Industrialization and the return to labor: Evidence from Prussia*

Ann-Kristin Becker† Erik Hornung‡

Last update: 2025-10-30
Most recent version here.

Abstract

Industrialization boosts aggregate incomes, but its distributional effects remain debated. We study the impact of coal-driven industrialization on unskilled labor incomes using novel panel data on wages from 667 Prussian localities (1800–1879), extended with county-level data through 1914. Exploiting spatial variation in coal proximity in difference-in-differences and event-study designs, we find that wage gains in coal-rich regions emerged once industrialization accelerated in the 1850s and continued to grow until WWI. Evidence from 3,000 household accounts shows that coal proximity raised labor incomes primarily for low-skilled workers, with weaker effects for high-skilled and mechanical occupations. This pattern suggests that industrialization reduced wage inequality by compressing the local skill premium. Mediation analysis indicates that wage gains for unskilled workers were primarily driven by technology adoption and the increasing demand for low-skilled labor, rather than by sectoral change or the spread of the factory system.

Keywords: Industrialization, Labor income, Energy transition, Structural change, Technological change, Deskilling, Nineteenth-century Prussia

JEL Codes: C23, J31, N33, N73, N93, O13

*We would like to thank Sandra de Pleijt, Alan Fernihough, Philippe Quirion, Jakob Molinder, Ulrich Pfister, Carola Stapper, and seminar participants at Bonn, Cologne, the EHS Annual Conference 2023 (Warwick, UK), the FRESH Meeting 2023 (Cologne, Germany), the EHES Conference 2023 (Vienna, Austria), the EHA Annual Meeting 2023 (Pittsburgh, USA), the EAYE Annual Meeting 2024 (Paris, France), the ETRANHET workshop 2024 (Paris, France), and the Workshop in Economic History 2025 (Uppsala, Sweden) for helpful comments and suggestions. Erik Hornung acknowledges that his research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2126/1 - 390838866.

†University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany (ann-kristin.becker@wiso.uni-koeln.de)

‡University of Cologne, Albertus-Magnus-Platz, 50923 Cologne, Germany (hornung@wiso.uni-koeln.de)

1 Introduction

While industrialization is an important driver of economic development that boosts aggregate incomes, its distributional effects are contested. Widely used British data suggest that real wages stagnated during the early 19th century, despite rising per capita GDP, before increasing from the mid-19th century onward.[1] These dynamic patterns have been linked to the role of technology in shaping the demand for workers,[2] and to the elasticity of the labor supply, particularly of workers moving from the agricultural sector into manufacturing.[3] However, much of our knowledge about the impact of industrialization on labor incomes is inferred from aggregate data, which obscure regional heterogeneity that could shed light on how industrialization affects wages.

To gain a better understanding of the dynamic effects of industrialization on labor incomes, we focus on an industrial follower nation, the German state of Prussia from pre-industrial times to the height of the Second Industrial Revolution in Germany. According to Pfister (2018), German real wages followed a trajectory broadly in line with Britain, i.e., stagnant or even modestly declining until mid-century, gradually growing until 1880, and growing strongly thereafter. While our focus is to analyze the response of unskilled wages to expanding industrialization, we also examine whether industrialization affected wages of skilled workers. Finally, we explore *how* industrialization affected unskilled wages by investigating a range of mediating factors, including the roles of technological change, sectoral change, deskilling, and the centralization of production in shaping the wage response.[4]

One of the virtues of our study is that we exploit a novel panel dataset, transcribed from administrative sources, covering average day-laborer wages of male seasonal workers in 667 Prussian state forests, observed at decadal intervals from 1800 to 1879. This is unique as much other work is either much more limited in scope or has to infer incomes from occupations. We argue, and provide supporting evidence, that our wage data reflect local labor market conditions for unskilled workers, due to the high sectoral mobility of this group. To extend the analysis beyond 1879, we expand our dataset into a longer panel with newly collected county-level wage data from 1883 to 1914, allowing us to trace wage dynamics up to the First World War. These wage data are particularly suitable for our research question, as they span the period preceding widespread industrialization, the early phase of industrial development, and the Second Industrial Revolution in Prussia. To

[1] This notion relies on day wages, as used for example in Allen (2009b), Broadberry et al. (2015), Clark (2007), and Feinstein (1998). More recent evidence accounting for the length of the working year finds rising annual wages also during the early Industrial Revolution in England (Humphries & Weisdorf, 2019). In this paper, we abstain from interpreting day wages as informative about living standards to avoid having to make assumptions about the number of working days.

