Rage Against the Machine: How Inequality Affected Technological Progress in Nineteenth-Century Prussia

Jonathan Adams

University of Cambridge

July 2018

Abstract

This dissertation examines the relationship between land inequality and the Industrial Revolution in nineteenth-century Prussia. An instrumental variable approach, using the heterogeneity in soil quality, suggests that there is a causal negative impact of land inequality on the spread of the Industrial Revolution as measured by the occupational structure. This impact is not detectable when examining investments in new technologies or methods of textile production.

Contents

1 Introduction

1.1 Motivation

In recent years, inequality has become a key part of the debate around economic policy. Works such as Piketty (2014) and Stiglitz (2012) have emphasised the potential for inequality to cause harm to social cohesion and economic growth. This paper uses detailed quantitative data to investigate the effect of inequality on technological development at the point when modern economic growth began: the Industrial Revolution. New technologies were developed and diffused, and some countries raced ahead of others in terms of development. There was both the 'Great Divergence', where Europe developed significantly faster than other regions of the world, and the 'Little Divergence', where even within Europe there were significant disparities in growth rates between nations. Whilst many factors are likely to have played a role in the advantages earned by some countries over others, including effective economic institutions and resource endowments, assessing the effect of inequality is important both in attempting to understand the Industrial Revolution, and in its relevance to modern economic policy.

The setting for this analysis is nineteenth-century Prussia. This is due to the empirical richness and heterogeneity of the Prussian Empire, allowing for a rigorous quantitative approach. This paper utilises detailed censuses collected by the Prussian Bureau of Statistics, and applies modern econometric techniques designed to assess causality. To address possible issues of endogeneity we adopt an instrumental variables (IV) approach, using the varying soil quality in Prussia as an exogenous influence on the inequality in landholdings.

The results of the analysis are stark, finding that the effect of inequality differs substantially depending on the measure of technological development examined. In particular, the analysis finds strong relationships between inequality and occupational structure, but no similarly strong relationships between inequality and investments in new technology, or inequality and methods of textile production. There are several potential explanations for this result, it could be caused by the limited diffusion of certain technologies, or the effect of inequality is greater on slowing down changes in occupational structure than in changing other measures. Importantly, it suggests that there is a high sensitivity to the variable we use as a proxy of technological progress - there is no necessity that inequality (or other explanatory factors) should affect all measures equally.

As a likely causal mechanism for these effects, this paper looks in particular at institutional economics. There is historical evidence that large landholdings in Prussia afforded the owners significant political power, and this political power could be used by landowners to prevent the spread of the Industrial Revolution when they would financially benefit from doing so. Their ability to prevent this spread may have been more effective at slowing occupational change than in preventing industrialists investing in new machinery. This is in accordance with Acemoglu et al. (2005)'s theory of political institutions, and is the most likely causal mechanism by which inequality reduced the spread of technological development.

1.2 Structure

The paper proceeds as follows: Section 2 reviews the relevant literature pertaining to both the empirical approach and the historical background. Section 3 provides an outline of the data and summary statistics, and Section 4 presents the quantitative analysis. Finally, Section 5 concludes.

2 Related Literature

2.1 Historical Background

Prussia is used for this analysis due to its rich empirical data. As detailed in Section 3, the records maintained by the Prussian Statistical Office are highly detailed, and permit a rigorous quantitative approach. Additionally, Prussia also provides a distinctive commentary on the Industrial Revolution. As Voth (2001, p.109) notes:

> At the end of the eighteenth century, the area that was to be unified in 1871 was remarkably backwards in economic terms. At a time when less than half the English population was still working the land, the figure in Germany was closer to 80 per cent. Incomes and life expectancy were low even by the standards of the time; infant mortality was high. (...) By the end of the nineteenth century, however, it had undergone a classic 'rags-to-riches' transformation. National income and population size had rapidly increased, and the prowess of its export industries earned it admiration and caused consternation abroad.

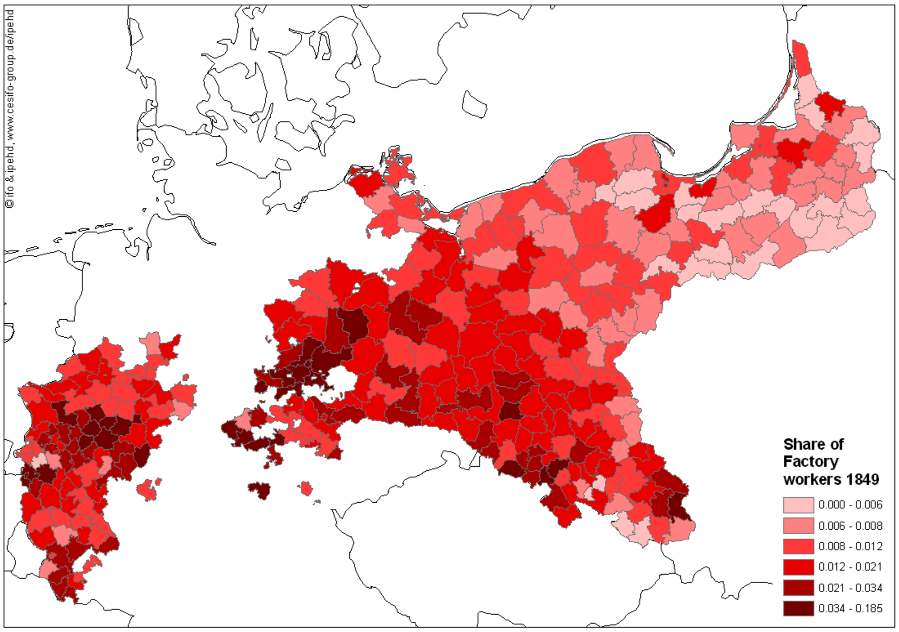
This illustrates the rapidly changing nature of Prussia throughout the nineteenth century, and the setting for the empirical analysis conducted here. The measures of industrialisation used in the analysis stem primarily from an 1849 census of Prussian counties, dating them in the middle of the national transformation.

Also of particular note is the regional heterogeneity. Lecce and Ogliari (2015) detail the long history of regional variation amongst the Germanic counties, in particular noting the varying change in institutions within Prussia over time. They note in particular (p.9) that "the transition from the ancien regime to the modern era evolved at different speeds across territories and, in some states, the transplant of French institutions failed." The widely varying religions and institutions across Prussia provide an ideal setting for empirical analysis, and the weaknesses of not adopting a regional approach have been highlighted by Tipton (1976) as a limitation of previous assessments of Prussia. The variation in technological development of particular interest in this analysis is highlighted in Figure 1, which shows the proportion of workers employed in factory occupations in 1849 by county; in Prussia in 1849 some counties had close to no recorded factory workers, whereas others had more than 15 per cent of their workers in factories.

2.2 Inequality and Development

Economists have long been interested in the relationship between inequality and economic growth, both on the effect of inequality on economic growth and how economic growth impacts inequality. The literature is vast and conflicting,

Figure 1: Share of Factory Workers in Prussian Counties (1849)



Notes: Reproduced from https://www.cesifo-group.de/dms/ifodoc/iPEHD/Maps/. This map uses different county boundaries to the aggregated counties used in the quantitative analysis for reasons discussed in Section 3.1, so the figures are not directly comparable.

highlighting the endogenous relationship between the two factors, of critical importance when conducting empirical analysis.

