DISCUSSION PAPER SERIES

No. 9730

LANDOWNERSHIP CONCENTRATION AND THE EXPANSION OF EDUCATION

Francesco Cinnirella and Erik Hornung

DEVELOPMENT ECONOMICS, ECONOMIC HISTORY and PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9730.php**

ISSN 0265-8003

LANDOWNERSHIP CONCENTRATION AND THE EXPANSION OF EDUCATION

Francesco Cinnirella, Ifo Institute, CESifo, CAGE and CEPR
Erik Hornung, Max Planck Institute for Tax Law and Public Finance

Discussion Paper No. 9730
November 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

CEPR Discussion Paper No. 9730

November 2013

ABSTRACT

Landownership Concentration and the Expansion of Education*

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 serfdom, has a negative effect on schooling. This effect diminishes substantially towards the end 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 present several robustness checks which shed some light on possible mechanisms.

JEL Classification: I25, N33, O43 and Q15
Keywords: education, institutions, land concentration, peasants' emancipation, Prussian economic history and serfdom

Francesco Cinnirella
Ifo Institute for Economic Research
Poschingerstrasse 5
81679 Munich
GERMANY

Email: cinnirella@ifo.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=168123

Erik Hornung
Max Planck Institute
Marstallplatz 1
80539 Munich
GERMANY

Email: erik.hornung@tax.mpg.de

For further Discussion Papers by this author see:
www.cepr.org/pubs/new-dps/dplist.asp?authorid=175851

*We gratefully acknowledge comments from Brian A'Hearn, Sascha O. Becker, Davide Cantoni, Greg Clark, Oliver Falck, Oded Galor, Mike Hoy, Marc Piopiunik, Florian Ploeckl, Ludger Woessmann and 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. An earlier version of this paper is a chapter of Erik Hornung's PhD dissertation. We are grateful for financial support from the Pact for Research and Innovation of the Leibniz Association. Francesco Cinnirella is also grateful for financial support from the Fritz Thyssen Stiftung.

Submitted 04 November 2013

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 the population being 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 these areas 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 us 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

[1] However, we will also provide alternative estimates of the relationship between landownership inequality and education using the Gini coefficient.

that large landowners' opposition to mass education and the condition of limited freedom that characterized the peasantry can explain the strong negative effect during the first half of the nineteenth century; the fading effect of large landownership in the second half of the nineteenth century is consistent with mechanisms such as more capitalistic agricultural production, larger stakes of the landed nobility in the industrialization process (Galor et al., 2009), and the emancipation of the peasantry.

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 that is extensively investigated in agricultural economics, and 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 also introduce controls for land productivity in different crops. Instrumental variable 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 us 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 and the effect of landownership concentration declines over time. Further panel specifications demonstrate that the eastern regions, treated by the agricultural reforms, are responsible for the declining effect of landownership over time.

Prussia provides an interesting case, since it is representative of the different agricultural institutions that characterized the eastern and the western parts of Europe. Historically, the 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 abolition 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 in which to study the effect of different institutions on education. Our variable for landownership concentration aims to capture such variation in local institutions.

Section 2 reviews the related literature; Section 3 provides the theoretical and historical background emphasizing the Prussian reforms in agriculture and education; Section 4 describes the data; Section 5 introduces the model and presents OLS estimates; Section 6 addresses the issue of causality showing instrumental variable estimates and several robustness checks; Section 7 presents estimates of panel models; Section 8 discusses possible mechanisms; Section 9 concludes.

2 Related literature

The literature on the long-run economic consequences of inequality for 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]

Similarly, Ramcharan (2010) studies the effect of economic inequality on redistributive policy. The author studies the relationship in the United States between

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

land inequality (measured by the Gini coefficient) and education expenditure using census data, at the county level, for the period 1890-1930. The paper shows that greater inequality is strongly associated with less redistribution and therefore with less expenditure on education. Interestingly for our case, the author shows that using a measure for land concentration similar to ours, 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 for the United States using variation in the distribution of landownership and educational expenditure, across states and over time, for the period 1900-1940. Consistent with their theoretical predictions, they find that greater 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. Kourtellos et al. (2013) also test the hypothesis in a panel cross section of countries finding that higher levels of land inequality delay the extension of schooling.

Go and Lindert (2010) explain differences in enrollment rates between the North and the South by analyzing US counties in 1850. They point to local governments having more autonomy and the population having a more equally distributed political voice in the North. Among other things, they test whether extending the voting power to lower-income groups raises the taxes paid for schooling and thus the enrollment rates in primary schools. Similarly, Vollrath (2012) 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. By estimating the

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

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

correlation between schooling and literacy rates with inequality in political power, Mariscal and Sokoloff (2000) show how the extension of the franchise in Latin America increased schooling. 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. Aghion et al. (2012) argue that democratization does not play an important role and suggest instead military rivalry as an important factor behind countries' decision to invest in mass primary education.

Using agricultural endowments as an instrument for inequality, Easterly (2007) tests the hypothesis of Engerman and Sokoloff (1997). 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 rates 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 us to study how this relationship evolved throughout the nineteenth century, a period characterized by important institutional changes.

3 Theoretical and historical background

Our paper fits into the literature that emphasizes the role of inequality for public education. According to Sokoloff and Engerman (2000), inequality in the distribution of income indirectly affects education through the quality of institutions, which in turn have a direct effect on public schooling. Galor et al. (2009) instead claim that inequality in the distribution of landownership adversely affects expenditure in public schooling, due to low complementarities between agriculture and education. In agrarian economies, landowners can influence the political process and therefore might affect the implementation of public policies such as the expansion of education. In fact, Prussia is not well suited to test the model of Galor et al. since (as explained in subsection 3.2), large landowners were exempted from the payment of

taxes. In our case the concentration of landownership serves more as a proxy for the extent of serfdom.[5]

Similar to the model of Galor et al., our framework assumes that education increases the productivity of workers. Yet, analogous to the work of Lagerlöf (2009) which models the long-run evolution of property rights, we assume that under serfdom (which in Lagerlöf's model is a mild form of slavery) the landed elite 'owns' part of the peasants' labor and they pay these workers less than their marginal product.[6] Under these conditions landowners oppose the spread of education among the peasantry and the serfs themselves have few incentives to invest in education.[7]

With the abolition of serfdom and the emancipation of the peasants, the workers can regain full possession of their labor and (or) choose a different occupation where they can earn the competitive market wage. At the same time the landed nobility changes its attitude towards a more capitalistic agricultural production and increases its stakes in the industrialization process. In this framework we expect to find an increase in investments in education and a lessened adverse effect of landownership concentration on education.

3.1 Landownership and land reform

Prussian 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 the 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 widely within these areas (Rosenberg, 1944). The amount of servile dues and services usually depended on the attributes of the contracted land, and were heaviest in the regions with large noble demesnes (Melton, 1988, p 149). Most of the peasants were serfs who were not allowed to relocate without compensation payment and the permission of the lord. Becoming the legal owner of the land they cultivated was rarely possible.

Servile 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

[5] It is also important to note that our study focuses on enrollment rates and not on education expenditure.

[6] Note that the model of Lagerlöf does not consider investments in human capital.

[7] The model of Lagerlöf (2009) also predicts that, ceteris paribus, low population density makes slavery (that is, serfdom) more likely, compared to free labor. This is entirely consistent with nineteenth century Prussia where serfdom was most prevalent in the East, which is characterized by a relatively low population density.

servants, a boy, and four horses to work, daily, on the lord's demesne. 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' households. In the provinces of Pomerania and Silesia, Junkers drafted the most promising and capable children of the peasantry to work as full-time servants (Carsten, 1988, p 63). Additionally, children were also burdened with the corvée (here: *Scharwerk*) which, as the parish priest of Dombrowken (a village in East Prussia) complained in 1773, was so much work that children were kept from attending school even in the winter (Carsten, 1988, p 65).[8]

In the eastern parts of Prussia, serfdom was abolished for the first time towards the end of the eighteenth century on the royal domains. After the defeat by Napoleon, a series of agricultural reforms was initiated which freed agricultural workers at a piecemeal rate. The 'Oktoberedikt' of 1807 made all peasants free people, starting from November 11th (St Martin's Day) 1810. In 1811, the reformer Hardenberg passed the Edict of Regulation on landownership according to which peasants gained the legal right to own landed property of any type.[9] This reform applied to peasants with weak rights of ownership, i.e. non-heritable tenancy or cancelable tenancy, who became able to redeem the land by ceding one-third or respectively up to half of the land to their landlord.[10] Due to the opposition of large landholders to the Edict, its scope was restricted by excluding peasants on small parcels in the Declaration of 1816.

Much more burdensome than weak rights of legal landownership were the servile duties. The landlords demanded that any reform of this labor-extracting institution needed to clarify the amount of compensation for foregone duties. Hence, the agricultural reforms did not gain momentum, also due to the ongoing war on Napoleon, until the Dissolution Edict of 1821.[11] This edict allowed peasants to redeem servile duties by compensating the manorial lord for forgone earnings with the amount of 25 times their annual value in either money or land. However, the release from servile duties only applied to peasants with strong rights of ownership. The emancipation of the peasantry with weak rights thus resembled a two step process where peasants

[8] Henning (1969) compares the servile duties in regions of East Prussia and Westphalia in the eighteenth century. Even within East Prussia there is a lot of variation where 77 percent of peasants in the *Oberland*, 4 percent in the *Binnenland*, and 8 percent in the *Küstengebiet* had to provide between 251 and 500 days of manorial service per year. In Westphalia, roughly 30 per cent of the peasants had to provide manorial services for 26 to 52 days per year; roughly 50 percent had to work 9 to 15 days per year, and the rest provided even fewer services, whenever required.

[9] "Edikt, die Regulierung der gutsherrlichen und bäuerlichen Verhältnisse betreffend."

[10] Because of this practice, and due to the dissolution of the common lands, the average farm size of the large landholders grew strongly until the middle of the nineteenth century (Schiller, 2003).

[11] "Ordnung wegen Ablösung der Dienste, Natural- und Geldleistungen von Grundstücken, welche eigenthümlich, zu Erbzins- oder Erbpachtsrecht besessen wurden."

first had to redeem land to become legal owners and then had to redeem servile duties (Harnisch, 1984, p 154). Thus, by 1848 many peasants still claimed that they had no personal freedom, since it did not matter if they had to fulfill their duties either as serf-tenants or as owners of the land (Carsten, 1988, p 98).

Excluded from both edicts were peasants on small parcels who thus remained obliged to provide servile duties. This limitation reduced the share of peasants affected by the reform to one third of the peasantry. 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[12] that all peasants ultimately gained complete legal emancipation and were able to redeem small parcels as well as servile duties (Pierenkemper and Tilly, 2004; Bowman, 1980).

