 1886	(5) 1896
Enrollment rate (6-14)	0.603	0.802	0.753	0.935	0.944
	(0.195)	(0.117)	(0.104)	(0.061)	(0.057)
Large landholdings (share)	0.017	0.024	0.025	0.009	0.008
	(0.021)	(0.027)	(0.026)	(0.009)	(0.007)
Gini coefficient	0.646	0.704	0.733	0.742	0.756
	(0.120)	(0.114)	(0.091)	(0.117)	(0.113)
Gini coefficient (adjusted)	0.868	0.867	0.884	0.872	0.878
	(0.073)	(0.122)	(0.108)	(0.088)	(0.082)
Protestant (share)	0.616	0.605	0.600	0.597	0.595
	(0.402)	(0.394)	(0.391)	(0.386)	(0.382)
Urban (share)	0.244	0.246	0.260	0.285	0.305
	(0.182)	(0.186)	(0.194)	(0.202)	(0.207)
Industrial (share)	0.009	0.072	0.080	0.116	0.123
	(0.023)	(0.039)	(0.048)	(0.058)	(0.059)
Agricultural (share)	0.088	0.433	0.186	0.203	0.195
	(0.038)	(0.174)	(0.066)	(0.073)	(0.076)
Child dependency ratio	0.631	0.646	0.602	0.903	0.782
	(0.074)	(0.077)	(0.067)	(0.111)	(0.100)
Population density	0.760	1.774	2.236	2.931	3.228
	(1.855)	(8.430)	(11.238)	(15.633)	(16.600)
School density	0.131	0.177	0.195	0.146	0.150
	(0.263)	(0.730)	(0.756)	(0.276)	(0.231)
Inheritance (dummy)	0.246	0.245	0.245	0.245	0.245
	(0.432)	(0.431)	(0.431)	(0.431)	(0.431)
First language not German (1890)	0.109	0.132	0.132	0.132	0.132
	(0.223)	(0.249)	(0.249)	(0.249)	(0.249)
Observations	272	335	335	335	335

Note: Standard deviations in parenthesis.
Source: See Appendix A.2 for data sources and details.

Table 4: Landownership concentration and enrollment rates (OLS)

	(1)	(2)	(3)	(4)	(5)
Dep Var: Enrollment rates	1816	1849	1864	1886	1896
Large landholdings (share)	-3.195***	-1.329***	-1.181***	-0.616*	-0.708*
	(0.507)	(0.253)	(0.190)	(0.328)	(0.409)
Protestant (share)	0.172***	0.049***	0.040***	-0.004	0.007
	(0.031)	(0.016)	(0.013)	(0.008)	(0.007)
Urban (share)	-0.068	-0.092**	0.047	-0.116***	-0.099***
	(0.063)	(0.039)	(0.056)	(0.022)	(0.020)
Industrial (share)	-0.478**	0.062	0.322	0.619***	0.609***
	(0.233)	(0.142)	(0.203)	(0.111)	(0.081)
Agricultural (share)	0.207	-0.067	0.312*	0.589***	0.550***
	(0.265)	(0.044)	(0.167)	(0.103)	(0.066)
Child dependency ratio	-0.300**	-0.038	-0.180**	0.188***	0.230***
	(0.132)	(0.083)	(0.090)	(0.031)	(0.028)
Population density	-0.064***	-0.008***	-0.004***	-0.002***	-0.002***
	(0.014)	(0.002)	(0.001)	(0.001)	(0.001)
School density	0.407***	0.087***	0.046***	0.116***	0.184***
	(0.087)	(0.015)	(0.010)	(0.034)	(0.052)
Inheritance (dummy)	-0.009	0.019	0.029**	-0.002	-0.008
	(0.026)	(0.012)	(0.012)	(0.007)	(0.007)
First language not German (1890)	-0.271***	-0.163***	-0.140***	-0.057***	-0.010
	(0.059)	(0.030)	(0.025)	(0.012)	(0.009)
Constant	0.769***	0.892***	0.782***	0.611***	0.597***
	(0.099)	(0.063)	(0.089)	(0.060)	(0.042)
Observations	272	335	335	335	335
R-squared	0.49	0.38	0.40	0.58	0.63

Note: OLS estimates at the county level. Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.
Source: See Appendix A.2 for data sources and details.