The purpose of the present analysis is to assess the impact of inequality on the spread of industrialisation, using a broad range of measures rather than focusing just on economic growth. Nonetheless, the literature on the relationship between these two key economic indicators is important to discuss, as the mechanisms by which economic growth and inequality interact bear relevance for the effect of inequality on other measures of the spread of industrialisation.

The effect that growth has upon inequality is most famously investigated by Kuznets (1955), who examines the relationship between economic growth and inequality and puts forward the 'Kuznets Curve', an inverted-U relationship between growth and inequality over time, suggesting that as countries begin to industrialise, inequality increases, before economic growth diffuses and begins to make the nation more equal. On the other hand, Okun (1975, p.1) posits that there is a trade-off between efficiency and equality. He states that "in many cases, the institutional arrangements represent uneasy compromises [between equality and efficiency], rather than fundamental inconsistencies. The contrasts among American families in living standards and in material wealth reflect a system of rewards and penalties that is intended to encourage effort and channel it into socially productive activity". Hence where Kuznets sees economic growth driving inequality, Okun posits that inequality is driving economic growth.

From a cross-sectional perspective, Dabla-Norris et al. (2015) find that inequality in the current day is negatively associated with the growth rate, and that policies which reduce inequality do not necessarily lower the growth rate. However this has been contested, for instance Forbes (2000) examines a comprehensive panel dataset and find that there is in fact a positive relationship between economic growth and inequality.

In the same historical period examined by this paper, Voigtländer and Voth (2006) examine the reasons that England was the first nation to industrialise, tangentially investigating the effects of inequality on technological development. Putting forward a theoretical model of endogenous growth, they posit that the key reasons that England industrialised prior to other nations were chiefly demographic factors, rather than resource endowments or unpredictable events (such as a bumper harvest year). Specifically on the effects of redistributions they find that the likelihood of very good outcomes was higher with a more generous system of redistributive institutions, but that the mean rate of technological progress was unaffected. However, it is important to note that these predictions result from a theoretical model calibrated to parameters of nineteenth-century England, so it is unclear how well they bear relevance to Prussia, which had a very different set of political and economic institutions.

What is clear throughout all of the literature on the relationship between economic growth and inequality is that there is an inherent endogeneity that impedes simple cross-sectional analysis. The Kuznets curve predicts that the inequality of a society will change as it grows, and scholars like Okun (1975) posit that the inequality in a nation is a key instigator of economic growth. When attempting to analyse how one affects the other, it is therefore key to adopt econometric tools designed to cope with endogeneity issues - this analysis will adopt an instrumental variable approach to address these concerns.

2.3 Land Inequality as a Measure of Inequality

The analysis follows Cinnirella and Hornung (2016) in focusing on land inequality over other measures as the variable of interest. Whilst Cinnirella and Hornung (2016) rely on this measure as a proxy for the strength of serfdom in Prussia during the period, this measure is not widely used as a proxy used in the economic institutions literature, and at the time the extent of serfdom throughout Prussia varied widely, not necessarily coinciding with large landed estates. For instance, Baranowski (2001, p.150) details how large landed estates relied on a combination of permanent workers and day-labourers, where "a sufficient labour pool at the disposal of primary producers was essential, but growers increasingly assessed that supply in terms of cost effectiveness, choosing short-term for long-term commitments wherever possible". This compromises the argument that large landed estates necessarily employed serfs, and that an increased share of large estates is indicative of the stronger incidence of serfdom in a region. Baranowski (2001, p.148) specifically highlights the "creation of a low-cost, flexible labour force" as one of the four components to agrarian modernisation, driven in part because there was a significant change in the make-up of estate owners, away from the tradition of only noblemen owning serf run estates to the point where "commoners significantly outnumbered noblemen [as estate owners], save for estates exceeding 1,000 hectares, which remained in the possession of the oldest Junker families". These commoner run estates could have significant land holdings, and frequently employed salaried workers, not serfs. Tipton (1976, p.24) also provides evidence for the limited relevance of serfdom as an economic institution in Prussia, noting "[s]ince the large landlords and large peasants gained virtually free disposition over their land through the law of 1807 on the one hand and rapid emancipation on the other, however, it appears that nearly all the land stood unencumbered by 'motley feudal ties' very early in the century". Therefore the available historical evidence undermines the association that Cinnirella and Hornung (2016) assign between their constructed measure of land inequality and the economic institution of serfdom.

Nonetheless, landholding inequality is an interesting measure in its own right for several reasons. As Grant (2002, p.5) notes, there are significant issues with trying to measure income inequality in the time period (specifically using tax data to construct such a measure):

> There are three main problems with the use of tax data to examine income inequality. Firstly, part of the population is normally exempt from the tax, due to income levels below the tax threshold, or some other reason for exemption. So, coverage is inevitably incomplete. Secondly, as the income assessment is linked to the tax payment, there is an obvious incentive to minimise declared income and conceal sources of income wherever possible. Some downwards bias must be anticipated. Thirdly, the tax system tends to change over time, which has effects on coverage, incentives to evade, on definitions of income, on procedures for checking returns and so on.

None of these issues arise using landholding inequality: there is no exemption as all estate sizes are measured, the declaration of the estate size does not face the same incentives for untruthful reporting, and the size of an estate is not affected by change over time.[1] Additionally, a Gini coefficient measures the degree of inequality through the whole distribution, whereas this analysis focuses in particular large landholders. This is relevant when considering the spread of industrialisation as these large landholders would be important investors in the new technologies, and important in affecting the economic and political institutions, as detailed below. Finally, Cinnirella and Hornung (2016) do find casual impacts of these large landed estates on educational development, so even

[1] Grant (2002) continues to address possible means of overcoming these issues, but notes that all of these means are limited to some extent.

if there is no direct tie to their existence and the preservation of serfdom, there are likely other channels through which powerful landowners exerted influence.

2.4 Institutions and Inequality

Institutional economics provides a highly plausible causal mechanism through which inequality can affect technological development. Acemoglu et al. (2005) state that the economic and political institutions are a key determinant of economic growth in societies, and that the institutions present in a given society are closely related to the degree of inequality in political and economic power. They note that (p.451), "the distribution of political power at any date is influenced to a large degree by the distribution of resources in society, since those with greater resources can command more power both through legitimate and intimate means". Here the authors' mechanism by which actors profit from inequality relates to their ability to exercise political power. They further note that (p.451) "the distribution of resources at this point is influenced by economic institutions and economic outcomes of the past". Hence, economic institutions and outcomes influence the inequality in society, and the inequality in society influences the distribution of political power amongst economic agents. These groups then leverage this political power to implement policies which appropriate economic rents for themselves, and the cycle begins anew.

Ogilvie and Carus (2014) highlight the key importance of generalised institutions over particularised institutions for economic growth, where generalised institutions consist of rules applying uniformly to all economic agents and particularised institutions apply to only a subset of economic agents. Whilst generalised institutions are frequently better for economic growth, agents may seek to introduce particularised economic institutions in order to appropriate economic rents for distributional purposes. Therefore inequality could represent the presence of particularised institutions, in that these institutions benefiting some subset of agents at the expense of others creates inequality. For instance, slavery represents a particularised economic institution, applying only to slaves, which inevitably creates income inequality. Alternatively, inequality could be the precursor to particularised institutions being created, since preexisting inequality afford some economic agents with the political power and economic influence necessary to sway rulers to preserve or create particularised economic institutions. As with Acemoglu et al. (2005), the cyclical nature is clear, with inequality being both an output of economic and political institutions and a determining factor of the institutions in society.