During the Napoleonic occupation between 1794 and 1814, the western parts, which belonged to Prussia after 1815, introduced the Napoleonic Code which basically meant the abolition of feudal privileges. Serfdom was therefore abolished *de jure* in Rhineland in 1805 and in Westphalia in 1808. Redemption of servile duties was based on the payment of 15 times their annual value in Rhineland, and 25 times their 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 west of the River Rhine, all other areas (re-)established Prussian law (*Allgemeines Landrecht*).[13]

3.2 Education reform and school financing

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 expense (Lindert, 2003, 2004). The war against France triggered a 'decade of feverish activity' (Schleunes, 1979, p. 317) during which the expansion of schooling grew very quickly. Wilhelm von Humboldt, head of cultural and educational affairs at the time, favored a general education over specialization and advocated a unified schooling system, which would have educated everybody equally.[14]

[12] "Gesetz, betreffend die Ablösung der Reallasten und die Regulierung der gutsherrlichen und bäuerlichen Verhältnisse."

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

[14] 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, as it often happens in life, to change from one to another (own translation from von Humboldt (1964, p. 218)).

The composition of resources 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 such as estates, entitlements, or capital rents. Assets were managed by so-called schooling societies, which were bodies of the municipality that received additional taxes from the heads of each household (excluding noble landowners) proportional to their wealth and income, and independent of their religious denomination and number of children. When 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 that 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 smaller school funds compared to western provinces. For example, primary school funds amounted to about 1.4 Marks per capita in East and West Prussia in 1861,[15] whereas it amounted to 1.8 Marks in the western province of Rhineland. However, it is important to note that this difference may reflect differences in the cost of living. If we deflate school funds per capita with rye prices in 1861, funds in the provinces of East and West Prussia are 10 and 14 percent higher than in Rhineland, respectively.[16]

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 East Prussia, West Prussia, and Posen benefited to a larger extent from State contributions than did the western provinces of Westphalia and Rhineland. Also,

[15] Note that both provinces of East and West Prussia belong to the East Elbe part of the Kingdom of Prussia.
[16] The gap between school funds in East and West Prussia also decreases if we deflate by wheat prices in 1861.

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 per cent of the total school funds, respectively; instead, in East and West Prussia tuition fees accounted for only 11 and 13 percent of the total school funds, respectively.

Thus, the aggregate data show that school funds in eastern areas were relatively smaller in nominal terms with respect to the richer and more urban regions of the west, though in real terms the gap might have been non-existent. 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 Figure 2 we plot the number of schools and the number of teachers per 100 children (6 to 14) for 25 districts by year (1816, 1849, 1864, 1886, and 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. In section 8 we discuss this issue more in detail.

Therefore, the lower (nominal) level of financial support in the rural eastern areas arguably neither had an impact on the availability of schools nor of teachers.[17] Supply of education was not restrained in the eastern regions of Prussia.[18]

4 Data

4.1 Prussian census data

In order to test whether the concentration of large landownership delayed the expansion of education in Prussia, we collected county-level census data. Our data refer to five points in time (1816, 1849, 1864, 1886, and 1896) spanning the entire nineteenth century and they come from various sources published on behalf of the Royal Prussian Statistical Bureau in Berlin.[19] 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 us to study how the effect of landownership concentration varied during the

[17] Of course we cannot exclude that it affected their quality.
[18] See Chaudhary (2009) about the role of private and public spending on education in India.
[19] See Appendix A.2 for more details about data sources.

nineteenth century. However, we are able to test our hypothesis by organizing the data as a panel, thus accounting for unobserved heterogeneity.

Prussian county-level data are challenging because of several changes of 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. For 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 have enrollment rates for the years 1816, 1849, 1864, 1886, and 1896. Primary school enrollment rates represent school attendance of the 6 to 14 year-old.[20]

Data on landownership by size are 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.[21] The first full census in 1816 classified landholdings in three groups: properties or leasehold estates up to 15 Prussian Morgen (henceforth PM), from 15 to 300 PM, and more than 300 PM.[22] 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 three to five bins. From 1882 onwards, the census considered only the arable land which was classified into six bins, increasing to seven in the 1895 census. Unfortunately, the total area

[20] 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 per cent. This could be due to children commuting from neighboring counties or because of children older than 14 years being enrolled in school.

[21] It is important to note that the original census data contain 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 (or landholders), since they might own (or 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 us to identify the landowner. See Eddie (2008) for a discussion of farm statistics and property statistics.

[22] One PM is equal to circa 0.25 hectare.

of all holdings by bin category was not published until 1882, so the calculation of conventional measures of inequality is not straightforward.[23]

Our measure of landownership concentration is the number of landholdings larger than 300 PM (circa 75 hectares) 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.[24] Our indicator is therefore bounded between 0 and 1: 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 relied greatly on serf labor, and where peasants' freedom and mobility were 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 from 1858 on the average size of a farm by category. This information is not 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 level 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.[25]

Our covariates aim to control for different aspects of the demand and supply of education. We include the share of the Protestant population—as Protestants are expected to have a higher demand for education—the share of the urban population, the share of the population employed in the industrial sector,[26] the share of the population employed in agriculture, and population density.[27] The age structure of the county might also influence the demand for education.[28] In our case, this factor is proxied by the child dependency ratio, calculated as the share of the young

[23] 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 might also capture one dimension of political inequality.

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

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

[26] 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.

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

[28] About the relationship between fertility and education in pre-demographic transition Prussia see Becker et al. (2010, 2012, 2013).

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.[29] This variable controls for differences between the former Polish parts of Prussia (the provinces of Poznan, Silesia, and 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 issue might have an impact on our analysis. In fact, inheritance laws follow a geographic 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 the value 1 in counties with partible inheritance.

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 to 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 percent in 1816 to 94 percent in 1895. Interestingly, the enrollment rate in 1864 is lower than at the previous census in 1849.[30]

The variable *large landholdings* indicates the share of the largest holdings in the county. In 1816, 1.7 percent of 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 percent in 1858.[31] This evidence is consistent with accounts of the agrarian reform in Prussia, which stress how the landed elite benefited most from the reforms (Schiller, 2003).

[29] Unfortunately, the earliest available census data about spoken languages are from 1890.
[30] This peculiarity has already been found by Lindert (2004, p. 91, Tab. 5.1).
[31] The year 1864 in the table headings refers to the date for which data on enrollment rates are available.

As mentioned already, a direct comparison of 1882 with 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.[32] This explains why our concentration index decreases to 0.8 percent in 1882 and to 1 percent in 1895.

The progress of industrialization is shown by the increase of urbanization which rises from 24 percent to 30 percent, and industrial employment which increases from 1 percent to 12 percent. 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 Figure 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 they 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 diametrical to Figure 3, as areas with lower enrollment rates are now characterized by a relatively high concentration of large holdings. Western areas are strongly characterized by small holdings, though western counties also 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 first present simple correlations for the five periods (Table 4). We immediately observe a strong negative correlation, which weakens towards the end of the nineteenth century and becomes insignificant by 1896.

Successively, 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 county i in a given year t:

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

[32] 100 hectares are equal to 392 PM.

β_1 is the coefficient of interest. The vector X contains the share of Protestants in the county, the share of the 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 the value 1 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 5. We consistently find a significant negative association between the share of large landholdings and the enrollment rate. This relationship, very large in 1816, seems to fade throughout the nineteenth century. As already mentioned, it is possible to compare the magnitudes of the coefficients for 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 wrong conclusions. Since the dependent variable is theoretically bounded between 0 and 1, 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 coefficients also in terms of standard deviations.

In fact, the estimated relationships are quite large. In 1816, an increase in the share of large estates by 1 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 1 standard deviation in landownership concentration is associated with a decrease of the county enrollment rate by 0.09 standard deviations.[33]

As already shown in Becker and Woessmann (2009), we find 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 is highly significant from 1886. This result is consistent with the study of Becker

[33]To further account for the fact that the dependent variable has an upper bound, we also controlled for "initial conditions" and included enrollment rates in 1816 in the specifications for 1849, 1864, 1886, and 1896. The results indicate the same pattern as in Table 5 (estimates available upon request).

et al. (2011) who argue that, differently from Britain, education had a causal effect on industrialization in Prussia.

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

It is important to note the significant positive relationship between school density and enrollment rates throughout the whole period. This result suggests that the supply of schools was not the mechanism through which large landowners opposed the spread of mass education (see Section 8 for a detailed discussion). If that was the case, the coefficient β_1 attached to large landholdings should decrease when controlling for school density. As can be seen from Table 6 where we show estimates without school density, the coefficients are virtually unchanged both in the magnitude and in the pattern over time.

Clearly, school density is to a large extent endogenous to enrollment rates. However, it is reasonable to argue that a 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 to 14 decreased substantially, at least until 1886.[34] Therefore, it is justifiable to use school density as a covariate in our model.[35]

In sum, OLS estimates show that landownership concentration has a strong negative association with enrollment rates in nineteenth century Prussia. Results suggest that the association of landownership concentration with education diminished throughout the century. This result is consistent with the hypothesis that landowners might have changed their attitude towards education due to increasing stakes in the industrial sector (Galor et al., 2009).

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

[34] Schools per child (6 to 14) were 1.025 in 1816; 0.817 in 1849; 0.716 in 1864; 0.646 in 1886; and 0.754 in 1896.

[35] We decided to use school density instead of school per child (6 to 14) because we believe that the former is better suited to capturing 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.

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-educated peasants might have been more able to redeem their land. Finally, higher-educated peasants might have had a stronger 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, the 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 and clay) tends 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 having a higher soil quality; areas with brighter colors are dominated by soil of lower quality.

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 poorer quality soil, therefore with a 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 that was followed by north-eastern European regions during the Middle Ages where the local lords, in order to attract more agricultural workers and encourage immigration, granted the peasants 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.

On the other hand, regions with relatively higher soil quality experienced a stronger demand for land which determined a more accentuated land fragmentation and more secure property rights. According to Boserup (1965), Binswanger

and McIntire (1987), and Binswanger and Rosenzweig (1986), increasing population pressure results in increasing intensification of land use and in growing pressure for security of land tenure (Eastwood et al., 2010). In fact, we find that in regions of the Rhineland and parts of Saxony and Silesia, where the soil is dominated by loam and clay (that is, higher quality), farm size is small on average, leading to a low concentration of large landholdings (compare Figures 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 remains a choice variable, although heavily dependant on the type of terrain.[36]

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 as in equation 1. The exclusion restriction demands that soil texture has no direct effect on enrollment rates. In order to ensure that condition, 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 the soil, namely the different shares of loam and 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 and 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)[37] in order to recover two components that are uncorrelated by construction.[38] The first component explains 66.4 percent of the total variation of the three soil variables, whereas the second component explains an

[36] 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.

[37] 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 contain most of the variance. See Appendix A.1 for more details.

[38] 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.

additional 32.9 percent of the variation. Thus, these two components capture more than 99 percent of the variation in soil texture.[39]

6.2 Instrumental variable results

First stage estimates of equation 2 are presented in the upper panel of Table 7. The two components are significantly correlated with landownership concentration. The first component is negative for loam and 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.

Second stage estimates are presented in the lower panel of Table 7. 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 us to be definitive on this matter.

Our IV estimates imply that OLS estimates are downward biased. This result is consistent with the findings of Easterly (2007). Instrumenting inequality using the ratio between the suitability of wheat to sugar crops, Easterly finds that IV estimates are about three times larger when 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 1 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.[40] It implies that if the share of large landholdings in 1849 had doubled (for instance from the average value of 2.4 to 4.8 percent), the enrollment rate in primary schools would have decreased by almost 9 percentage points. A similar increase at the end of the century (from the average of 0.8 to 1.6 percent) would have resulted in only 1 percentage point lower enrollment rates.