[2] See, e.g., Caprettini and Voth (2020), Mokyr et al. (2015), Nuvolari (2002), and Ridolfi et al. (2022) on the impact of labor-saving technologies, De Pleijt and Weisdorf (2017), Frey (2019), and Goldin and Katz (1998) on the impact of skill-saving technologies, and Franck and Galor (2022) and Galor (2011) on the emerging complementarity between technology and skill.

[3] This idea, sometimes associated with the 'industrial reserve army' in Marxian terms, is more frequently used in the context of labor scarcity and the resulting adoption of labor saving technologies in seminal work by Habakkuk (1962) and Acemoglu (2002) and more recently by Franck (2024) and Voth et al. (2023).

[4] Given the range of plausible channels through which industrialization may affect wages, we view our question as an empirical one that asks which forces dominate in this historical context.

adjust for differential changes in consumer prices, we construct a forestry-level price index based on wood prices from the same sources and expand it with fine-grained data on grain prices when analyzing the long panel.

For the analysis, the wage panel is combined with cross-sectional variation in industrialization potential, captured by proximity to carboniferous rock strata, a coal-bearing strata formed around 300 million years ago. This proxy for industrialization, widely used in the literature, exploits plausibly exogenous variation in access to hard coal deposits (De Pleijt et al., 2020; Esposito & Abramson, 2021; Fernihough & O'Rourke, 2021).[5] The transition from wood to coal was central to Prussia's industrial development, enabling the adoption of new technologies, the expansion of the industrial sector, and the rise of large-scale enterprises (Ayres & Warr, 2009; Pomeranz, 2000; Wrigley, 2013).

We use the resulting dataset to estimate the dynamic effect of coal-driven industrialization on unskilled wages in an event-study framework. Our results show no relationship between proximity to coal and wages until the mid-19th century, after which regions closer to coal deposits experience a significant and sustained increase in wages. This trend continues at least until 1914, as we demonstrate using the long panel, and holds despite declining transport costs in the second half of the 19th century. To interpret the results in absolute terms, we estimate a standard Difference-in-Difference (DiD) specification. We find that wages in regions located within 30 kilometers of the nearest coal deposit rise by approximately 11 percent between 1850 and 1879, relative to regions further away. Put differently, a one standard deviation (SD) increase in proximity to coal is associated with a 0.2 SD increase in unskilled wages by 1879, and a 0.36 SD increase by 1914.

By focusing on wages for forestry workers, we are able to use data for highly comparable tasks from a single source across a large number of regions and time periods, but we are limited to drawing conclusions about unskilled day laborers. To allow for broader conclusions, we supplement our analysis with more than 3,000 household accounts collected by H. K. Fischer (2011), which include detailed information on annual labor incomes and occupational skill levels between 1859 and 1914. Using a pooled cross-sectional regression approach, we confirm that wages of low-skilled workers are higher in closer proximity to coal, whereas wage gains for high-skilled workers and those in mechanical occupations were more limited. Coal-driven industrialization is thus found to be associated with a lower skill premium. We supplement these findings with employment data from the 1882 occupational census suggesting that the share of high-skilled workers in industrial employment is smaller in closer proximity to coal. These results suggest that the demand for low-skilled labor is elevated in industrial regions, supporting the idea that coal-driven industrialization in Prussia was unskill-biased.