Table 5: Instrumenting landownership concentration with soil texture

	(1) 1816	(2) 1849	(3) 1864	(4) 1886	(5) 1896
First stage	Dep Var: Landownership concentration				
PCA component 1	0.003***	0.003***	0.002***	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
PCA component 2	0.004**	0.003**	0.003***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Second stage	Dep Var: Enrollment rates				
Large landholdings (share)	-8.858***	-3.683***	-2.281**	-1.943*	-1.478
	(2.183)	(1.129)	(1.037)	(1.054)	(1.165)
Protestant (share)	0.251***	0.081***	0.050***	0.003	0.012
	(0.043)	(0.021)	(0.014)	(0.009)	(0.009)
Urban (share)	0.045	-0.050	0.092	-0.111***	-0.094***
	(0.119)	(0.048)	(0.076)	(0.023)	(0.022)
Industrial (share)	-1.222**	-0.395	0.257	0.521***	0.572***
	(0.475)	(0.263)	(0.213)	(0.122)	(0.096)
Agricultural (share)	0.584*	-0.014	0.505**	0.582***	0.559***
	(0.299)	(0.051)	(0.237)	(0.103)	(0.068)
Child dependency ratio	-0.091	0.100	-0.034	0.201***	0.235***
	(0.167)	(0.128)	(0.160)	(0.033)	(0.030)
Population density	-0.084***	-0.009***	-0.003***	-0.002**	-0.002***
	(0.020)	(0.002)	(0.001)	(0.001)	(0.001)
School density	0.524***	0.096***	0.040***	0.112***	0.187***
	(0.123)	(0.017)	(0.014)	(0.035)	(0.054)
Inheritance (dummy)	-0.050	-0.013	0.022*	-0.009	-0.012
	(0.031)	(0.019)	(0.013)	(0.008)	(0.008)
First language not German (1890)	-0.184***	-0.153***	-0.144***	-0.054***	-0.008
	(0.065)	(0.032)	(0.026)	(0.012)	(0.009)
Constant	0.629***	0.848***	0.675***	0.619***	0.598***
	(0.122)	(0.078)	(0.129)	(0.060)	(0.042)
Observations	272	335	335	335	335
Kleibergen-Paap F statistic	8.236	10.37	9.521	14.92	13.69

Note: IV estimates instrumenting landownership concentration with soil texture. Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.
Source: See Appendix A.2 for data sources and details.

Table 6: Robustness check – Agricultural productivity

	(1) 1816	(2) 1849	(3) 1864	(4) 1886	(5) 1896
First stage	Dep Var: Landownership concentration				
PCA component 1	0.003***	0.004***	0.002**	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
PCA component 2	0.003***	0.003**	0.004***	0.002***	0.002***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
Second stage	Dep Var: Enrollment rates				
Large landholdings (share)	-9.016***	-3.308***	-2.357**	-1.166	-0.442
	(2.515)	(1.118)	(0.976)	(0.948)	(1.068)
Wheat yield per hectare	0.014	0.008**	0.007**	0.000	0.003**
	(0.011)	(0.004)	(0.003)	(0.002)	(0.001)
Rye yield per hectare	-0.010	-0.015**	-0.010***	-0.000	-0.004**
	(0.012)	(0.006)	(0.004)	(0.002)	(0.002)
Barley yield per hectare	0.015*	0.005	0.001	-0.004**	-0.002
	(0.009)	(0.005)	(0.004)	(0.002)	(0.002)
Oat yield per hectare	-0.023	-0.001	-0.001	0.002	0.001
	(0.015)	(0.005)	(0.005)	(0.002)	(0.002)
Potato yield per hectare	0.001	0.001**	0.001**	0.000	0.000
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
Observations	269	329	329	329	327
Kleibergen-Paap F statistic	6.357	9.611	9.321	15.33	13.47

Note: IV estimates instrumenting landownership concentration with soil texture. Robust standard errors in parentheses. Constant omitted. Additional controls: % protestant, % urban, % industrial, % agricultural, child dependency ratio, population density, school density, inheritance, first language not German (1890). In columns 1-4 yields refer to 1886; in column 5 yields refer to 1896. Significance: *** p<0.01, ** p<0.05, * p<0.1.
Source: See Appendix A.2 for data sources and details.

Table 7: Instrumenting landownership concentration with soil texture (excluding city counties)