[39] For a similar approach see Galor et al. (2009). 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.

[40] These findings do not depend on the threshold of 300 PM which we chose 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.

6.3 Robustness checks

6.3.1 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 variations in the employment of child labor in agriculture and thus drove children out of school. In order to tackle this problem, we additionally include land productivity measures.

We have information on land productivity from the '*Grundsteuerreinertrag*' (henceforth GRE) determined in 1865. GRE is defined as the income from agrarian use of land less the costs of farming (see the Appendix A.2 for further details).[41] The GRE was stipulated by the tax administration as an assessment base for the land tax. By assessing the GRE, the tax administration explicitly aimed at determining the net earnings per Prussian acre of land for different kinds of farm land (arable, pasture, meadow, and horticulture), in different parts of the State (Kopsidis and Wolf, 2012, p. 6). In brief, the GRE approximates net farm income per acre before taxes, and thus corresponds to the concept of land rent. For our analysis, we use the deflated average land rent per unit of farmland.

Furthermore, we have detailed county-level information on yields for the most important crops (rye, wheat, oats, barley, and potatoes) for the years 1886 and 1896. The information on yields and the GRE are strongly positively correlated. For instance, the correlation between GRE and rye yields in 1886 is 0.84; the correlation with wheat yields is 0.72.

We include these controls for land productivity in our instrumental variable estimates presented in Table 8. In the upper panel we show the effect of landownership concentration, controlling for the GRE in 1865; in the lower panel, controlling for the yields in 1886 and 1896.[42] As one can see, using different controls for land productivity does not change the results. In fact, controlling for land productivity provides similar results to those in Table 7. The effect of landownership concentration on enrollment rates is highly significant until 1864, declining substantially towards the end of the century.

Compared to the IV estimates of Table 7, the point estimates for 1886 and 1896 are considerably smaller, suggesting a possible mediating role for land productivity. However, the point estimate for 1864, for which we have a precise measure of land

[41] We thank Niko Wolf and Michael Kopsidis for making these data available to us.

[42] In the lower panel, we assign the yields in 1886 to the years 1816, 1849, and 1864. 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 in 1886 and 1896.

productivity (upper panel), shows a larger coefficient (-3.048 versus -2.281). This result should mitigate any concern about the mediating role of land productivity.

6.3.2 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 in 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 9, column 6). In order to tackle this issue further, we can exclude those counties which consist only of a city.[43] The estimates, presented in Table 9, are consistent with the results discussed so far. As expected, with the exception of 1864, the first-stage F-statistics suggest that the instruments are stronger when omitting city counties.

6.3.3 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 Rhineland and Westphalia, the results (both for OLS and IV) are qualitatively similar, confirming the findings discussed throughout the paper.

The counties in the Rhineland that had been under the Napoleonic influence had more liberal institutions and serf labor was less prevalent (Acemoglu et al., 2011). We exploit this feature for the 1816 cross section and test whether the effect of landownership concentration is weaker in these counties. We construct a binary variable which takes on value one if the county lies (completely or in part) West of the river Rhine, that is if it belongs to one of the following districts: Aachen, Düsseldorf, Koblenz, Köln, and Trier. We interact this binary variable with large landownership. We can show that the effect of land concentration is indeed weaker in these counties: The main effect is -3.23 whereas the interaction term is 1.51.[44]

Analogously, one could argue about the existence of an east-west gradient regarding school financing. However, as discussed in more detail in Section 3.2, the differences in school funds in real terms were probably small. In fact, we observe

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

[44] Due to the low number of counties the interaction effect is imprecisely estimated (std. error 6.41) and is statistically insignificant.

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).

6.3.4 The Gini coefficient

We have argued that our measure of land concentration is better suited to capture the extent of labor exploitation than standard measures of inequality are. 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 holdings and assume that all the holdings belong to the largest category (>300 PM). In this case, the Gini coefficient would be equal to 0, since land is equally distributed among the 100 large landholders. Our concentration variable, instead, would take the value of 1, since large landholdings represent 100 percent 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 subject, under serfdom, to the will of the few large landowners.

In order to compare our measure of land concentration with standard inequality measures, we compute the Gini coefficient and estimate its effect on education.[45] Results are reported in the upper panel of Table 10. 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.[46]

Yet, it is important to note that the Gini coefficient computed across all landholdings does not consider landless agricultural workers. We correct for this by using a method proposed by Vollrath (2009, p. 8-9) which incorporates the total number of adult males.[47] 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 smaller 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 10 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

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

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

[47] One could argue that our concentration index should also be adjusted for landless agricultural workers. However, this does not apply in 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 appears to be a rather weak instrument for the standard and the adjusted Gini. Therefore, instrumental variable estimates generate coefficients with a very large bias.[48] We can, however, show that the pattern of the results is very similar to the results discussed above.

7 Panel models

In this section, we construct a panel dataset which allows us to address the issue of unobserved heterogeneity. Time-invariant 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 enrollment 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):

$$Edu_{it} = \beta_3 Land_{it} + \alpha_i + \tau_t + X_{it}\gamma_3 + \nu_{it} \tag{3}$$

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

Panel estimates are presented in Table 11. We show, alternatively, specifications using landownership concentration and the adjusted Gini. For completeness, in columns 1 and 4 we present pooled models with neither county nor time fixed effects. In columns 2 and 3, and 5 and 6 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 (columns 3 and 6) does not change the results qualitatively. In fact, the time fixed effects reinforce the coefficient of the adjusted Gini (column 6), whereas they reduce the effect of landownership concentration (column 3). The latter can be explained by the removal of the scale effect which is triggered by the change in the definition of landownership in the last two censuses.

[48] Results are not presented here, but are available upon request.

In columns 7 and 8 we estimate a fully flexible model—that is, we interact all variables with time dummies. In this way, we estimate the relationship between landownership concentration and enrollment rates for each time-period. The notion behind this approach is to test whether decreases in the negative effect of landownership concentration over the century can be observed also in a panel setting. As the reference year is 1816, the model shows the change in the relationship of interest as compared to the year 1816. This means that if our hypothesis is correct, the effect of landownership concentration should be less negative in later periods than at the time of the reforms in 1816. The results of the fully flexible model confirm the pattern estimated so far: The increasing coefficients of large landholdings interacted with time imply that the negative effect prevalent in 1816 decreases over the course of the nineteenth century (column 7). Because of the changes in landholding classifications and the contextual change of the unit of measurement discussed in Section 4, the adjusted Gini might be a more suitable measure for this approach. We find that the fully flexible model behaves similarly when estimated with the adjusted Gini (column 8).

8 The mechanism

We have shown that high levels of landownership concentration have a significant negative effect on education and that this effect declines towards the end of the nineteenth century. In this section we discuss possible mechanisms that can explain our reduced form results.

8.1 Supply factors

Earlier, we claimed that in the case of Prussia, restricting the supply of schools was not the key mechanism. To substantiate this claim, we present OLS estimates of the effect of landownership concentration on (*i*) school density, and (*ii*) the number of teachers per child (6-14). The results are presented in Table 12. We do not find any systematic relationship between landownership concentration and school density in any given year. The results show that, if anything, counties with a high concentration of large landholdings provided *more* teachers to each child (6-14).[49] This evidence does not lend support to the hypothesis that restrictions from the supply

[49] Regressions for the year 1886 suggest that teacher wages in counties with higher concentrations of large landholding were significantly lower, even after accounting for differences in prices and urbanization. This might indicate a lower quality of teachers, despite a higher teacher-child ratio.

side could explain the negative relationship between landownership concentration and enrollment rates.

8.2 Market integration

Increasing market integration with declining agricultural prices could cause a shift from an agricultural to an industrial economy and therefore explain the rise in the demand for education. To test this hypothesis, we estimate equation 1 including average agricultural prices for different crops for the period 1837-60 (wheat, rye, oats, and potatoes). In addition, we add a binary variable that is 1 if a county had access to a railroad line.[50] Since prices are averaged over the period 1837-60, we run estimates only for the 1849 and 1864 cross sections. The results are presented in Table 13. Including agricultural prices and railroad access in columns 1 and 2 reduces the effect of landownership concentration on enrollment rates by 28 percent in 1849 and by 35 percent in 1864. However, the effect remains highly significant both economically and statistically indicating that our relationship of interest cannot be entirely explained by increased market integration.[51]

In columns 3 and 4 we include an agricultural price index which is the average of the four agricultural prices weighted by the number of hectares devoted to each crop. The coefficients of the agricultural price index are not significant whereas the coefficients for large landownership become similar in size to the baseline OLS estimates presented in Table 5. Again the same pattern of the results holds.

8.3 Peasants' emancipation - A demand factor

Another plausible mechanism works through the emancipation of the peasantry which, we argue, increased the demand for education. We expect peasants to exhibit a higher demand for schooling for their children depending on the level of freedom achieved. We test this hypothesis by analyzing new cross-sectional data on emancipation achieved by 1848. The emancipation was a consequence of the implemented agricultural reforms discussed in subsection 3.1. These additional data allow not only to identify the number of emancipated peasants but also to distinguish between different levels of achieved freedom. The data depict the number of the peasants who redeemed their servile duties thus achieving the highest level of freedom. Furthermore, the data also depict the number of peasants who only redeemed their land but were yet obliged to provide servile duties thus achieving only

[50] See Hornung (2012) for more details.

[51] The coefficient on railroad access becomes insignificant in 1864, presumably because of the rapid railroad expansion in Prussia. This leads to a decrease in the county-level variation of that variable in subsequent years.

a limited level of freedom. Thus we construct two variables indicating (*i*) the share of emancipated landholders who redeemed their servile duties (variable *fully emancipated*, mean 0.039, s.d. 0.052) and (*ii*) the share of emancipated landholders who only redeemed their land (variable *partly emancipated*, mean 0.007 and s.d. 0.011).[52]

Indeed results presented in Table 14 confirm our hypothesis. Holding landownership concentration fixed, we find that counties with a larger share of emancipated peasants exhibit significantly higher enrollment rates. As a start, we regress enrollment rates on the share of fully emancipated peasants (those who redeemed servile duties) in a model similar to our baseline equation 1. As seen in the previous results of Table 13, agricultural prices represent relevant proxies for market integration. Furthermore, it might be the case that peasants in more prosperous regions were able to redeem their duties at an earlier stage. Therefore, we include controls for agricultural prices and a variable for per-capita tax revenues, which are expected to account for regional variations in income.[53] Results presented in column 1 show a significant positive relationship between fully emancipated peasants and enrollment rates.[54]

As discussed above, the burden of servile duties was much heavier in the regions east of the river Elbe. Consequently, the level of freedom achieved by redeeming servile duties in these regions should be comparatively higher. Excluding the provinces of Rhineland, Saxony and Westphalia from our sample, we find a significantly larger coefficient on emancipated peasants confirming our hypothesis (column 2).