A key difficulty in the analysis of institutions is the scarcity of quantitative observations. To that extent, landholding inequality may represent a proxy for the presence of political power and paticularised economic institutions that Acemoglu et al. (2005) and Ogilvie and Carus (2014) note as harmful for economic growth. There is historical evidence that this relationship existed. Dumke (1988, p.7) details how income inequality in nineteenth-century Germany directly resulted in favourable treatment from political institutions:

> In the German political context, where severe property, residence, and income restrictions were placed on voting rights for the state and local elections, a tendency towards income concentration automatically implied a greater degree of political inequality. This was particularly true of the Prussian three-class electoral system employed in the Prussian state and local elections. There were three classes of voters in these indirect elections. The top third comprised the total number of well-to-do persons contributing one-third of the local tax revenues; the second class of voters were those contributing the second third of revenues, etc. Each class of voter selected the same number of electors, who, in turn, chose the representatives. In this system, therefore, the vote of a rich industrialist like Friedrich Krupp was weighed by his contribution to local taxes.

Therefore, by examining the inequality in Prussia and its effect on the Industrial Revolution, it is highly plausible that the chief mechanism through which inequality affected this spread is through the preservation or creation of economic institutions. With the data available it is not possible to separate competing claims of causal mechanisms, only to investigate whether causality can be measured in the first place.

3 Data

3.1 Overview

The data used for the analysis consists of the ifo Prussian Economic History Database (iPEHD). This dataset contains information at the county level on a variety of socioeconomic indicators, including education, religiosity and industrial output. These data were collected by the Royal Prussian Statistical Office through censuses throughout the nineteenth century, both before and after the unification of Germany. The dataset is documented thoroughly in Becker et al. (2012).

The iPEHD data contain records on 574 separate counties of Prussia in 1901 (the end of the dataset), but this number is not reflective of earlier years for two reasons. Firstly, over time the Prussian Statistical Office moved to more disaggregated counties, providing more detail. Secondly, in 1871, the unification of Germany added 103 counties to the dataset. For the purpose of comparing across counties using multiple censuses it is necessary to merge the disaggregated counties together to the earliest census used, and exclude the counties which were added in the unification of Germany. Therefore for the analysis conducted there are approximately 280 separate counties (varying slightly depending on the measure used). This aggregation of counties is the approach recommended for using data across multiple periods by Becker et al. (2012).

Whilst the iPEHD data contain censuses from sixteen different years in the nineteenth century, the 1849 census in particular contains many details on industrialisation that are not recorded for other periods.[2] Therefore, for the majority of the analysis conducted throughout the paper, 1849 is used. For the analysis presented here, the data of chief concern are those relating to employment, manufacturing, soil quality, landholdings and crop yields.

The low-level of the data permits a cross-sectional analysis, as there is significant regional heterogeneity across the counties throughout the data. The population of the counties totalled approximately 50,000 in 1849, which allows us to look at effects on a micro-scale.

Following Cinnirella and Hornung (2016), as a measure of inequality the analysis relies on the share of large landed estates relative to the total number of estates. Specifically, the estates are measured in Prussian Morgen (henceforth PM, where 1 PM is approximately 0.27 Ha). The 1849 census records five separate categories of estate sizes: <5 PM, 5-30 PM, 30-300 PM, 300-600 PM, and >600 PM. "Large" estates are defined in this analysis as estates which have an area greater than 300 PM, or 81 Ha.[3]

[2]For instance, details on the number and power of steam engines in Prussia is recorded only in the 1849 census, and not in later years.

[3]This is the definition of large estates used in Cinnirella and Hornung (2016) for 1849, although they adopt a different measure in later periods due to changes in the measurement of estates from PM to Ha.

Table 1: Descriptive Statistics of Industrial, Demographic and Soil Data in Prussian Counties, 1849

	Statistics				
	mean	sd	min	max	count
Population	58057.11	42093.19	13647	423902	279
Share of Large Estates	0.021	0.022	0	0.114	277
Proportion of Loamy Soil*	0.284	0.217	0	0.911	280
Proportion of Sandy Soil*	0.240	0.229	0	0.809	270
Number of Ironworks	2.581	8.177	0	82	279
Number of Steelworks	1.065	8.003	0	114	279
Number of Textile Handlooms	607.480	1366.814	0	8122	279
Number of Textile Power Looms	15.821	84.742	0	928	279
Power Looms-Handlooms Ratio	0.021	0.099	0	0.938	277
Number of Factories	150.689	149.917	0	1557	280
Share of Workers in Factories	0.024	0.040	0.001	0.364	279
Share of Workers in Agriculture	0.412	0.166	0	0.849	279
Number of Steam Engines	7.003	18.944	0	171	280
Steam Engine Power (HP)	239.557	902.020	0	9864	280

Notes: All data come from 1849 censuses, except where marked above with * referring to the variables from the soil assessment carried out in 1866.

3.2 Summary Statistics

Summary statistics for the data are provided in Table 1. Of particular note is the high variance in the share of landed estates, which is essential for reliable estimation of causal effects. With regard to the measures of investments in technology outlined in Section 3.4.2 below, a number of them have numerous counties with a recorded value of zero. The implications of these low values are investigated further when conducting the analysis in Section 4.3.

Table 1 also highlights the regional heterogeneity discussed earlier. The standard deviation for most of the variables is greater than the mean, indicating that the variation across the data is extremely high. Looking at population for instance, whilst the mean population of a county is 58,000, the standard deviation is 42,000, with a range from 13,647 to 423,902. Thus the largest county in the data has more than thirty times the population of the smallest, with a wide range of values in between. The variation in these variables greatly facilitates conducting quantitative analysis, as a smaller variation in the data would not permit precise estimation of effects. For instance, if the soil quality was relatively homogenous throughout Prussia then the statistical significance of the IV estimate would be reduced.

3.3 Soil Quality as an Instrumental Variable

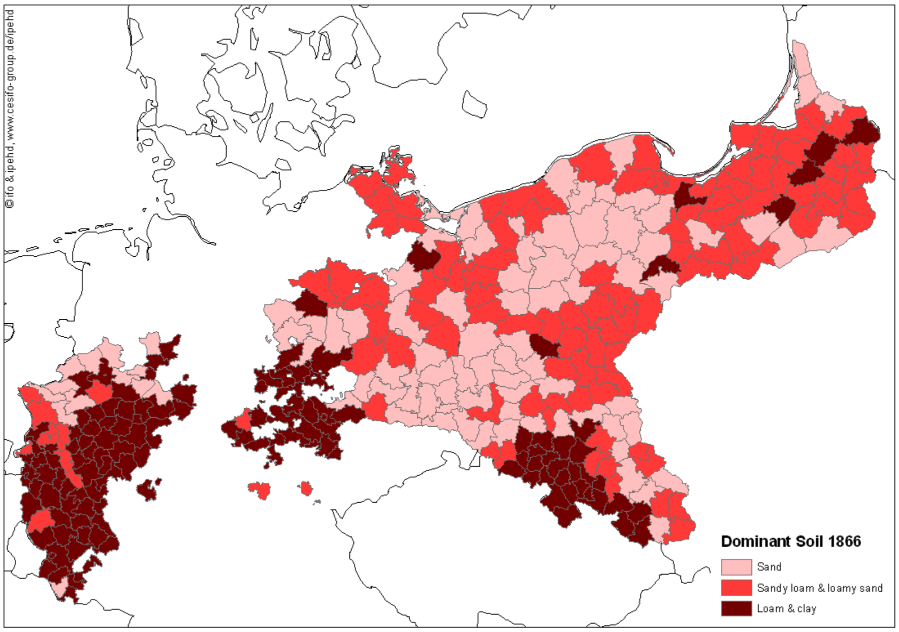
As an instrument for the land inequality in Prussia, this paper uses the relative soil quality. The relationship between the suitability of the land for growing crops and inequality is put forward initially by Engerman and Sokoloff (1994), who discuss how these resource endowments can affect the resulting societal structure. Whilst their paper does state that factor endowments can explain the widely varying economic success of economies across the American continent, this is not a prerequisite for its use in empirical analysis. For its use as an instrumental variable, it is only required that it has a measurable effect, not that the effect dominates those of other possible determinants of economic growth (as they attest).