To understand *how* industrialization affects unskilled wages, we conduct a mediation analysis using a two-period panel dataset. We examine several potential channels through which coal-based industrialization may have influenced wages, focusing on transformation on the firm side

[5] We confirm that proximity to carboniferous strata is strongly correlated with a range of other indicators of industrial activity in Prussia, suggesting that transportation costs and local linkages played a decisive role.

and changes on the worker side. On the firm side, we consider technological change, proxied by the adoption of steam engines; organizational change, proxied by average firm size; and market integration, proxied by market access. On the worker side, we examine sectoral change, measured by shifts in the manufacturing employment share, and deskilling, proxied by changes in the low-skilled employment share. In addition to these core mechanisms, we also consider health degradation, captured by infant mortality rates, and migration as potential mediators of wage dynamics.

Our results suggest that the effect of coal-driven industrialization on unskilled wages operates primarily through technological change and an increased demand for low-skilled labor. The finding that industrialization affects unskilled wages through technological change is consistent with the idea that productivity gains outweighed the displacement effects of technology adoption. Mediation through deskilling aligns with the workshop-to-factory hypothesis, which suggests that early industrialization was largely unskill-biased and increased the demand for low-skilled workers (Acemoglu, 2002; Atack et al., 2004; Goldin & Katz, 1998; O'Rourke et al., 2013). By contrast, we find little evidence that, conditional on technological change and deskilling, organizational change, market integration, or sectoral reallocation contributed to rising unskilled wages. Moreover, we do not find that the effect is driven by migration or health degradation.

Contribution to the literature. This paper contributes to debates on the distributional consequences of industrialization. A central focus of this debate is the observed stagnation or decline in labor incomes during the early stages of industrialization, followed by rising incomes and living standards in later phases (Broadberry et al., 2015; Clark, 2007; Voth, 2004).[6] Much of the existing work centers on timing and measurement issues for aggregate real wages in England and offers limited empirical insight into the underlying drivers of wage dynamics. A recent exception is Ridolfi et al. (2022) who show that the adoption of steam engine in France raised employment rates, male wages, and the labor share at the same time. In comparison, our paper provides a rigorous approach that exploits fine-grained regional variation in industrialization to examine its dynamic effects on wages and shows that patterns of the industrial follower Prussia resemble the English example, while also shedding light on the mechanisms underlying the distributional effects of industrialization.

Capital was a scarce factor during the early stages of the industrial revolution and its accumulation was necessary to raise output as part of a broader transition process (Galor, 2005). This relative scarcity of capital may have resulted in stagnating labor incomes as emphasized by Allen (2009b). The literature also documents an emerging complementarity between capital and skill as the industrial revolution progressed, which raised incomes of skilled workers as emphasized by Franck and Galor (2022), Galor (2011), and E. Lewis (2011). A related literature highlights that industrialization required certain highly-skilled mechanics to install, operate, and maintain the new

[6] The so-called standard-of-living debate contrasts pessimist accounts, which emphasize that price dynamics eroded real wages despite rising nominal incomes, with more optimistic interpretations, often focused on the period after 1850, which argue that industrialization spurred specialization, urban agglomeration, and productivity growth, ultimately raising labor incomes (Cinnirella, 2008; Ericsson & Molinder, 2020; Feinstein, 1998; Gallardo-Albarrán & De Jong, 2021; Humphries & Weisdorf, 2019; Hunt, 1986; Lindert & Williamson, 1983; Nicholas & Steckel, 1991).

machines (Kelly et al., 2023; Mokyr et al., 2022). Considering these different aspects of the literature, predictions about the dynamics of labor incomes as well as about the dynamics of skilled and unskilled wages are unclear. This paper aims to shed some new light on the debate by showing that coal-based industrialization dynamically affected the wages of unskilled workers whereas wage gains for high-skilled workers and mechanically-trained occupations were more limited, suggesting a lower skill premium in close proximity to coal deposits.