Our second measure of emancipation — the partly emancipated — regards the freedom obtained by redeeming only land, effectively gaining legal ownership. In column 3 we find no significant relationship between the share of partly emancipated peasants and school enrollment. However, this variable is inflated by a high number of zeros since peasants from the regions west of the river Elbe had effectively achieved legal landownership without any compensation due to Napoleon's reforms. Thus, estimates restricted to the sample of eastern provinces presented in column 4 are

[52] We chose the denominator to be the sum of landowners and persons employed in agriculture to calculate the share of the agricultural population that became emancipated. However, using as denominator the total number of farms does not change the results qualitatively.

[53] Per-capita tax revenues, only available for 1867, are constructed from the total amount of tax revenues in Prussian Thaler from direct taxes—such as income tax, class tax, trade tax, building tax, and property tax—divided by the total population (Meitzen, 1868, vol. 4).

[54] We acknowledge that these result should not be interpreted as causal. The instrument previously used to identify exogenous variation in landownership concentration is a poor instrument in this setting. Soil composition is not able to predict the different timing and extent of serfs' emancipation, which depends, among other things, on the formal and informal institutional characteristics of each region. In fact, a possible source of county-level variation of peasants' emancipation could be the level of education in the county, raising concerns of reverse causality. To account for this problem, we include controls for lagged education, namely enrollment rates in 1816. Our results, however, remain virtually unchanged. Results including lagged education are available from the authors upon request.

more informative. Indeed we find a positive coefficient, though not significant, for partly emancipated peasants in the eastern provinces. However, this coefficient is smaller compared to the estimates for the fully emancipated (column 2). This result is thus consistent with our hypothesis that a higher level of gained freedom results in a higher demand for education.

It is important to note that landownership concentration consistently remains significant across all specifications, confirming that there is a negative relationship with enrollment rates, independently from the degree of emancipation granted to the peasantry. This suggests also that in these specifications we might be disentangling the supply from the demand effect.

The models in columns 5 and 6 include both types of emancipated peasants.[55] The results remain unchanged: the higher the share of full emancipation, the higher the enrollment rates; the share of partly emancipated peasants, thus with a lower degree of freedom, has a positive relationship with education, yet smaller and imprecisely estimated. The negative effect of landownership concentration remains highly significant and unaltered in magnitude.

In columns 7 and 8 we include an alternative measure for large estates, namely the county-level share of knight estates (*Rittergüter*) in 1866. We know from historical accounts that knight estates relied heavily on serf labor. Accordingly, this variable shows a negative relationship with enrollment rates consistent with the notion that large landholders opposed the spread of education. The coefficients for emancipations confirm our hypothesis on the positive demand effect, which is larger and significant for those who gained a higher level of freedom.

With the data at hand, any explanation of the exact reason why the demand for education increased with the emancipation of the peasantry would be highly speculative. However, one might argue that the freedom of occupational choice (*Gewerbefreiheit*) opened up the possibility to reap the benefits of an education and thus set the incentive to invest in children's human capital. In combination with fewer time constraints, due to the repeal of servile duties which allowed children to attend school, the main obstacles to investment in human capital were reduced.

8.4 Peasants' emancipation in a panel

Previous results from the panel estimates presented in Table 11 indicate that, also accounting for county fixed effects, a significant negative relationship between landown-

[55] It is also important to remember that the different degrees of emancipation were not county specific, i.e. within the same county there were both fully and partly emancipated peasants which allows us to estimate the models in column 5 and 6.

ership concentration and enrollment rates remains and decreases over time. If our hypothesis that the demand for education increased with the extent of freedom is correct, we would expect the eastern counties, "treated" by the agricultural reform, to drive the declining effect of landownership concentration on education. We have seen that the results for the 1848 cross section presented in Table 14 are consistent with this story. To add evidence in support of this hypothesis, we further exploit the panel dimension of our data and estimate equation 3 interacting landownership concentration with an indicator for the regions treated by the agricultural reform and time fixed effects. Consistently with the historical accounts (see section 3.1), we consider the counties treated by the Prussian agricultural reform to be those east of the river Elbe (variable *East Elbe*). In this way, we compare changes in enrollment rates *within* counties with different levels of landownership concentration in east Prussia for each year. In fact this estimation strategy resembles a differences-in-differences approach with a triple interaction term.

Peasants on very small parcels became able to redeem land and servile duties only after 1850, a development which we cannot observe in the previous cross-sectional estimate for 1848. Yet, these peasants were often forced to abandon their small holdings because of the impossibility to finance the land redemption, thus leading to increasing landownership concentration as observed in our data. However, as the burden of servile duties was relieved, they gained complete freedom and were able to pursue a new activity and their children had more time to attend school. Therefore, if our hypothesis is correct, counties for which landownership concentration increased over time due to the annexation of small parcels should exhibit a relative increase in enrollment rates over time. This would be consistent with the notion that, within the eastern regions, demand for education increased more where the concentration of large landownership increased, i.e. where the surplus of emancipated workers was larger.

Estimates of this model are presented in column 1 of Table 15. To facilitate interpretation, this model is again estimated relative to the baseline year 1816. Indeed we find that the coefficients of the interaction terms increase over time and are highly significant, demonstrating that in those counties affected by the reform (East Elbe), an increasing share of large landownership is associated with *increasing* enrollment rates. Part of the increment in the size of the coefficient between 1864 and 1886 could be due to the change in the unit of measurement of land. Thus, we present estimates including the adjusted Gini in column 2. Estimates of this specification show a more gradual change of the effect over time reflecting the first wave of

emancipation until the mid of the century and the second wave of emancipation for small landholders after 1850.

One could argue that the divergence between East and West in the effect of landownership concentration is due to increasing investments in primary education which affected particularly the laggard counties of the East.[56] Estimates in columns 3 and 4 of Table 15 use teachers and schools per 100 children (6-14) as dependent variable, respectively. The estimates do not show any significant pattern within counties east of the river Elbe. There is no evidence that in counties where landownership concentration increased because of the agricultural reform, more teachers had been employed or more schools had been built. These results are consistent with our hypothesis that the increase in enrollment rates in the eastern counties was driven by a higher demand for education from emancipated peasants.

In summary, we have shown that schooling infrastructure in the form of school density and the teacher-to-child ratio cannot explain the negative effect of landownership concentration on education in Prussia. Market integration, proxied by agricultural prices and railroad access, has a mediating effect but cannot fully explain the relationship of interest. We present some cross-sectional evidence that the emancipation of serfs possibly generated a stronger demand for primary schooling. We find that this effect is particularly strong and significant for those who gained complete emancipation compared to those who became legal owner of their land but were still obliged to provide servile duties. Panel estimates analyzing the effect of landownership concentration on education in the regions treated by the agricultural reform are consistent with the proposed mechanism.

9 Conclusion

Nineteenth century Prussia offers the possibility to study how the evolution of landownership concentration in relationship with agricultural reforms affects the spread of primary education. Using a unique county-level database that covers the entire 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 linguistic heterogeneity.

To overcome biases due to omitted variables and reverse causality, we adopt an instrumental variable approach. We identify the causal effect of landownership con-

[56] Cross-sectional results presented in Table 12 already reject this hypothesis.

centration on education using exogenous variation in farm size due to differences in the geological composition of the soil (soil texture). Regions which exhibit relatively poorer soil quality, therefore with a lower marginal value of land, historically experienced a lower population density and a weaker demand for land and are thus characterized by higher average farm sizes. On the contrary, regions with relatively higher soil quality experienced a denser population and 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.

We show that the negative effect of landownership concentration weakened over time, which is consistent with the existing notion of a diminishing opposition of the landed elite against mass education. We propose an alternative mechanism and present evidence consistent with the hypothesis that the declining effect of landownership concentration is the result of an increasing demand for education from the emancipated peasants.

County and time fixed effects estimates also confirm the existence of a significant negative effect of landownership concentration on enrollment rates within counties. Fully flexible specifications support the finding that the effect of landownership concentration on education weakened towards the end of the nineteenth century. Further panel specifications show that the counties east of the river Elbe, strongly affected by the Prussian agricultural reform, are responsible for the decreasing relationship between landownership concentration on education.

References

Acemoglu, D., D. Cantoni, S. Johnson, and J. A. Robinson (2011): "The Consequences of Radical Reform: The French Revolution," *American Economic Review*, 101, 3286–3307.

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.

Aghion, P., T. Persson, and D. Rouzet (2012): "Education and Military Rivalry," Working Paper 18049, National Bureau of Economic Research.

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, 29–44.

——— (2013): "Does Women's Education Affect Fertility? Evidence from Pre-Demographic Transition Prussia," *European Review of Economic History*, 17, 24–44.

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*, New York: Aldine Pub. Co.

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*, Frankfurt a. M.: Suhrkamp.

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: 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*, Weimar: Hermann Böhlaus Nachfolger.

Henning, F.-W. (1969): *Dienste und Abgaben der Bauern im 18. Jahrhundert*, Stuttgart: Gustav Fischer Verlag.

Hornung, E. (2012): "Railroads and Micro-regional Growth in Prussia," Ifo Working Paper 127, Ifo Institute.

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

Kopsidis, M. and N. Wolf (2012): "Agricultural Productivity Across Prussia During the Industrial Revolution: A Thuenen Perspective," *Journal of Economic History*, 72, 634–670.

Kourtellos, A., I. Stylianou, and C. Ming Tan (2013): "Failure to Launch? The Role of Land Inequality in Transition Delays," *European Economic Review*, 62.

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.

Lagerlöf, N.-P. (2009): "Slavery and Other Property Rights," *Review of Economic Studies*, 76, 319–342.

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: 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.

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

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

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*, New York and Oxford: 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*, Berlin: Akademie Verlag.

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

Sokoloff, K. L. and S. Engerman (2000): "Institutions, Factor 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.

——— (2012): "School Funding and Inequality in the United States, 1890," *Explorations in Economic History*, forthcoming.

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.

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 Variable definitions and data sources

We present variable definitions and data sources in Appendix Tables 1-5. This county-level database covers 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 sections using censuses from the years 1816, 1849, 1864, 1886 and 1896, if available, or 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). Some of the data used for the 1886 and 1896 cross section were taken from Galloway (2007).

Table 1: 1816 variables

Enrollment rate (6-14)	Number of attending students in public elementary schools (Öffentliche Elementarschulen), 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) in 1816 over number of children at recommended school age between 6 and 14 in 1816 (Public elementary schools were the only school type equally available in rural areas and towns at the time. Data on private and middle schools are available for the 172 medium and large towns.)
Large landholdings (share)	Share of estates larger than 300 PM. 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.
Protestant (share)	Share of total population that is Protestant in 1816.
Urban (share)	Share of total population living in cities which held city rights in 1816.
Industrial (share)	Number of looms in 1819 over total population in 1816 (since Prussia was not industrialized in 1816, we choose this as an indicator of proto-industrial occupations).
Agricultural (share)	Number of servants in agriculture in 1819 over total population in 1816 (no data is available for the total number of people in agriculture for this period).
Child dependency ratio	Ratio of the population younger than 15 years to the population between 15 and 60 years in 1816.
Population density	Number of people per square kilometer of land area in the county in 1816.
School density	Number of elementary and middle schools per square kilometer of land area in the county in 1816.
Inheritance (dummy)	Unity for counties that predominantly practiced partible inheritance (*Naturalteilung*) and zero for counties that predominantly practiced non partible inheritance (*Anerbenrecht*). (Coded using county-level maps of historical inheritance laws from ca. 1900 (Sering, 1897-1905).)
First language not German (1890)	Share of total population whose mother tongue is not German in 1890 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 121a).
GRE	The *Grundsteuerreinertrag* in real terms per area. The *Grundsteuerreinertrag* is defined as the income from agrarian use of land less all costs of farming and was originally judged by experts for tax purposes. The Prussian administration established the GRE to determine the net earnings per acre for different classes of farm land (Meitzen, 1868, vol. 4). The GRE is deflated with a county-level index of agricultural prices.
Teachers per children 6-14	Number of elementary- and middle-school teachers per 100 children at recommended school age (6 to 14 years) in 1816.