The empirical relationship between soil quality and inequality is tested by Easterly (2007), who proposes its use as an instrumental variable for assessing the effects of inequality. Utilising data on the suitability of soil types for growing wheat and sugar across a large cross section of countries, he finds not only that soil quality predicts inequality, but that this inequality also affected schooling and economic institutions. Cinnirella and Hornung (2016) use this empirical relationship in an application to the Prussian census data used in this paper, using the area of highly fertile loamy soil in a county relative to its total size, as given by an 1866 census. They note that since regions with a poor soil quality have a lower marginal value of land, there was a lower demand for land, causing these areas to become developed later than lands with higher quality soil. This later development resulted in larger average farm sizes and higher inequality in landholdings. Regions with a good quality of soil for growing crops experienced a greater demand for land, causing greater fragmentation in landholdings and earlier security of property rights. In their analysis they investigate the relationship between landholding inequality and the expansion of education, finding that a higher share of large estates in a region causes a reduction in the expansion of education throughout the nineteenth century.

In order for an instrument to be effective at creating a causal estimate, it must be both relevant and valid. The instrument must be relevant in that there is a correlation between the quality of soil and the land inequality; The instrument must be valid in that there is no correlation between the quality of soil and the independent variables through no other means than the variable being instrumented, accounting for controls.

Importantly, the exogeneity of the instrument is almost guaranteed. The underlying loamy property of soil is relatively unchanged by agricultural processes in the short term, and by using contemporaneous estimates taken in 1866 there are no resultant interactions between the long term use of the soil that would arise if modern figures were used. For instance, if instead the analysis used recent assessments of the soil quality across modern day Germany, it is possible that areas with a lower soil quality in the nineteenth century have since been treated or otherwise changed such that this variation is no longer detectable. By using the 1866 assessment these possible interactions could not have occurred. Hence the relative loamy soil proportion of the land can be viewed as an exoge-

Figure 2: Predominant Soil Characteristic in Prussian Counties



Notes: Reproduced from
https://www.cesifo-group.de/dms/ifodoc/iPEHD/Maps/ipehd_soil1866.gif.
This map uses different county boundaries to the aggregated counties used in the quantitative analysis for reasons discussed in Section 3.1.

nous influence upon inequality, so long as validity is also guaranteed. Figure 2 shows the variation in the predominant soil characteristic in each of the Prussian counties as recorded in 1866 (note that for the quantitative analysis the exact proportion of loamy soil is used, not the predominant type).

3.3.1 Validity of the Instrument

For the instrument to be valid it is necessary for there to be no correlation between the soil quality and the residuals of the regression. This means that soil quality's only effect on any of the dependent variables examined operates only through its effect on the inequality of a region.

The main channel through which validity may be violated is controlled for - by crop yields. Higher soil quality affects the productivity of the agricultural land and without controlling for this difference in yields it would be highly likely that there would be substantial correlation between the residuals and the instrument. It is also possible that there is some remaining correlation between the quality of crops and the quality of soil. For instance, if the yield remains the same but the crops grown on loamy soil are of higher quality, then there would also be residual correlation. It is not possible to control for this potential source of endogeneity however, as there are no recorded details on the price of the crops sold or the relative quality of the crops grown.

3.3.2 Relevance of the Instrument

For the instrument to be relevant there needs to be both a theoretical and an empirical relationship between the quality of soil and the concentration of landholdings. The theoretical relationship is given by the literature as discussed above. The empirical relationship is investigated here.

To measure the quality of soil, this analysis uses the proportion of land area in a county that has soil described in the 1866 census as either loamy, or loamy with sand. This is the same methodology used in Cinnirella and Hornung (2016), as the relative loam content for the soil is a good indicator of its suitability for growing crops. The suitability for growing crops then induces increased economies of scale for landholders, creating incentives for greater concentration of landholdings.

Figure 3 displays the observed positive correlation in the Prussian data between the proportion of loamy soil in a county and the standardised share of large estates. Whilst this does not account for other potential factors (which are controlled for in the first stage of the IV regression), this does highlight the strong relationship in the data between these two variables.

3.4 Measurements of the Spread of Industrialisation

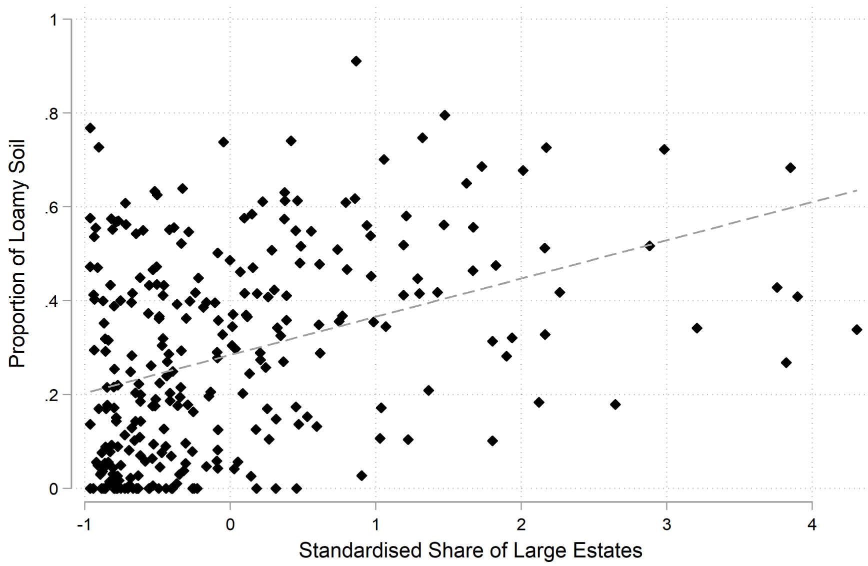
The goal of this analysis is to assess the causal effect of inequality on the spread of industrialisation, not just on a single measure but on a broader assessment, to see if inequality has greater impact on certain aspects of industrialisation than on others. The iPEHD dataset contains many plausible indicators for this spread, which means it is possible to undertake the analysis using a variety of approaches.

In order to assess the impact on a broad range of measures, this paper uses three categories of dependent variables: occupational structure, technological investments, and textiles by methods of production. These are discussed separately below.