One set of arguments explaining wage dynamics during the industrial revolution highlights the role of labor demand, specifically, whether technological progress was directed at saving or augmenting labor (Acemoglu, 2010; Allen, 2009a; Hicks, 1932; Zeira, 1998). Labor-saving technological motives and their potential displacement effects feature prominently in Acemoglu and Restrepo (2020a), Allen (2021), Hornbeck and Naidu (2014), E. Lewis (2011), and Ridolfi et al. (2022). However, the so-called 'deskilling hypothesis' suggests that technical progress during the early stages of the industrial revolution actually generated a higher demand for low-skilled labor and should thus increase wages at least among low-skilled workers (Allen, 2015; Atack et al., 2004, 2019; De Pleijt et al., 2020; De Pleijt & Weisdorf, 2017; Goldin & Katz, 1998; O'Rourke et al., 2013). We contribute to this literature by documenting first that industrialization had a positive effect on unskilled wages through the adoption of the steam engine and second that this effect extends into the later stages of the industrial revolution.

Another set of arguments highlights the role of labor supply. Early industrialization is portrayed as a period of labor abundance during which a large reserve of agricultural labor limited wage growth in manufacturing (Acemoglu, 2002; Habakkuk, 1962; W. A. Lewis, 1954). In contrast, recent work emphasizes the role of labor scarcity in spurring technological progress (Andersson et al., 2022; Franck, 2024; San, 2023; Voth et al., 2023). While we have little to say on the latter topic, our findings suggest that the abundance of low-skilled workers in Prussia coincided with stronger wage growth of unskilled workers, consistent with rising industrial demand after 1850.

We also contribute to the literature that examines whether natural resources are a blessing or a curse (van der Ploeg, 2011). Several studies have investigated recent booms and busts in oil and gas fracking and gold mining (Allcott & Keniston, 2018; Aragón & Rud, 2013; Bartik et al., 2019; Caselli & Michaels, 2013; Jacobsen & Parker, 2014; Michaels, 2011). This literature frequently finds that resource booms raise employment and aggregate incomes in the short run, but that long-run effects are more heterogeneous. In particular, local extraction and reliance on coal often entails adverse consequences in the long run (Berbée et al., 2024; Black et al., 2005; Esposito & Abramson, 2021; Franck & Galor, 2021; Fritzsche & Wolf, 2023). While this literature typically focuses on aggregate incomes, the few studies that examine wages find that positive effects during boom phases tend to vanish after the bust (Jacobsen et al., 2023). Our paper contributes to this debate by providing a historical perspective and by adding nuance regarding differential effects on high- and low-skilled workers. In contrast to modern resource booms, coal in 19th-century Prussia fueled broader industrial development, thereby reshaping local labor markets and the direction of technological change.

The remainder of this paper is organized as follows. Section 2 provides some background industrialization and labor markets in Prussia; Section 3 describes our main dataset; Section 4 presents our main results with respect to the dynamics of unskilled wages; Section 5 presents some evidence on skilled wages and the skill premium; Section 6 provides evidence for potential channels through which industrialization affected wages; Section 7 concludes.

2 Historical background in 19th-century Prussia

In this section, we describe the foundations and beginnings of the industrial take-off in Prussia, emphasizing the central role of coal mining in Prussia's industrialization. Furthermore, we examine the development of the labor market in Prussia in the 19th century.

2.1 Institutional pre-conditions for industrialization

During the 19th century, Prussia rapidly transformed from an agrarian society into one of Europe's leading industrial economies. Institutional reforms enacted in the early decades of the century laid the groundwork for this transition (Henning, 1973; Tilly & Kopsidis, 2020). The Stein-Hardenberg reforms (1807–1821) abolished feudalism and curtailed the power of guilds, thereby promoting occupational freedom, labor mobility, and facilitating factor markets (Ashraf et al., 2025; Tilly & Kopsidis, 2020).

Following the end of the Napoleonic Wars, Prussia gained resource-rich western territories that later became its key coal mining regions. The abolition of internal tariffs in 1818 and the establishment of the Zollverein in 1834 enhanced market integration. Infrastructure improvements, particularly the rapid expansion of the railway network from 1,600 km in the mid-1840s to over 20,000 km by 1879, further reduced transportation costs and facilitated both goods and labor mobility.

The German Revolution of 1848/49, though politically unsuccessful, triggered concessions that resulted in a more liberal constitution and institutional stabilization. This period also marked the liberalization of coal mining. The Co-ownership Act (*Miteigentümergesetz*) of 1851 transferred mine management from the state to private actors, incentivizing investment and technological advancement (Schulz, 1911; Tilly & Kopsidis, 2020). These institutional developments laid the structural foundations for Prussia's broad-based industrialization which accelerated in the mid-19th century (Hoffmann, 1963; Tilly & Kopsidis, 2020).