Note: Unless otherwise specified, the source of the 1816 and 1819 census data is Mützell (1823-1825, vol. 5-6).

Table 2: 1849 variables

Enrollment rate (6-14)	Number of attending students in public elementary schools (Öffentliche Elementarschulen) and public middle schools for boys or girls (Öffentliche Mittelschulen für Söhne oder Töchter) in 1849 over number of children at recommended school age between 6 and 14 in 1849.
Large landholdings (share)	Share of estates larger than 300 PM. (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.)
Protestant (share)	Share of total population that is Protestant in 1849.
Urban (share)	Share of total population living in cities which held city rights in 1849.
Industrial (share)	Number people working as craftsmen or in factories in 1849 over total population in 1849
Agricultural (share)	Number of people in agriculture in 1849 over total population in 1849 (including subsidiary occupations in agriculture, all family members, servants and farm-laborers).
Child dependency ratio	Ratio of the population younger than 15 years to the population between 15 and 60 years in 1849.
Population density	Number of people per square kilometer of land area in the county 1849.
School density	Number of elementary and middle schools per square kilometer of land area in the county 1849.
Inheritance (dummy)	Unity for counties that predominantly practiced partible inheritance (*Naturalteilung*) and zero for counties that predominantly practiced non partible inheritance (*Anerbenrecht*). (Coded using county-level maps of historical inheritance laws from ca. 1900 (Sering, 1897-1905).)
First language not German (1890)	Share of total population whose mother tongue is not German in 1890 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 121a).
GRE	The *Grundsteuerreinertrag* in real terms per area. The *Grundsteuerreinertrag* is defined as the income from agrarian use of land less all costs of farming and was originally judged by experts for tax purposes. The Prussian administration established the GRE to determine the net earnings per acre for different classes of farm land (Meitzen, 1868, vol. 4). The GRE is deflated with a county-level index of agricultural prices.
Teachers per children 6-14	Number of elementary- and middle-school teachers per 100 children at recommended school age (6 to 14 years) in 1849.
Agricultural prices	Average market prices in Prussian *Silbergroschen* for bushels of wheat, rye, oat, and potatoes, over the period 1837-1860. (Prices were collected annually during the 15-day period of *Martinimarkt* (Meitzen, 1868, vol. 4).)
Railroad access (dummy)	Unity if a county was crossed by at least one railroad line in 1849 (Hornung, 2012).

Note: Unless otherwise specified, the source of the 1849 data is (Statistisches Bureau zu Berlin, 1851-1855, vol. 1-6b).

Table 3: 1864 variables

Enrollment rate (6-14)	Number of attending students in 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) in 1864 over number of children at recommended school age between 6 and 14 in 1864.
Large landholdings (share)	Share of estates larger than 300 PM. (Meitzen (1868) 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.)
Knight estates (share)	Share of knight estates (*Rittergüter*) over the total number of holdings (Meitzen, 1868).
Protestant (share)	Share of total population that is Protestant in 1864.
Urban (share)	Share of total population living in cities which held city rights in 1864.
Industrial (share)	Number of people employed in mining and industry (including construction) in 1867 over total population in 1867 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 16b).
Agricultural (share)	Number number of people employed in agriculture, forestry and hunting, and fishing in 1867 over total population in 1867 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 16b).
Child dependency ratio	Ratio of the population younger than 15 years to the population between 15 and 65 years in 1864.
Population density	Number of people per square kilometer of land area in the county in 1864.
School density	Number of elementary and middle schools per square kilometer of land area in the county 1864.
Inheritance (dummy)	Unity for counties that predominantly practiced partible inheritance (*Naturalteilung*) and zero for counties that predominantly practiced non partible inheritance (*Anerbenrecht*). (Coded using county-level maps of historical inheritance laws from ca. 1900 (Sering, 1897-1905).)
First language not German (1890)	Share of total population whose mother tongue is not German 1890 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 121a).
GRE	The *Grundsteuerreinertrag* in real terms per area. The *Grundsteuerreinertrag* is defined as the income from agrarian use of land less all costs of farming and was originally judged by experts for tax purposes. The Prussian administration established the GRE to determine the net earnings per acre for different classes of farm land (Meitzen, 1868, vol. 4). The GRE is deflated with a county-level index of agricultural prices.
Teachers per children 6-14	Number of elementary- and middle-school teachers per 100 children at recommended school age (6 to 14 years) in 1864.
Agricultural prices	Average market prices in Prussian *Silbergroschen* for a bushel of wheat, rye, oat, and potatoes, over the period 1837-1860. (Prices were collected annually during the 15-day period of *Martinimarkt* (Meitzen, 1868, vol. 4).)
Railroad access (dummy)	Unity if a county was crossed by at least one railroad line in 1864 (Hornung, 2012).
Tax revenues per capita (1867)	Total amount of revenues from direct taxes in Prussian Thaler over the total population (Meitzen, 1868, vol. 4).

Note: Unless otherwise specified, the source of the 1864 data is (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 10). 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.

Table 4: 1886 variables

Enrollment rate (6-14)	Number of attending students in public primary schools (Öffentliche Volksschulen) in 1886 over number of children at recommended school age in 1886 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 101).
Large landholdings (share)	Share of farms larger than 100 ha arable land 1882. (Königliches Statistisches Bureau in Berlin (1861-1934, vol. 76c) 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.)
Protestant (share)	Share of total population that is Protestant in 1880 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 66).
Urban (share)	Share of total population living in cities which held city rights in 1880 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 61).
Industrial (share)	Number of people employed in mining and industry 1882 over total population in 1882 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 76b).
Agricultural (share)	Number number of people employed in agriculture and animal husbandry in 1882 over total population in 1882 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 76b).
Child dependency ratio	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).
Population density	Number of people per square kilometer of land area in the county in 1882 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 76b).
School density	Number of schools per square kilometer of land area in the county in 1886 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 101).
Inheritance (dummy)	Unity for counties that predominantly practiced partible inheritance (*Naturalteilung*) and zero for counties that predominantly practiced non partible inheritance (*Anerbenrecht*). (Coded using county-level maps of historical inheritance laws from ca. 1900 (Sering, 1897-1905).)
First language not German (1890)	Share of total population whose mother tongue is not German in 1890 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 121a).
GRE	The *Grundsteuerreinertrag* in real terms per area. The *Grundsteuerreinertrag* is defined as the income from agrarian use of land less all costs of farming and was originally judged by experts for tax purposes. The Prussian administration established the GRE to determine the net earnings per acre for different classes of farm land (Meitzen, 1868, vol. 4). The GRE is deflated with a county-level index of agricultural prices.
Agricultural yields	Total yields of winter wheat, winter rye, summer barley, oats and potatoes per hectare in 1886 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 92).
Teachers per children 6-14	Number of school teachers per 100 children at recommended school age (6 to 14 years) in 1886 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 101).

Table 5: 1896 variables

Enrollment rate (6-14)	Number of attending students in public primary schools (Öffentliche Volksschulen) in 1896 over number of children at recommended school age in 1896 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 151b).
Large landholdings (share)	Share of farms larger than 100 ha arable land 1895. (Königliches Statistisches Bureau in Berlin (1861-1934, vol. 142b) 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.)
Protestant (share)	Share of total population that is Protestant in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 148a).
Urban (share)	Share of total population living in cities which held city rights in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 143).
Industrial (share)	Number of people employed in mining and industry 1895 over total population in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 142a).
Agricultural (share)	Number number of people employed in agriculture and animal husbandry 1895 over total population in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 142a).
Child dependency ratio	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).
Population density	Number of people per square kilometer of land area in the county in 1895 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 148a).
School density	Number of schools per square kilometer of land area in the county in 1896 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 151b).
Inheritance (dummy)	Unity for counties that predominantly practiced partible inheritance (*Naturalteilung*) and zero for counties that predominantly practiced non partible inheritance (*Anerbenrecht*). (Coded using county-level maps of historical inheritance laws from ca. 1900 (Sering, 1897-1905).)
First language not German (1890)	Share of total population whose mother tongue is not German in 1890 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 121a).
GRE	The *Grundsteuerreinertrag* in real terms per area. The *Grundsteuerreinertrag* is defined as the income from agrarian use of land less all costs of farming and was originally judged by experts for tax purposes. The Prussian administration established the GRE to determine the net earnings per acre for different classes of farm land (Meitzen, 1868, vol. 4). The GRE is deflated with a county-level index of agricultural prices.
Agricultural yields	Total yields of winter wheat, winter rye, summer barley, oats and potatoes per hectare in 1896 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 147).
Teachers per children 6-14	Number of school teachers per 100 children at recommended school age (6 to 14 years) in 1896 (Königliches Statistisches Bureau in Berlin, 1861-1934, vol. 151b).

Data on peasants' emancipation

In order to analyze the mechanism through which landownership concentration affected education in Section 8, we use additional data that assess the achievements of the agrarian reform taken from Meitzen (1868, vol. 4). After 1848, a full scale county-level assessment was undertaken to measure the outcome of the agrarian reform.

The data report the number of peasants that had become emancipated from serfdom separately by their extent of emancipation at the end of 1848. We denote peasants with weak rights of ownership that redeemed their land and gained legal landownership due to the Edict of 1816 as 'partly emancipated'. We denote peasants that held legal rights of ownership and already redeemed all servile duties due to the Edict of 1821 as 'fully emancipated'. Peasants on small parcels were exempted from the reform due to the Edict of 1816 until the Commutation Law in 1850 and are thus not covered by these data.

Table 6 below reports the increasing number of emancipated peasants over the second part of the nineteenth century. The fact that the absolute number of partly emancipated increases only marginally over time demonstrates that the large majority of the peasantry was already partly emancipated by 1848 while the process of full emancipation continued to gain momentum towards the end of the century.

Table 6: State-level emancipated peasants

Year	Partly emancipated	Fully emancipated
1848	70582	289651
1855	79951	936333
1865	83300	1348178
1885	87110	2527685

Source: Meitzen and Grossmann (1901).

The number of partly emancipated peasants is effectively zero in all areas west of the river Elbe and in the district of Stralsund (Swedish Pommerania). Here the Napoleonic reforms had already established legal landownership without compensation.

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.

Hornung, E. (2012): "Railroads and Micro-regional Growth in Prussia," Ifo Working Paper 127, Ifo Institute.

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, Berlin: Verlag von Paul Parey.