3.4.1 Occupational Structure

These measures use detailed employment statistics to examine heterogeneity in occupational structure across Prussian counties. Of particular interest is the proportion of workers employed in factories, and the proportion of workers employed on farms. As industrial technologies became more widespread, increasing number of workers were employed in factories and fewer were required to work in agriculture, so these measures have the potential to examine that effect of industrialisation. As Tipton (1976, p.17) notes “[t]he most dramatic change in the distribution of the labor force connected with modern economic growth

Figure 3: Relationship between Soil Quality and Land Inequality



Notes: Scatter plot of the relationship at the county level between the proportion of loamy soil in the county and the standardised share of large estates. Grey dashed line indicates a linear fit to the data. The share of large estates is standardised such that the mean is zero and the variance is one.

has been the decline of the share of agriculture in employment, commonly from seventy-five to less than ten per cent".

To calculate these measures, for factory employment the number of employees across all the types of factories recorded in the 1849 census is aggregated. For farm employment the specific figure is the number of people recorded with farming as their main occupation. To calculate the proportion of workers, the denominator used is the total 1849 population. The estimated proportion therefore is an underestimate of the true value therefore, as the total population will include those not of working age.

Figure 4 shows how the proportion of workers employed in agriculture and factories varies across Prussia. There is significant heterogeneity in both proportions, with the mean proportion of farm workers in a county as 42%, and the mean proportion of factory workers as just 2%. Intuitively, there is a negative correlation between the proportion employed in factories and the proportion employed in farms, with a coefficient of -0.45, suggesting that some counties had a more agricultural economy than others.

Figure 5 shows the importance of the changing occupational structure over the course of the nineteenth century, by comparing the number of agricultural employees in 1849 and 1882. The dashed line on the figure indicates the 45-

Figure 4: Relationship between the Relative Employment of Agricultural and Factory Workers in 1848

Notes: Grey dashed line indicates a quadratic fit of the data.

degree line, so counties below this line had fewer agricultural jobs in 1882 than in 1849. As is clear, almost all of the counties saw an absolute decline in the total number of agricultural jobs (271 out of 280), despite a rising population over the entire period. This highlights the importance of changing occupational structure as an key facet of the Industrial Revolution.

3.4.2 Technological Investments

These measures refer to the possession and utilisation of new industrial technologies. In order to assess as broad a base of technological investments as possible, this paper utilises four separate measures. These are:

1. **Steam Engine Power**. This variable measures the power of the steam engines in a county, in horsepower. The aim of including this variable is to see if there is an effect not just on the quantity of investments in new technologies but additionally on the quality of investments.

2. **Number of Steam Engines**. This variable measures the total number of steam engines in a given county, and seeks to measure the rate of investment in new machinery.

3. **Number of Ironworks**. This variable records the number of ironworks in a county.

Figure 5: Decline in the Number of Agricultural Employees by County from 1849 to 1882

Notes: Total number of employees with farming as their main occupation in 1849 compared to the total of employed persons in agriculture in 1882. Grey dashed line shows the 45 degree line. A single county with very high employment in both periods is excluded to aid readability.

4. **Number of Steelworks**. This variable records the number of steelworks in a county.

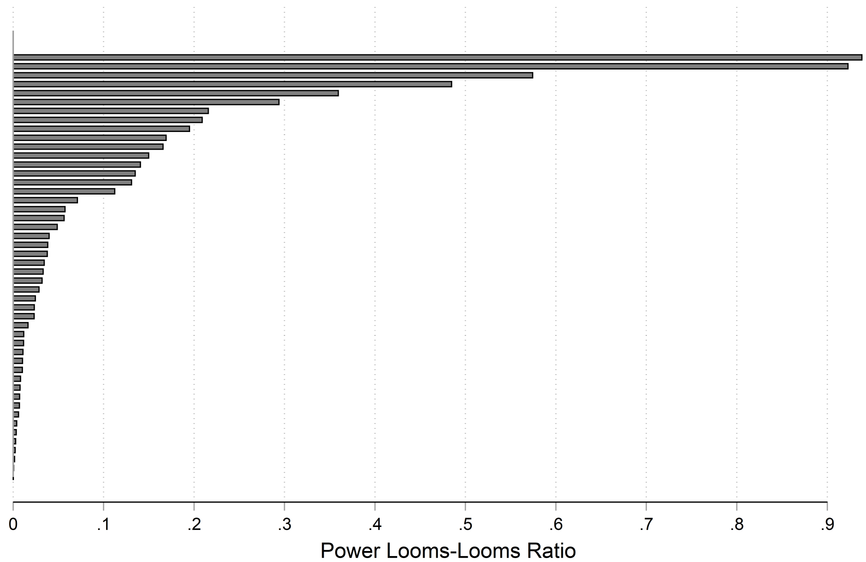
3.4.3 Textiles by Method of Production

These measures refer to the method of output for various goods recorded in the data. For certain products, the total production of a good is broken down into the various methods of producing the good, allowing the measurement of the adoption of new production processes as a proxy of the spread and adoption of new industrial technologies.

The measures used in the analysis relate to the production of textiles. Traditionally manufactured by handloom, in 1849 there was the beginning of a shift towards factory production on powered looms.[4] To analyse this trend, the analysis examines the number of handlooms used to produce textiles in a county, the

[4]The iPEHD data contains records on "looms", "power looms", and "handlooms". The documentation of the data implies that "looms" refer to unpowered handlooms, of a larger size than the separately distinguished "handlooms" (although still operated by hand). For the analysis in this paper, the total of "looms" and "handlooms" are together combined to

Figure 6: Power Loom-Handloom Ratio in 1849, by County



Notes: This chart shows the power loom-handloom ratio for the 48 counties which had any power looms recorded in 1849. The remaining 232 counties with a ratio of zero are not shown to aid readability.

number of power looms used to produce textiles in a county, and the proportion of the total production of textiles manufactured in factories.[5]

The data exhibits a significant long tail in the proportion of textiles manufactured by power loom. As Figure 6 shows, there are a small number of states with a very high number of power looms per handloom (almost at parity amongst the highest), a moderate number with between 0.1-0.2 power looms per handloom, and the vast majority with no power looms whatsoever.

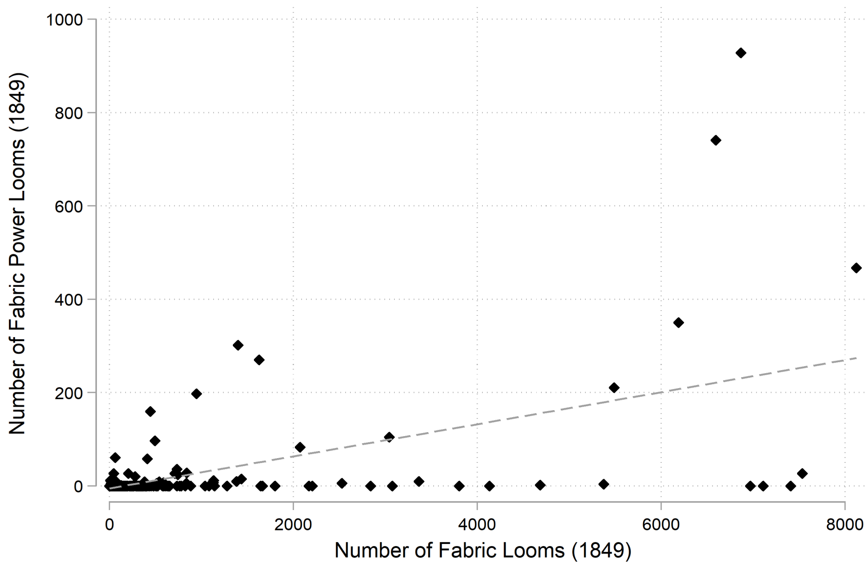
Furthermore, the relationship between textile manufacturing becoming industrialised and the total textile production is complex. As Figure 7 shows, whilst there is a positive correlation between the textiles produced by either output method, there are many counties with a relatively small production of textiles by handloom which nonetheless were leading in the number of power looms, and counties producing vast quantities of textile by handloom with relatively little output by factory.