2.2 Labor market developments

The same institutional reforms that liberalized markets also reshaped Prussian labor relations. The abolition of serfdom and the introduction of occupational freedom dismantled traditional labor constraints and lifted mobility restrictions. However, they also forced many smallholders and landless rural workers into wage labor, particularly as redemption payments and the enclosure of common lands reduced disposable incomes. As a result, labor mobility and the supply of wage workers increased across sectors (Gutberlet, 2014; Pierenkemper, 2009).

This growing labor supply facilitated the structural transformation of the Prussian economy throughout the 19th century. Concurrently, the industrial workforce grew sharply: the population share employed in manufacturing rose from around 8% in 1849 to 12% by 1882, reflecting the rapid expansion of factory-based production.[7] Additionally, new technologies played a pivotal role in mechanizing manual labor, especially in the textile sector (Gutberlet, 2014).

Despite these developments, labor market regulation remained limited. Labor market conditions were characterized by freely negotiated employment contracts. Regulatory protections for employees were nearly absent, and collective wage agreements remained uncommon until the turn of the century. Worker associations faced significant legal obstacles and most were banned until 1869 (Pierenkemper, 2009).[8] Consequently, we argue that wages in 19th-century Prussia were predominantly determined by market forces, reflecting shifts in labor supply and demand rather than institutional or regulatory interventions.

2.3 Transition towards coal

The transition to fossil fuels was a central driver of Prussia's industrialization. Although coal had been used for domestic heating since the 18th century, it remained of limited industrial importance until the mid-19th century. Prior to that, water, animal, and charcoal power dominated, each strongly constrained by geography and seasonal fluctuations. Several factors delayed the rise of coal, including immature smelting technologies, state control of mining, and the limited diffusion of steam engines (Kopsidis & Bromley, 2017).[9]

From mid 19th century onward, coal became an increasingly important source of power. Between 1850 and 1880, hard coal production increased more than ninefold and lignite production increased more than sixfold (Figure A1). Steam engines enabled deeper mining and improved energy conversion, while coal-fired transport and industrial machinery facilitated larger factories and faster production (Kander & Stern, 2014; Malm, 2016). The shift to coal also relaxed locational constraints based on water availability and facilitated production with increasing returns (Ayres & Warr, 2009; Pomeranz, 2000; Wrigley, 2013). At the same time, coal remained a geographically concentrated resource. High transport costs and infrastructure limitations initially constrained its use.[10]

[7] In 1843, Germany produced only 10% of its rail material domestically and imported over 88% from Great Britain; by 1863, this scenario reversed, with domestic production accounting for 85.4% and imports falling below 13% (Fremdling, 1979, p. 211).

[8] The first collective wage agreements appeared relatively late, beginning with construction workers in 1899, followed by metalworkers in 1906 (Schneider, 1989, p. 108-110). Forestry workers in state forests gained collective agreements in 1919 (Treitschke, 1928, p. 215 ff.).

[9] The number of steam engines operating in Prussia increased from 231 (total horsepower 3,670) in 1830 to 1,445 (29,482 HP) in 1849 and to 35,431 (958,366 HP) in 1878. In 1846/47 (1875), 43.8 (33.4)% of all steam engines were installed in mining, 18 (30.5)% in metal, 14.9 (10.3)% in textile, and 4.3 (3.9)% in the machinery industry (Banken, 1993). Steam power was first used to reach deeper, higher-quality coal in 1839–40 (Tilly & Kopsidis, 2020).

[10] Transport costs for coal fell considerably in the last decades of the 19th century. For example, the cost of

Coal mining also reshaped local labor markets. First, mining itself was labor intensive. Second, mining required a wide range of intermediate inputs, the production of which benefited from proximity to the mines. Third, high transport costs for coal encouraged firms using it as an input to locate nearby. Finally, spatial proximity facilitated knowledge spillovers and innovation, especially in mining technology (Morris et al., 2012).[11] As we will show below, these forces favored industrial development in coal-rich regions and contributed to an increasing demand for labor.