	(1) 1816	(2) 1849	(3) 1864	(4) 1886	(5) 1896	(6) 1849 rural
First stage	Dep Var: Landownership concentration					
PCA component 1	0.003***	0.002**	0.002**	0.000	0.000	0.004***
	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)
PCA component 2	0.004**	0.004***	0.004***	0.002***	0.002***	0.005***
	(0.002)	(0.001)	(0.001)	(0.000)	(0.000)	(0.002)
Second stage	Dep Var: Enrollment rates					
Large landholdings (share)	-7.652***	-3.644***	-2.422**	-1.441	-1.005	
	(1.841)	(1.155)	(0.984)	(0.886)	(0.921)	
Large landholdings (share in rural)						-2.642***
						(0.844)
Protestant (share)	0.195***	0.055**	0.029**	0.002	0.009	0.048**
	(0.043)	(0.022)	(0.014)	(0.008)	(0.008)	(0.021)
Urban (share)	0.105	0.029	0.173**	-0.113***	-0.090***	0.076
	(0.101)	(0.063)	(0.074)	(0.025)	(0.020)	(0.069)
Industrial (share)	-1.036**	-0.300	0.708***	0.464***	0.499***	-0.206
	(0.512)	(0.229)	(0.133)	(0.120)	(0.092)	(0.204)
Agricultural (share)	0.684**	-0.073	0.434**	0.540***	0.459***	-0.090*
	(0.275)	(0.050)	(0.219)	(0.107)	(0.060)	(0.050)
Child dependency ratio	-0.073	0.104	0.014	0.171***	0.206***	0.103
	(0.153)	(0.123)	(0.161)	(0.032)	(0.027)	(0.117)
Population density	-0.184***	-0.109***	-0.082***	0.001	-0.003	-0.116***
	(0.050)	(0.027)	(0.011)	(0.007)	(0.005)	(0.027)
School density	1.453***	0.757***	0.181	0.144	0.161*	0.784***
	(0.374)	(0.251)	(0.135)	(0.122)	(0.098)	(0.233)
Inheritance (dummy)	-0.049*	-0.017	0.026**	-0.007	-0.009	-0.009
	(0.030)	(0.019)	(0.013)	(0.008)	(0.008)	(0.017)
First language not German (1890)	-0.163**	-0.138***	-0.136***	-0.052***	-0.011	-0.151***
	(0.067)	(0.031)	(0.025)	(0.013)	(0.009)	(0.030)
Constant	0.568***	0.863***	0.668***	0.652***	0.651***	0.847***
	(0.119)	(0.071)	(0.120)	(0.058)	(0.039)	(0.069)
Observations	267	325	325	324	324	325
Kleibergen-Paap F statistic	11.51	10.48	9.658	15.39	17.56	11.40

Note: IV estimates instrumenting landownership concentration with soil texture.
Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1.
Source: See Appendix A.2 for data sources and details.

Table 8: Gini coefficient and enrollment rates (OLS)

	(1)	(2)	(3)	(4)	(5)
Dep Var: Enrollment rates	1816	1849	1864	1886	1896
Gini coefficient	-0.074	-0.026	0.042	0.003	0.039
	(0.110)	(0.064)	(0.070)	(0.033)	(0.030)
Observations	272	335	335	335	335
R-squared	0.40	0.32	0.35	0.58	0.63
Gini coefficient (adjusted)	-0.500**	-0.154**	-0.077	-0.050	0.003
	(0.206)	(0.065)	(0.068)	(0.034)	(0.033)
Observations	272	335	335	335	335
R-squared	0.42	0.33	0.35	0.58	0.63

Note: OLS estimates at the county level. Robust standard errors in parentheses. The adjusted Gini accounts for propertyless farm workers, see text. Constant omitted. Additional controls: % protestant, % urban, % industrial, % agricultural, child dependency ratio, population density, school density, inheritance, first language not German (1890). Significance: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.
Source: See Appendix A.2 for data sources and details.

Table 9: Panel analysis

Dep Var: Enrollment rates	(1) Pooled	(2) FE	(3) TFE	(4) Pooled	(5) FE	(6) TFE
Large landholdings (share)	-2.618***	-2.163***	-1.038**			
	(0.262)	(0.476)	(0.455)			
Gini coefficient (adjusted)				-0.249***	-0.131**	-0.167***
				(0.034)	(0.055)	(0.058)
Protestant (share)	0.100***	0.578**	0.505**	0.090***	0.523**	0.486**
	(0.009)	(0.234)	(0.199)	(0.010)	(0.235)	(0.200)
Urban (share)	-0.099***	0.004	0.019	-0.088***	0.025	0.025
	(0.026)	(0.067)	(0.065)	(0.025)	(0.064)	(0.063)
Industrial (share)	1.184***	1.310***	-0.351**	1.408***	1.334***	-0.401**
	(0.074)	(0.108)	(0.159)	(0.074)	(0.106)	(0.161)
Agricultural (share)	0.361***	0.302***	0.003	0.271***	0.256***	-0.040
	(0.030)	(0.024)	(0.042)	(0.030)	(0.024)	(0.042)
Child dependency ratio	0.307***	0.401***	0.254***	0.333***	0.479***	0.271***
	(0.019)	(0.035)	(0.073)	(0.020)	(0.033)	(0.074)
Population density	-0.001*	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)
School density	0.029***	0.026**	0.023*	0.020**	0.021	0.016
	(0.008)	(0.013)	(0.012)	(0.008)	(0.014)	(0.013)
Time dummies	no	no	yes	no	no	yes
Observations	1387	1387	1387	1387	1387	1387
R-squared	0.52	0.59	0.70	0.48	0.58	0.70
Number of counties	280	280	280	280	280	280

Note: Panel estimates: FE indicates county fixed effects, TFE adds time fixed effects. Robust standard errors in parentheses. The adjusted Gini accounts for propertyless farm workers, see text. 13 observations drop out from the analysis because of missing information in the 1816 data. Significance: *** p<0.01, ** p<0.05, * p<0.1.
Source: See Appendix A.2 for data sources and details.

European Historical Economics Society

EHES Working Paper Series

Recent EHES Working Papers

2011

All papers may be downloaded free of charge from: www.ehes.org