The volume of textiles by means of production would be a more effective measure of technological development than the number of handlooms compared

form a total number of unpowered looms. This paper adopts the term "handlooms" to refer to this aggregated figure for the purposes of providing a clear distinction between powered and unpowered looms.

[5]Throughout the analysis, textiles refer to the total across silk, cotton, linen and wool.

Figure 7: Textile Production by Production Method, 1849



Notes: The grey dashed line indicates a linear fit of the data.

to the number of power looms, but this information was not recorded in the 1849 census. Whilst it may be possible to construct estimates of the relative productivity of looms and power looms, this would entail enforcing strict assumptions regarding the productivity of the various methods, for which there are not detailed records. Furthermore, this conversion would not remain constant across Prussia, as there was significant variation in power loom quality at the time.

4 Analysis

This section presents the results of the quantitative analysis. It reports estimates of a two stage least squares regression, with the measure of inequality instrumented by the proportion of loamy soils.

4.1 Model Specification

The model used is broadly the same across the analysis of all the various dependent variables. First, an OLS regression is performed, controlling for the population of the county and the agricultural yields. This OLS regression is as specified in Equation 1. Then, a two stage least squares instrumental variable regression is performed, using the proportion of loam in the soil as an instrument for the level of inequality in 1858, where the level of inequality is the relative proportion of large estate owners. The structural form of the instrumental variable regression is given in Equations 2 and 3. In the equations, Y_i refers to the relevant dependent variable of interest, $land\_share_i$ refers to the share of large estates as a proportion of all estates (standardised such that the mean is zero and the variance is one), $population_i$ is the population of a given county, $loam\_prop_i$ is the proportion of the total land area consisting of loamy soil or sandy-loamy soil, and $yields_i$ refers to the wheat yields. The regression output shows the first-stage F-stat as an indication of weak instruments, as well as the R squared in the first stage regression.

$$Y_i = \beta_0 + \beta_1 land\_share_i + \beta_2 population_i + \beta_3 yields_i + \varepsilon_i \tag{1}$$

$$\widehat{land\_share}_i = \gamma_0 + \gamma_1 loam\_prop_i + \nu_i \tag{2}$$

$$Y_i = \beta_0 + \beta_1 \widehat{land\_share}_i + \beta_2 population_i + \beta_3 yields_i + \nu_i \tag{3}$$

4.2 Occupational Structure

In order to assess the effects of inequality on occupational structure in Prussia, this paper utilises two constructed measures of the relative industrialisation of the workforce. The first relates to the number of factory workers in a given Prussian county relative to the total population in the county, and the second relates to the number of farm workers (as a primary occupation) relative to the total population in the county. The OLS and IV results are presented in Table 2.

Panel A of Table 2 shows the results of the OLS regression. Whilst the coefficients cannot be interpreted causally, it is interesting to note the magnitude of the coefficients compared with those in the later IV results.

Panel B of Table 2 shows that there is a highly significant relationship between the proportion of loamy soil in a region and the share of large estates in 1958, as predicted by the agricultural economics literature discussed in Section

Table 2: Effects of Inequality on Occupational Structure

	(1) Factory Workers (Proportion)	(2) Farm Workers (Proportion)
Panel A: OLS		
Share of Large Estates (1858)	-0.010***	0.049***
	(0.002)	(0.009)
Observations	279	279
R-squared	0.07	0.10
Panel B: First Stage IV		
Proportion of Loamy Soil	1.548***	1.548***
	(0.239)	(0.239)
Observations	279	279
R-squared	0.20	0.20
Panel C: Second Stage IV		
Share of Large Estates (1858)	-0.015**	0.059**
	(0.006)	(0.029)
Observations	279	279
First-stage F-stat	42.1	42.1

Notes: The table shows OLS and first and second stage IV results using county-level data. Landownership concentration is instrumented using the proportion of total land area consisting of loamy soils. Controls are included for the county population in 1849 and the wheat yields. The share of large estates is standardised such that the mean is zero and the variance is one. *, **, and *** indicate significance at the 10%, 5%, and 1% level respectively.

2.[6]

Panel C of Table 2 details the results of the second stage IV regression. The results of this analysis find that there is a causal negative relationship between the share of large estates and the proportion of the population working in factories in 1849, and a causal positive relationship between the share of large estates and the proportion of the population working on farms in 1849.

The effects of inequality on occupational structure are the most statistically significant of the regressions performed throughout this paper. Both the effect on the proportion of workers in factories and in farms are highly statistically significant, and intuitively the coefficients are of opposite direction. For every increase in the share of large estates by a standard deviation, there is an implied reduction of the proportion of factory workers by 1.1 percentage points. The effect on farm workers is even more significant - for every increase in the share of large estates by a standard deviation, there is an implied increase in the proportion of farm workers by 5.9 percentage points.

From the perspective of assessing the robustness of the two results, it would

[6]Panel B in each of Table 2, Table 3 and Table 4 reflect the same first stage regression of land share on loam proportion in a county, so the results across the three tables differ only marginally, reflecting the slight variation in the number of observations across the regressions.

appear that the factory worker result is less likely to suffer from potential omitted variable bias. Despite the fact that the regressions are controlling for differences in observable agricultural productivity (by including agricultural yields as a control), it is possible that there are other means by which the higher quality soil is causing both more farm jobs and larger landed estates.

For the factory workers, the result appears robust - mechanisms by which loamy soil affects the productivity of factories (asides from the controlled for factors of differing wealth and population) are harder to envisage, and this disparity in the robustness could explain some of the difference in the magnitude between the two coefficients, as to why the positive effect on farm workers is almost six times greater than the negative effect on factory workers. Other potential explanations for the disparity in the magnitude of the coefficients are the lower labour intensity of factory work relative to large-estate agriculture and the burgeoning services sector of the economy in the time period, either of which would mean that the shift away from agricultural jobs would be greater than the increasing employment in factories.

Table 3: Effects of Inequality on Investments in Technology

	(1) Steam Engine Power	(2) Number of Steam Engines	(3) Number of Ironworks	(4) Number of Steelworks
Panel A: OLS				
Share of Large Estates (1858)	-136.153***	-3.275***	-0.535*	-0.581**
	(50.161)	(0.968)	(0.276)	(0.294)
Observations	279	279	279	279
R-squared	0.10	0.23	0.03	0.02
Panel B: First Stage IV				
Proportion of Loamy Soil	1.548***	1.548***	1.548***	1.548***
	(0.239)	(0.239)	(0.239)	(0.239)
Observations	279	279	279	279
R-squared	0.20	0.20	0.20	0.20
Panel C: Second Stage IV				
Share of Large Estates (1858)	-123.593	-1.635	-2.381	-3.781*
	(119.232)	(2.805)	(1.572)	(1.973)
Observations	279	279	279	279
First-stage F-stat	42.1	42.1	42.1	42.1

Notes: The table shows OLS and first and second stage IV results using county-level data. Landownership concentration is instrumented using the proportion of total land area consisting of loamy soils. Controls are included for the county population in 1849 and the wheat yields. The share of large estates is standardised such that the mean is zero and the variance is one. *, **, and *** indicate significance at the 10%, 5%, and 1% level respectively.