3 Data

This section provides an overview of the main data used in this study. The data is either observed at the firm, i.e. forestry level, or at the county level. Additional information on the sources and the definition of all variables can be found in Appendix A.3. Descriptive statistics are reported in tables A1 to A4.

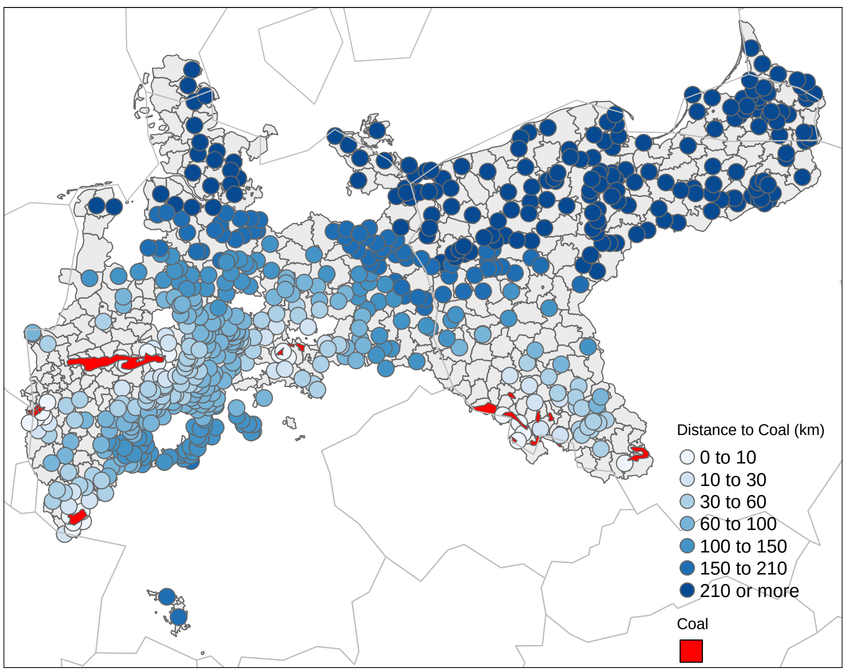


Figure 1: Distribution of Carboniferous strata and location of Prussian state forests

Notes: The figure shows the location of state forests (blue circles) and Carboniferous rock strata (red polygons).

transporting a tonne of Ruhr coal by rail per kilometer fell from 15 *Pfennig* in 1836 to 0.5 *Pfennig* in 1877. At the same time, the share of Ruhr coal transported by rail rose from 30% in 1853 to 77% in 1871 (Tilly & Kopsidis, 2020, p. 133). Electrification only began in the 1880s and was not widespread until the 1920s (Herzig & Ott, 1986).

[11] Underground coal mining required specialized infrastructure such as tunnels, shafts, elevators, railways, ventilation systems, and pumps to remove groundwater.

3.1 Labor income

Regionally disaggregated wage data are virtually non-existent for any of the German states for much of the 19th century. Only from 1883 onward were local day-laborer wages systematically collected to administer sick pay through the public health insurance. For this study, we uncovered and transcribed unique firm-level panel data on male day-laborer wages (seasonal workers) employed across all 667 state forests in Prussia. These data are reported as averages over nine decades between 1800 and 1879 in U. Eggert (1883).[12] Figure A2 in the Appendix shows the development of nominal unskilled wages over time, suggesting a strong upward trend throughout the 19th century.

To our knowledge, no other dataset documents wages so consistently for this period and with such extensive spatial coverage. Figure 1 shows the locations of all 667 state forests, which are distributed across 264 different counties (using 1871 administrative boundaries).