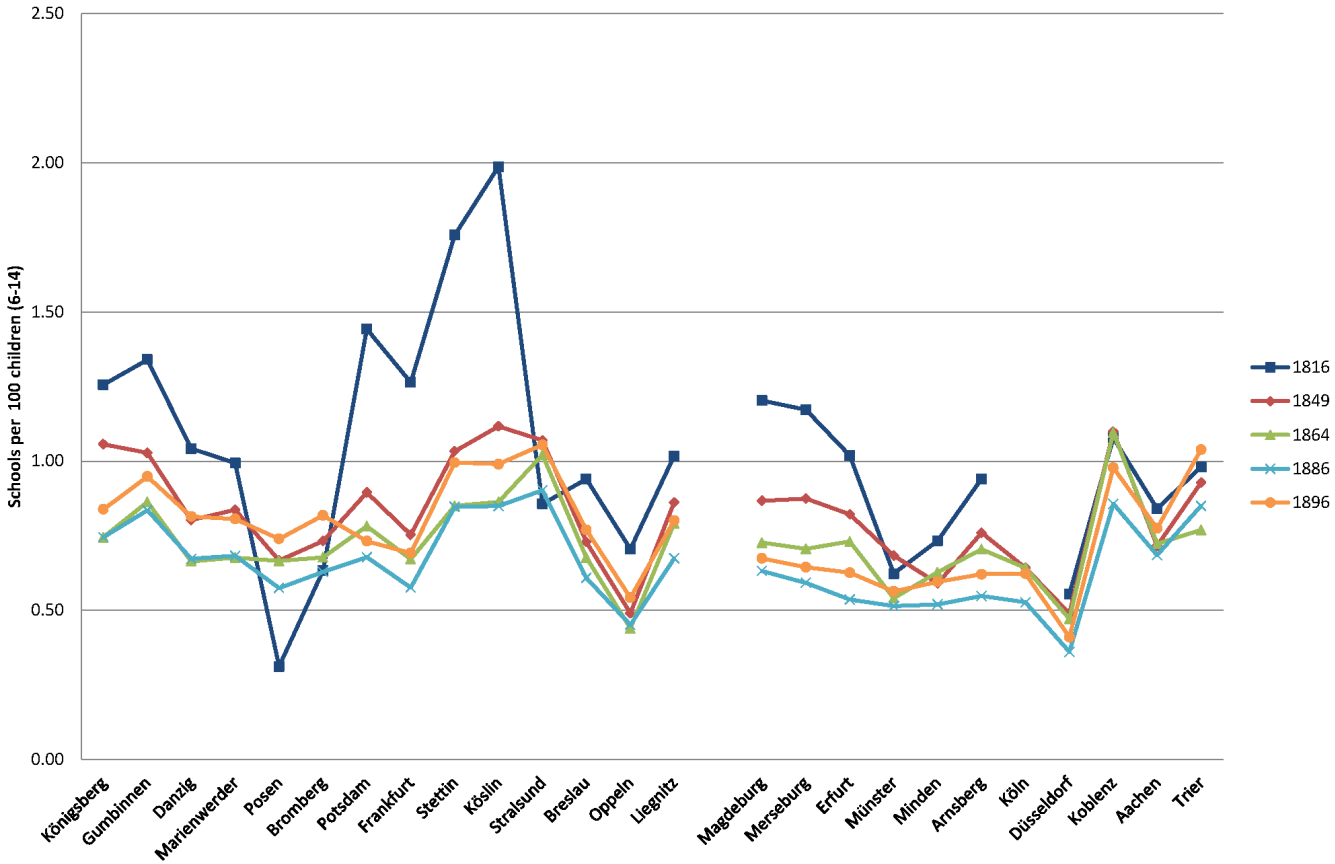
Meitzen, A. and F. Grossmann (1901): *Der Boden und die landwirthschaftlichen Verhältnisse des Preussischen Staates*, vol. 6, Berlin: Verlag von Paul Parey.

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

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

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

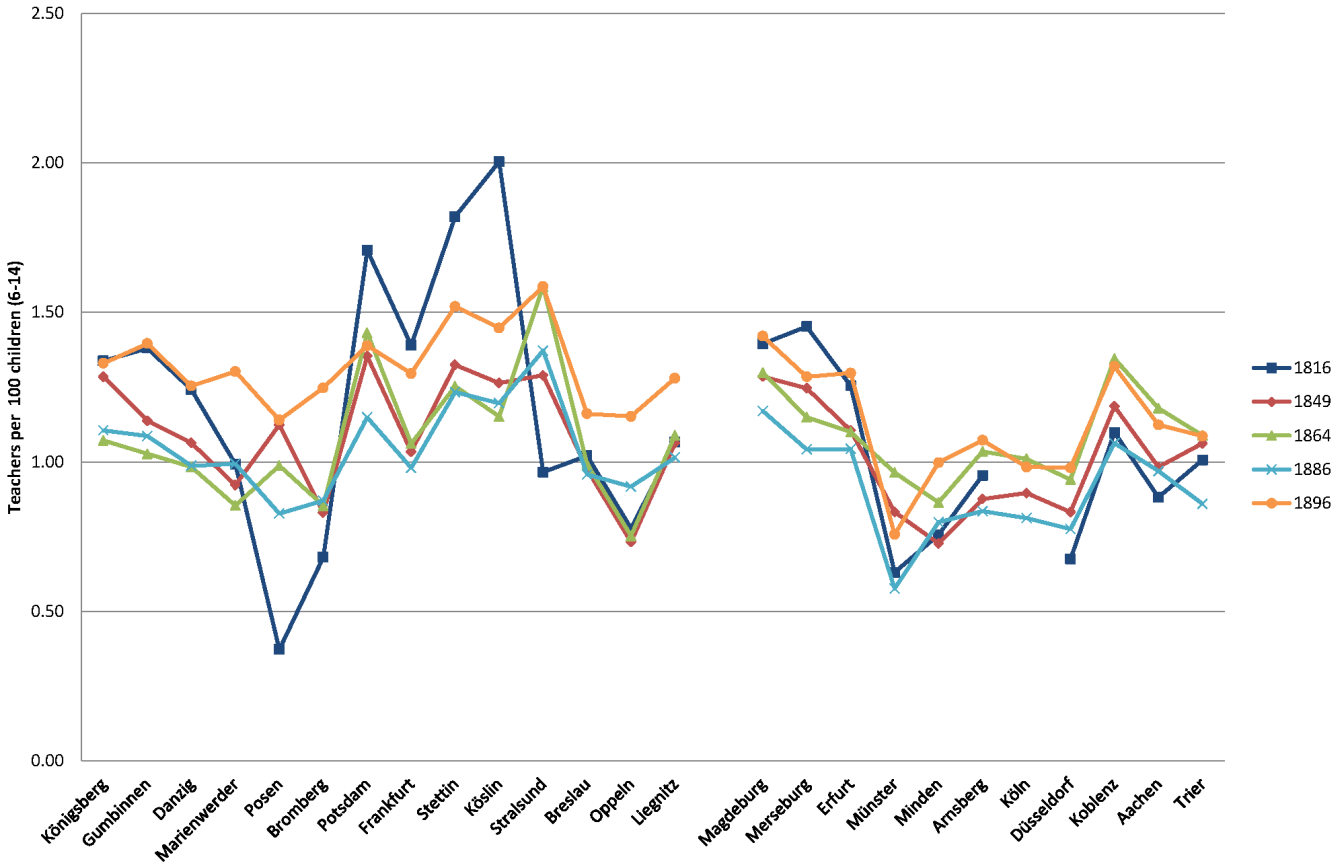


Note: Number of schools per 100 children at school-age (6-14) at the district level.

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

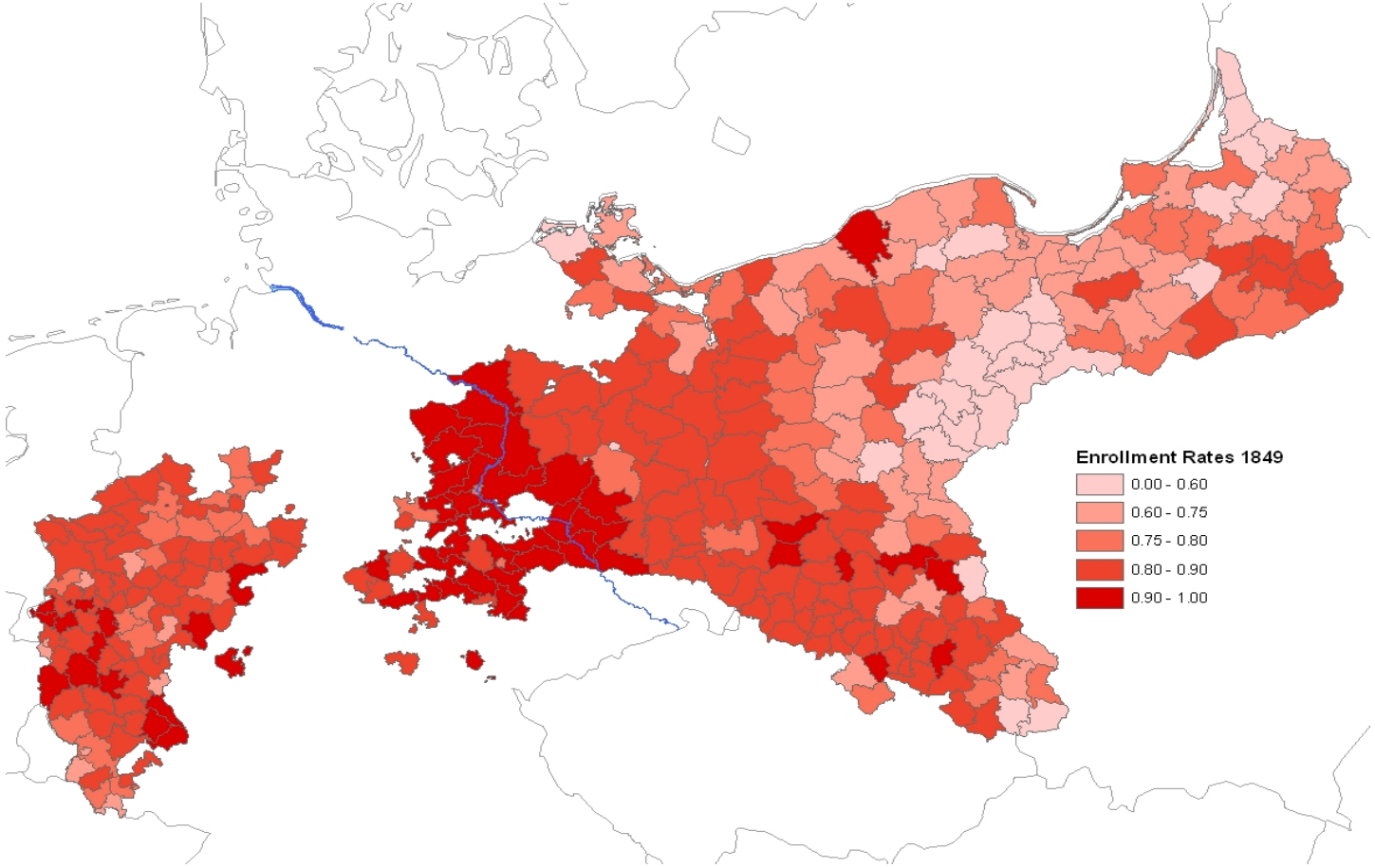


Note: Number of teachers per 100 children at school-age (6-14) at the district level.