4.3 Technological Investments

In order to assess the effects of inequality on technological investments in Prussia, the analysis examines four measures of new technologies across the various counties. The average horsepower of a steam engine in a county, the total number of steam engines in a county, the number of ironworks in a county, and the number of steelworks in a county.

Panel A of Table 3 shows the results of the OLS regression. There is a negative association between inequality and all four of the dependent variables, although the effect on the number of ironworks is only statistically significant at the 10 per cent level. The R-squared for the number of ironworks and steelworks is very low, indicating that the determining factors in these industries may not be captured by the regression.

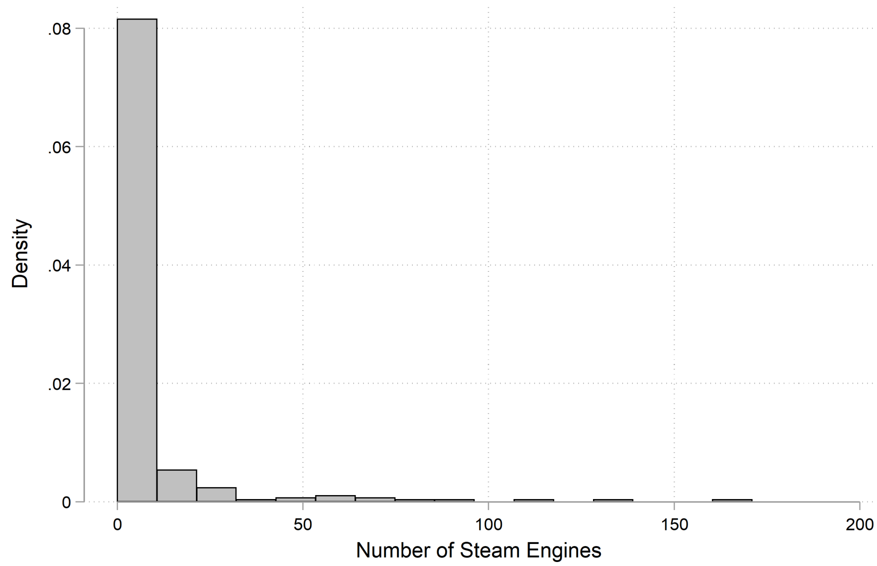
Panel B of Table 3 shows the first stage of the IV regression, the same as for Table 2 in that there is a highly statistically significant relationship between the proportion of loamy soil and the share of large estates in Prussian counties.

Panel C of Table 3 shows the IV estimates for the causal effect of the share of large landed estates on the various dependent variables. There is no statistically significant effect of inequality on the number of ironworks, and only at the 10 per cent level for the number of steelworks. The estimated effect on steelworks implies that for every standard deviation increase in the share of large estates, there are 3.8 fewer steelworks in a county. Whilst this coefficient has weak statistical significance, it is worthwhile noting that the magnitude is large, recalling from Table 1 that the mean number of steelworks in a county was just 1.1.

There is additionally no observed statistically significant effect of inequality on either the quantity or quality of steam engines. This is either because inequality had no effect on the decision to invest in steam engines, or perhaps because a large proportion of the counties in the data had no steam engines at all. The relative distribution of the number of steam engines is shown in Figure 8, and clearly highlights that in 1849 the modal number of steam engines in a county is close to zero, with a small number of counties accounting for the vast number of engines. Thus whilst this measure may in some regards account for the spread of new technology, in 1849 the technology was not sufficiently diffused for the relative inequality to have any material effect. This conjecture of insufficient diffusion is supported by historical evidence in the Saxon region by Tipton (1976, p.35), who notes "in 1846 there were one hundred ninety-seven steam engines in Saxony with two thousand four hundred horsepower and in 1862, two thousand four hundred engines with forty-six thousand horsepower". Thus the census in 1849 may not be measuring an effect of inequality upon steam engine purchases simply because there were so few purchases the effect is undetectable, and later censuses do not record the number of steam engines at a sufficiently disaggregated regional level to conduct quantitative analysis.

The lack of a statistically significant effect is in stark contrast to the effect found on occupational structures outlined in Table 2, and this disparity is discussed further in Section 5.

Figure 8: Distribution of Steam Engines Across Counties (1849)



4.4 Textiles by Method of Production

In order to assess the effects of inequality on the relative proportion of materials by method of production, the analysis utilises the data on textile output by production method. Table 4 shows these results.

Panel A of Table 4 shows the results of the OLS regression. This finds a highly statistically significant association between inequality and fewer handlooms and power looms. There is no statistically significant relationship between inequality and the ratio of power looms to handlooms.[7]

The results in Panel C of Table 4 are highly different to the results of the OLS regression. When attempting to find any causal impact of inequality on the methods of textile production, there is no statistically significant relationship for any of the examined dependent variables (at the 5 per cent level). This indicates that the OLS results in Panel A are not indicative of any structural relationship between the two variables, and could reflect endogeneity in the regressors instead.

This would seem to indicate that, from the perspective of the manufacturing of textiles, it is not possible to detect a significant impact of land inequality. There are negative coefficients on the effect of inequality on the quantity of

[7]The coefficient on the first stage result for the ratio of power looms to handlooms as reported in Table 4 appears to be exactly zero in the results table due to rounding, but the reported result was marginally negative (-0.0001028).

Table 4: Effects of Inequality on Textile Methods of Production

	(1) Handlooms	(2) Power Looms	(3) Power Loom - Handloom Ratio
Panel A: OLS			
Share of Large Estates (1858)	-312.450***	-11.826***	0.000
	(71.058)	(4.259)	(0.005)
Observations	279	279	277
R-squared	0.13	0.17	0.01
Panel B: First Stage IV			
Proportion of Loamy Soil	1.548***	1.548***	1.539***
	(0.239)	(0.239)	(0.239)
Observations	279	279	277
R-squared	0.20	0.20	0.20
Panel C: Second Stage IV			
Share of Large Estates (1858)	-497.921*	-4.308	0.014
	(272.483)	(16.922)	(0.014)
Observations	279	279	277
First-stage F-stat	42.1	42.1	41.38

Notes: The table shows OLS and first and second stage IV results using county-level data. Landownership concentration is instrumented using the proportion of total land area consisting of loamy soils. Controls are included for the county population in 1849 and the wheat yields. The first stage IV results vary slightly due to differences in the number of observations. The share of large estates is standardised such that the mean is zero and the variance is one. *, **, and *** indicate significance at the 10%, 5%, and 1% level respectively.

textiles manufactured by a loom and by a factory. This could indicate that there is no significant effect on the quantity of factory textiles, although it could also be a manifestation of the limited nature of factory manufacturing of textiles in 1849.