The panel is unbalanced, with relatively sparse coverage in the early decades of the 1800s and 1810s, prior to Prussia's territorial expansion. We show in Tables A5–A6 in the Appendix that missing wage data are less prevalent in forestries in closer proximity to coal deposits in each decade. Hence, it is unlikely that the pattern of results observed in our event study analysis below is produced by an increasing amount of coal regions with high wages entering the dataset. It is not the case that forestries located closer to coal deposits enter the dataset only in the later decades.[13]

Representativeness. We argue that wages for woodworkers in state forests are representative of local wages for unskilled male workers and provide both qualitative and quantitative evidence in support of this claim. First, as described in Section 2.2, labor markets in 19th-century Prussia were unregulated, and wages, especially for day laborers, were determined by market forces. Contemporaries noted that day laborers were highly mobile across sectors, frequently moving between factory work, mining, farm work, and forestry in response to wage differentials (Neuhaus, 1904).[14] Furthermore, we assume that the reported wages reflect total remuneration, as in-kind payments, such as firewood, were prohibited by the 1817 Ordinance for Woodworkers in the Royal Prussian Forests (*Hauordnung für die Holzhauer in den Königlich Preußischen Forsten*) (Treitschke, 1928, pp. 28–30).[15]

To substantiate these qualitative arguments, we compare forestry wages to those of ordinary day

[12] The data were collected retrospectively in 1879–1880, based on firm-level wage documentation, by decree of the Minister of Agriculture, Domain, and Forestry. They record average wages paid during the spring growing season, the most labor-intensive period in forestry (Treitschke, 1928). Typical work included cultivation of the soil, cutting grass and brush, and thinning surplus sprouts (Sparhawk, 1938). The chief forester (*Oberförster*) of each forestry was responsible for hiring workers and setting wages (Treitschke, 1928, pp. 31–33).

[13] Below, we show that our results are robust to restricting the sample to a balanced panel.

[14] The wage data source quotes the Minister of Agriculture, Domains, and Forestry: *"The Forest Administration does not pay more than it has to pay. [...] If we were to pay lower wages, we would not get any workers, according to experience. For the rest, strict attention is paid that the individual forestries do not pay higher wages than is urgently necessary."* (Treitschke, 1928, p. 59), own translation.

[15] A potential concern is that day laborers near coalfields may have worked longer hours as a result of coal-led industrialization beginning in the mid-19th century, which could explain wage increases in these areas. However, we have no reason to believe that working hours in forestries differentially increased with industrialization. Forestries hired day laborers flexibly for seasonal work, and anecdotal evidence suggests that they typically worked from sunrise to sunset, subject to weather conditions (Treitschke, 1928, p. 31).

laborers once such data become available in 1883. For this purpose, we digitized and transcribed county-level data on the wages of ordinary day laborers as reported in Schmitz (1888) and Zentralblatt (1892–1914) for the years 1883, 1897, 1905, 1910, and 1914. These data complement data on day laborer wages for 1892 and 1901 originally made available by S. O. Becker et al. (2014).[16] These additional wage data also allow us to extend our analysis until the outbreak of World War I.

Wages of ordinary day laborers are reported largely at the level of Prussian counties, but can be subdivided between urban and rural locations. To compare and combine both wage datasets, we aggregate the forestry-level data to the county level. We achieve broad coverage by using the simple average of wages across all forestries located within a 30 km radius around each county centroid.[17] This approach yields wage estimates for 275 out of 335 counties in 1849, for which consumer price data are also available.

The results of this comparison, presented in Figure A4 in the Appendix, suggest that wages paid in state forests in 1875–79 are highly correlated with those of ordinary day laborers in 1883. This relationship holds for both urban and rural areas within the same county, as well as for population-weighted averages. The correlation coefficients and the slope coefficient are close to one, indicating a high degree of alignment between wages in state forests and those in the broader labor market. In Table A7 in the Appendix, we further show that correlations between proximity to coal and wages from different data sources yield quantitatively similar coefficients. Based on this evidence, we conclude that day laborer wages in state forests are reflective of local market wages for unskilled male workers throughout 19th-century Prussia.