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.

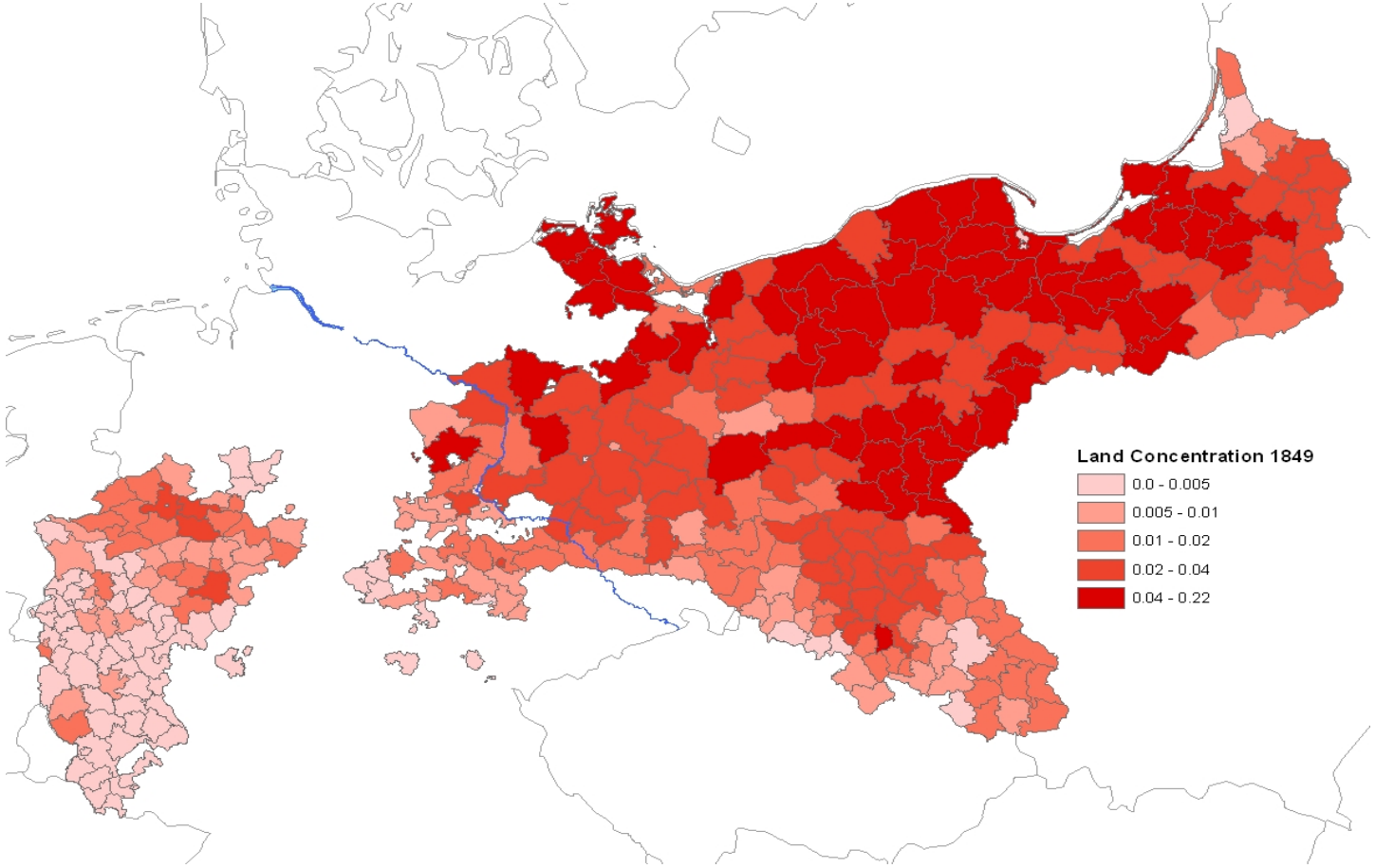
Figure 3: Enrollment rates in 1849



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

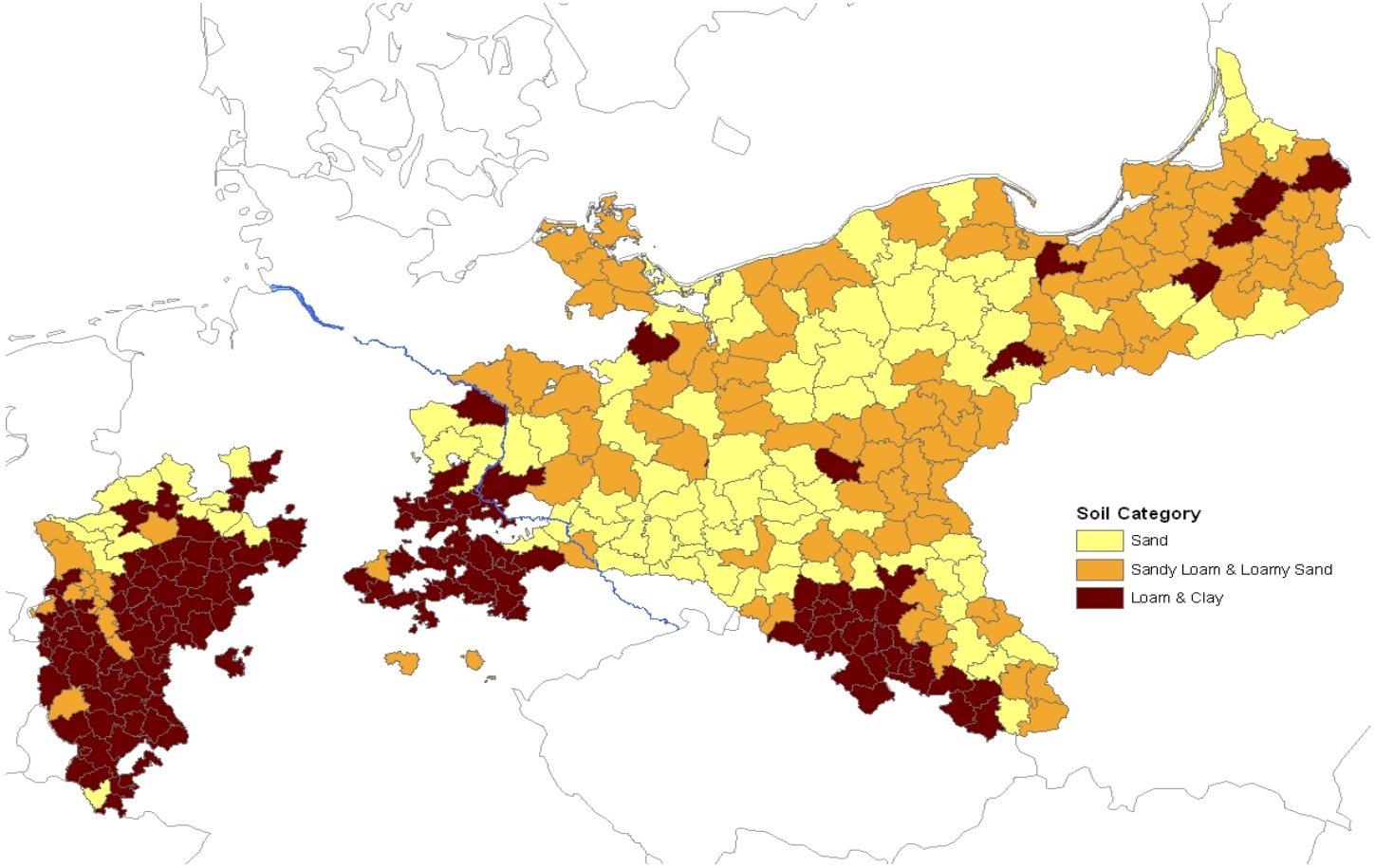


Note: The map shows the county structure of Prussia 1849.

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	0.626	0.186	0.203	0.195
	(0.038)	(0.194)	(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)
Teachers per 100 children (6-14)	1.124	1.062	1.071	0.981	1.230
	(0.528)	(0.269)	(0.264)	(0.211)	(0.243)
Observations	272	335	335	335	335

Note: Standard deviations in parenthesis.

Source: See Appendix A.2 for data sources and details.

Table 4: Bivariate correlations

	(1)	(2)	(3)	(4)	(5)
Dep. var: Enrollment rate	1816	1849	1864	1886	1896
Large landholdings (share)	-2.002***	-1.613***	-1.520***	-0.940**	-0.387
	(0.552)	(0.219)	(0.207)	(0.364)	(0.510)
Observations	291	335	335	335	335
R-squared	0.04	0.14	0.14	0.02	0.00

Note: OLS estimates at the county level. Constant omitted. 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: Landownership concentration and enrollment rates

	(1)	(2)	(3)	(4)	(5)
Dep. var: Enrollment rate	1816	1849	1864	1886	1896
Large landholdings (share)	-3.195***	-1.348***	-1.181***	-0.616*	-0.708*
	(0.507)	(0.250)	(0.190)	(0.328)	(0.409)
Protestant (share)	0.172***	0.048***	0.040***	-0.004	0.007
	(0.031)	(0.016)	(0.013)	(0.008)	(0.007)
Urban (share)	-0.068	-0.080**	0.047	-0.116***	-0.099***
	(0.063)	(0.041)	(0.056)	(0.022)	(0.020)
Industrial (share)	-0.478**	0.136	0.322	0.619***	0.609***
	(0.233)	(0.146)	(0.203)	(0.111)	(0.081)
Agricultural (share)	0.207	-0.029	0.312*	0.589***	0.550***
	(0.265)	(0.035)	(0.167)	(0.103)	(0.066)
Child dependency ratio	-0.300**	-0.035	-0.180**	0.188***	0.230***
	(0.132)	(0.084)	(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***	-0.167***	-0.140***	-0.057***	-0.010
	(0.059)	(0.030)	(0.025)	(0.012)	(0.009)
Constant	0.769***	0.873***	0.782***	0.611***	0.597***
	(0.099)	(0.070)	(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 6: Landownership concentration and enrollment rates, without the school density control

	(1)	(2)	(3)	(4)	(5)
Dep. var: Enrollment rate	1816	1849	1864	1886	1896
Large landholdings (share)	-2.952***	-1.316***	-1.220***	-0.648*	-0.581
	(0.556)	(0.250)	(0.197)	(0.330)	(0.451)
Observations	272	335	335	335	335
R-squared	0.44	0.35	0.39	0.57	0.59

Note: OLS estimates. Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1. Constant omitted. Controls: % protestant, % urban, % industrial, % agricultural, child dependency ratio, population density, inheritance, first language not German (1890). Source: See Appendix A.2 for data sources and details.