5 Conclusion

5.1 Avenues for Future Research

Whilst this paper has adopted an IV to permit a causal interpretation, it would also be possible to undertake the analysis using the panel nature of the data to control for county-level fixed effects. This approach would run into limitations relating to data availability. The majority of the data used throughout this analysis relates to the detailed manufacturing and industrialisation statistics which were only collected in 1849. Any panel-based approach could rely only upon the statistics which were collected at multiple time intervals by the Prussian statistics bureau. Therefore, it would be possible to utilise the panel-data approach for measuring the effects of occupational structure, but not on the effects of investments in new technologies or textiles by method of production.

As this paper has shown, the measured effects of inequality on industrialisation vary widely depending on the measure used, and therefore the panel approach could not be adopted to further investigate this phenomenon. Despite this limitation, the effect of inequality on occupational structure was the most statistically significant of the measures investigated, so it would warrant further research to see if this specific occupational effect persisted through later periods.

5.2 Final Remarks

The analysis has shown that inequality appears to have had a more significant impact on certain aspects of industrialisation than others. When looking at measures of occupational structure, there is a significant and robust negative impact upon industrialisation, both on increasing the proportion of agricultural jobs and reducing the proportion of factory jobs. However, these effects are not detectable when examining the manufacturing methods for textiles, nor the investments in new technologies such as steam engines, or the construction of new ironworks and steelworks. It would appear that the effect of inequality is limited solely to the occupational shift, and had no measurable effect on the other components of the Industrial Revolution. This paper investigates two candidate explanations for these measured effects: a lagged effect and institutional mechanisms.

The potential for a lag relates specifically to the timing of the analysis. In 1849 Prussia was transitioning from an agricultural to an industrial economy, and it is clear that this process was still in its infancy. The analysis throughout Section 4 shows that on the measures examined in this paper, many counties had yet to invest significantly in new technologies. It is therefore possible that these measures are simply slower to respond to technological change than occupational structure, which may have been more flexible. As noted in Section 2, many labourers in agriculture were on short term contracts, and could have changed jobs relatively quickly. Meanwhile, diffusion of new technologies such as the steam engine required more investment and occurred relatively later. It is possible then that whilst inequality impacted upon occupational structure and investments in new technologies, in 1849 only the effect upon occupational structure is detectable through quantitative analysis, and the investments in new technologies occurring later in the century are not. This possibility is supported by the large lagged effect of the institutional reforms examined by Acemoglu et al. (2011), who found that institutional reforms in parts of Germany at the start of the nineteenth century had no measurable effect on growth in 1850, but were detectable in 1875. Their paper therefore suggests that at least some effects of societal change take many years to propagate through society, and thus lends credibility to the hypothesis that inequality may effect occupational structure immediately, but take many years to effect the adoption of new technologies.

The potential institutional mechanisms by which large landholders could prevent industrialisation were outlined in Section 2. Inequality inevitably creates a disparity in political power, which is exacerbated by the Prussian political institutions which assigned more votes to larger landholders. These large landholders had a strong incentive to preserve the existing societal structure, even at the cost of economic growth, since they were able to appropriate a larger proportion of the economic rents from an agricultural society. It is feasible that when new technologies became available, these landholders chose to use their political power to ensure that the rate of change in occupational structure was reduced, delaying the spread of occupational industrialisation in these regions. These estate owners did not have influence to the same extent over the purchases of industrialists, so their ability to prevent investments in new technologies to reduce technological disruption was not as great as their ability to prevent a change in occupational structure. This also links to the lack of measured effects on textile production, as large estate owners also likely had less influence on the purchases of power looms than on their ability to prevent changes in occupational structure.

The interaction between industrialisation, economic institutions and inequality is complex, and the causal mechanisms by which inequality affected the occupational structure in Prussia are ultimately unobservable. The legal system in Prussia resulted in large landowners having a disproportionate say in political decisions, through which they would have been able to prevent a changing occupational structure which, whilst benefiting the economy as a whole, would have reduced their incomes substantially. The findings of this paper support the notion that inequality has an effect on industrialisation, though the mechanisms by which it does so are not as straightforward as impeding all means of technological development equally.

Bibliography

Acemoglu, D., Cantoni, D., Johnson, S., and Robinson, J. A. (2011). The Consequences of Radical Reform: The French Revolution. *American Economic Review*, 101(7):3286–3307.

Acemoglu, D., Johnson, S., and Robinson, J. (2005). Institutions as a fundamental cause of long-run growth. In Aghion, P. and Durlauf, S., editors, *Handbook of Economic Growth*, volume 1, Part A, chapter 06, pages 385–472. Elsevier, 1 edition.

Baranowski, S. (2001). Agrarian transformation and right radicalism: Economics and politics in rural prussia, 1830-1947. In Dwyer, P. G., editor, *Modern Prussian History: 1830-1947*, chapter 5, pages 109–125. Pearson Education Limited.

Becker, S., Cinnirella, F., Hornung, E., and Woessmann, L. (2012). iPEHD - The ifo Prussian Economic History Database. CAGE Online Working Paper Series 96, Competitive Advantage in the Global Economy (CAGE).

Cinnirella, F. and Hornung, E. (2016). Landownership concentration and the expansion of education. *Journal of Development Economics*, 121:135 – 152.

Dabla-Norris, E., Kochhar, K., Suphaphiphat, N., Ricka, F., and Tsounta, E. (2015). Causes and consequences of income inequality : A global perspective.

Dumke, R. (1988). Income inequality and industrialization in germany, 1850–1913: images, trends and causes of historical inequality. *Research in Economic History*, 11:1–47.

Easterly, W. (2007). Inequality does cause underdevelopment: Insights from a new instrument. *Journal of Development Economics*, 84(2):755–776.

Engerman, S. L. and Sokoloff, K. L. (1994). Factor endowments: Institutions, and differential paths of growth among new world economies: A view from economic historians of the united states. Working Paper 66, National Bureau of Economic Research.

Forbes, K. J. (2000). A reassessment of the relationship between inequality and growth. *American Economic Review*, 90(4):869–887.

Grant, O. W. (2002). Does Industrialisation Push Up Inequality? New Evidence on the Kuznets Curve from Nineteenth-Century Prussian Tax Statistics. Oxford University Economic and Social History Series 048, Economics Group, Nuffield College, University of Oxford.

Kuznets, S. (1955). Economic growth and income inequality. *The American Economic Review*, 45(1):1–28.

Lecce, G. and Ogliari, L. (2015). Institutional Transplant and Cultural Proximity: Evidence from Nineteenth-Century Prussia. CESifo Working Paper Series 5652, CESifo Group Munich.

Ogilvie, S. and Carus, A. (2014). Institutions and economic growth in historical perspective. In *Handbook of Economic Growth*, volume 2, chapter 08, pages 403–513. Elsevier, 1 edition.

Okun, A. M. (1975). *Equality and Efficiency: The Big Tradeoff*. Brookings Institution Press.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Harvard University Press.

Stiglitz, J. (2012). *The Price of Inequality: How Today's Divided Society Endangers Our Future*. W. W. Norton.

Tipton, F. B. (1976). *Regional Variations in the Economic Development of Germany During the Nineteenth Century*. Wesleyan University Press.

Voigtländer, N. and Voth, H.-J. (2006). Why England? Demographic Factors, Structural Change and Physical Capital Accumulation during the Industrial Revolution. *Journal of Economic Growth*, 11(4):319–361.

Voth, H.-J. (2001). The prussian zollverein and the bid for economic superiority. In Dwyer, P. G., editor, *Modern Prussian History: 1830-1947*, chapter 5, pages 109–125. Pearson Education Limited.