Table 7: 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 rate				
Large landholdings (share)	-8.858***	-3.774***	-2.281**	-1.943*	-1.478
	(2.183)	(1.125)	(1.037)	(1.054)	(1.165)
Protestant (share)	0.251***	0.082***	0.050***	0.003	0.012
	(0.043)	(0.021)	(0.014)	(0.009)	(0.009)
Urban (share)	0.045	-0.030	0.092	-0.111***	-0.094***
	(0.119)	(0.052)	(0.076)	(0.023)	(0.022)
Industrial (share)	-1.222**	-0.347	0.257	0.521***	0.572***
	(0.475)	(0.262)	(0.213)	(0.122)	(0.096)
Agricultural (share)	0.584*	0.024	0.505**	0.582***	0.559***
	(0.299)	(0.044)	(0.237)	(0.103)	(0.068)
Child dependency ratio	-0.091	0.110	-0.034	0.201***	0.235***
	(0.167)	(0.131)	(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***	0.095***	0.040***	0.112***	0.187***
	(0.123)	(0.017)	(0.014)	(0.035)	(0.054)
Inheritance (dummy)	-0.050	-0.014	0.022*	-0.009	-0.012
	(0.031)	(0.019)	(0.013)	(0.008)	(0.008)
First language not German (1890)	-0.184***	-0.152***	-0.144***	-0.054***	-0.008
	(0.065)	(0.032)	(0.026)	(0.012)	(0.009)
Constant	0.629***	0.813***	0.675***	0.619***	0.598***
	(0.122)	(0.090)	(0.129)	(0.060)	(0.042)
Observations	272	335	335	335	335
Kleibergen-Paap F statistic	8.236	11.10	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 8: Robustness check – Land productivity

	(1)	(2)	(3)	(4)	(5)
Dep. var: Enrollment rate	1816	1849	1864	1886	1896
Controlling for GRE	-8.944***	-3.700***	-3.048***	-1.828*	-0.352
	(2.354)	(1.168)	(1.149)	(1.108)	(1.312)
Observations	272	335	335	335	335
Kleibergen-Paap F statistic	7.352	10.43	7.514	13.54	11.42
Controlling for yields	-9.016**	-3.308***	-2.357**	-1.166	-0.442
	(2.515)	(1.118)	(0.976)	(0.948)	(1.068)
Observations	269	329	329	329	327
Kleibergen-Paap F statistic	6.357	9.611	9.321	15.33	13.57

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 the upper panel the GRE refers to 1865. In the lower panel, yields in columns 1-4 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 9: 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 rate					
Large landholdings (share)	-7.652***	-3.778***	-2.422**	-1.441	-1.005	
	(1.841)	(1.164)	(0.984)	(0.886)	(0.921)	
Large landholdings (share in rural)						-2.767***
						(0.855)
Protestant (share)	0.195***	0.056**	0.029**	0.002	0.009	0.049**
	(0.043)	(0.022)	(0.014)	(0.008)	(0.008)	(0.021)
Urban (share)	0.105	0.041	0.173**	-0.113***	-0.090***	0.092
	(0.101)	(0.066)	(0.074)	(0.025)	(0.020)	(0.072)
Industrial (share)	-1.036**	-0.246	0.708***	0.464***	0.499***	-0.143
	(0.512)	(0.231)	(0.133)	(0.120)	(0.092)	(0.205)
Agricultural (share)	0.684**	-0.032	0.434**	0.540***	0.459***	-0.045
	(0.275)	(0.041)	(0.219)	(0.107)	(0.060)	(0.042)
Child dependency ratio	-0.073	0.112	0.014	0.171***	0.206***	0.112
	(0.153)	(0.126)	(0.161)	(0.032)	(0.027)	(0.120)
Population density	-0.184***	-0.106***	-0.082***	0.001	-0.003	-0.114***
	(0.050)	(0.027)	(0.011)	(0.007)	(0.005)	(0.027)
School density	1.453***	0.724***	0.181	0.144	0.161*	0.747***
	(0.374)	(0.252)	(0.135)	(0.122)	(0.098)	(0.234)
Inheritance (dummy)	-0.049*	-0.018	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**	-0.143***	-0.136***	-0.052***	-0.011	-0.157***
	(0.067)	(0.032)	(0.025)	(0.013)	(0.009)	(0.030)
Constant	0.568***	0.844***	0.668***	0.652***	0.651***	0.828***
	(0.119)	(0.081)	(0.120)	(0.058)	(0.039)	(0.079)
Observations	267	325	325	324	324	325
Kleibergen-Paap F statistic	11.51	11.04	9.658	15.39	17.56	12.26

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 10: Gini coefficient and enrollment rates (OLS)

	(1)	(2)	(3)	(4)	(5)
Dep. var.: Enrollment rate	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. 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 11: Panel analysis

Dep. var.: Enrollment rate	(1) Pooled	(2) FE	(3) TFE	(4) Pooled	(5) FE	(6) TFE	(7) Fully flex	(8) Fully flex
Large landholdings (share)	-2.424***	-2.099***	-1.053**				-2.090***	
	(0.263)	(0.488)	(0.452)				(0.536)	
Gini adjusted				-0.245***	-0.147***	-0.164***		-0.343*
				(0.034)	(0.056)	(0.057)		(0.179)
Large landholdings (share)*1849							1.586***	
							(0.546)	
Large landholdings (share)*1864							2.020***	
							(0.543)	
Large landholdings (share)*1886							6.879***	
							(0.954)	
Large landholdings (share)*1896							7.324***	
							(1.504)	
Gini adjusted*1849								0.273
								(0.167)
Gini adjusted*1864								0.313*
								(0.163)
Gini adjusted*1886								0.558***
								(0.167)
Gini adjusted*1896								0.696***
								(0.174)
Observations	1,387	1,387	1,387	1,387	1,387	1,387	1,387	1,387
R-squared	0.50	0.59	0.70	0.47	0.57	0.70	0.81	0.80
Number of kreiskey1800		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. 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. Controls: % protestant, % urban, % industrial, % agricultural, child dependency ratio, population density, school density.

Source: See Appendix A.2 for data sources and details.

Table 12: The supply of education – Schools and teachers

	(1)	(2)	(3)	(4)	(5)
Dep. var.: School density	1816	1849	1864	1886	1896
Large landholdings (share)	0.597	0.364	-0.830	-0.281	0.690
	(0.751)	(0.270)	(0.621)	(0.666)	(0.647)
R-squared	0.86	0.90	0.90	0.94	0.92
Dep. var.: Teachers per child (6-14)					
Large landholdings (share)	5.476**	2.171***	1.464***	2.454*	2.036
	(2.662)	(0.547)	(0.501)	(1.456)	(1.927)
R-squared	0.46	0.42	0.51	0.48	0.55
Observations	272	335	335	335	335

Note: OLS estimates. Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1. Constant omitted. Controls: % protestant, % urban, % industrial, % agricultural, child dependency ratio, population density, inheritance, first language not German (1890).

Source: See Appendix A.2 for data sources and details.

Table 13: The demand for education – Market integration

	(1)	(2)	(3)	(4)
Dep. var.: Enrollment rate	1849	1864	1849	1864
Large landholdings (share)	-0.954***	-0.771***	-1.198***	-1.160***
	(0.311)	(0.275)	(0.303)	(0.273)
Avg. wheat price 1837-60	-0.005***	-0.004***		
	(0.002)	(0.001)		
Avg. rye price 1837-60	0.004**	0.004**		
	(0.002)	(0.002)		
Avg. oat price 1837-60	0.010***	0.013***		
	(0.003)	(0.003)		
Avg. potatoe price 1837-60	-0.002**	-0.003***		
	(0.001)	(0.001)		
Railroad access (dummy)	0.019*	-0.005	0.038***	0.009
	(0.010)	(0.009)	(0.011)	(0.009)
Agicultural price index			0.001	0.002
			(0.001)	(0.001)
Observations	295	295	295	295
R-squared	0.430	0.488	0.381	0.392

Note: OLS estimates. Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1. Constant omitted. Controls: % protestant, % urban, % industrial, % agricultural, child dependency ratio, population density, school density, inheritance, first language not German (1890).

Source: See Appendix A.2 for data sources and details.

Table 14: The demand for education – Peasants' emancipation in a cross section

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var: Enrollment rate	Full sample	East sample	Full sample	East sample	Full sample	East sample	Full sample	East sample
Fully emancipated	0.313**	0.850***			0.313**	0.856***	0.319**	0.919***
	(0.142)	(0.300)			(0.143)	(0.302)	(0.150)	(0.316)
Partly emancipated			-0.012	0.448	0.025	0.516	0.130	0.466
			(0.774)	(0.908)	(0.756)	(0.873)	(0.813)	(0.915)
Large landholdings (share)	-0.762***	-0.765**	-0.813***	-0.849**	-0.764**	-0.811**		
	(0.291)	(0.305)	(0.305)	(0.333)	(0.300)	(0.327)		
Knight estates (share)							-1.585*	-1.530
							(0.908)	(0.969)
Observations	243	162	243	162	243	162	243	162
R-squared	0.47	0.43	0.46	0.40	0.47	0.43	0.47	0.42

Note: OLS estimates. Robust standard errors in parentheses. Significance: *** p<0.01, ** p<0.05, * p<0.1. Constant omitted. Controls: % protestant, % urban, % industrial, % agricultural, child dependency ratio, population density, school density, inheritance, first language not German (1890), average agricultural prices (1837-60) for wheat, rye, oats, and potatoes, and tax revenues per capita in 1867. Models in columns 2, 4, 6, and 8 exclude the provinces of Rhineland, Saxony, and Westphalia.

Source: See Appendix A.2 for data sources and details.

Table 15: The demand for education – A diff-in-diff approach

	(1)	(2)	(3)	(4)
Dep. var.:	Enrollment rate	Enrollment rate	Teachers p.c.	Schools p.c.
Large landholdings*East Elbe	-2.360**		-1.244	-9.914
	(0.953)		(1.654)	(6.744)
Large landholdings*East Elbe*1849	3.456***		1.770	6.671
	(1.049)		(1.259)	(5.735)
Large landholdings*East Elbe*1864	3.419***		2.500	12.639*
	(1.100)		(1.943)	(6.812)
Large landholdings*East Elbe*1886	10.616***		6.376	13.832
	(2.618)		(4.343)	(9.869)
Large landholdings*East Elbe*1896	13.095***		0.160	12.799
	(2.645)		(3.787)	(11.341)
Gini adjusted*East Elbe		-0.874***		
		(0.300)		
Gini adjusted*East Elbe*1849		0.601**		
		(0.299)		
Gini adjusted*East Elbe*1864		0.900**		
		(0.430)		
Gini adjusted*East Elbe*1886		1.264***		
		(0.367)		
Gini adjusted*East Elbe*1896		1.229***		
		(0.370)		
Control variables (* Year FE)	Yes	Yes	Yes	Yes
Observations	1387	1387	1387	1387
R-squared	0.83	0.82	0.92	0.52

Note: Panel estimates. Robust standard errors in parentheses. 13 observations drop out from the analysis because of missing information in the 1816 data. The adjusted Gini accounts for propertyless farm workers. Significance: *** p<0.01, ** p<0.05, * p<0.1. The models include the main effects of large landholdings and the Gini adjusted, East Elbe, year fixed effects, and pairwise interactions. Control variables interacted with year fixed effects: % protestant, % urban, % industrial, % agricultural, child dependency ratio, population density, school density.

Source: See Appendix A.2 for data sources and details.