3.2 Consumer prices

To ensure that increasing wages in industrializing regions are not a mere result of increasing consumer prices, we deflate nominal wages using a local consumer price index. Our preferred index is based on wood prices reported in the same source as the wage data. Specifically, we digitized and transcribed prices for five varieties of wood, including two types of firewood (conifer and beech) and three types of timber (oak, pine, and spruce), as reported in U. Eggert (1883). The main advantage of this index is that it reflects market prices at the same spatial unit of observation as the wage data, namely for all 667 state forests, in decadal intervals from 1800–1809 to 1870–1879.[18] To construct the index, we first standardize prices by variety to have zero mean and unit SD, and then average them at the firm-decade level. Note that by controlling for wood prices, we also account for the possibility that wages may be influenced by firm profitability, for instance through profit sharing (Fuest et al., 2018).

[16] Data on the income of ordinary day laborers was collected under the provision of § 8 of the Health Insurance Act of June 15, 1883, which stipulated that the usual daily wage of ordinary day laborers served as the key factor in determining health insurance contributions and sickness benefits paid by health insurance funds (Neuhaus, 1904).

[17] Thirty kilometers is approximately the average walking distance a person could cover in a day in the 19th century. We show that results are robust to using other cut-offs in the Appendix.

[18] Firewood prices refer to one cubic meter of decomposed wood; timber prices refer to one cubic meter of pure wood mass. In a few cases, tax-assessed prices were entered when average selling prices were not available (U. Eggert, 1883).

We argue that wood prices are representative of local consumer prices. To substantiate this claim, we compare our wood price index to an alternative consumer price index based on eight staple goods: wheat, rye, barley, oats, rapeseed, potatoes, straw, and hay. The underlying market price data are available for the period 1837–1860. We compare the resulting consumer price index to our wood price index, averaged over the period 1830–1869, in Figure A5 in the Appendix. The correlation coefficient of 0.63 and the slope coefficient of 0.77 suggest that wood prices closely track overall consumer prices, supporting their use as a proxy in our main analysis.

Figure A3 in the Appendix plots nominal wages adjusted for local prices over time. The figure suggests real wages were largely stagnant until the 1860s, followed by modest increases thereafter.

When extending the wage panel to include data on ordinary day laborers between 1883 and 1914, where wood prices are not available, we deflate nominal wages using local wheat prices. For the period 1800 to 1879, we rely on annual wheat prices from 54 cities, as reported in Federico et al. (2021). For the period 1883 to 1914, we rely on newly digitized monthly wheat prices from 165 cities, as reported in Königlich Preussisches Statistisches Bureau (1861–1914). Each county-level wage observation is deflated using the average wheat price for the same period as the wage taken from the nearest city.[19]

3.3 Proximity to coal deposits

We use geographical variation in rock strata formed during the Carboniferous period as our preferred proxy for industrialization. We do so following the literature (De Pleijt et al., 2020; Esposito & Abramson, 2021; Fernihough & O'Rourke, 2021) which has shown that this coal-bearing strata is highly predictive of industrialization in Europe. Specifically, we focus on the late Carboniferous or Pennsylvanian period (approximately 323 to 300 million years ago) during which almost all hard coal deposits that were later commercially exploited in the region of study were formed. By using this measure rather than the historical location of coal mines, we avoid endogeneity arising from the possibility that mines may have been placed where labor was relatively cheap. This will allow us to estimate the dynamic reduced-form effect of coal abundance on wages in the absence of panel data on industrialization across 19th century Prussia.[20]

The spatial distribution of Carboniferous rock strata is derived from geospatial information provided by Asch and Bellenberg (2005). The distribution of these geological formations across Prussia is shown in Figure 1. For the empirical analysis, we compute the negative of the distance from each forest administration to the nearest area with Carboniferous rock layers. The transformation ensures that larger values indicate greater proximity and thus better access to coal.

Coal deposits and industrialization in Prussia. Coal consumption was a key driver of Prussia's industrialization, as discussed in Section 2.3. To validate that Carboniferous rock

[19]As wheat price data are missing for many markets in 1914, wages reported for that year are deflated using 1910 prices.

[20]Figure A6 in the Appendix confirms that distances between state forests and coalfields and distances between state forests and Carboniferous rock layers are highly correlated. Table A11 shows that results are qualitatively similar when using other definitions of coal availability.

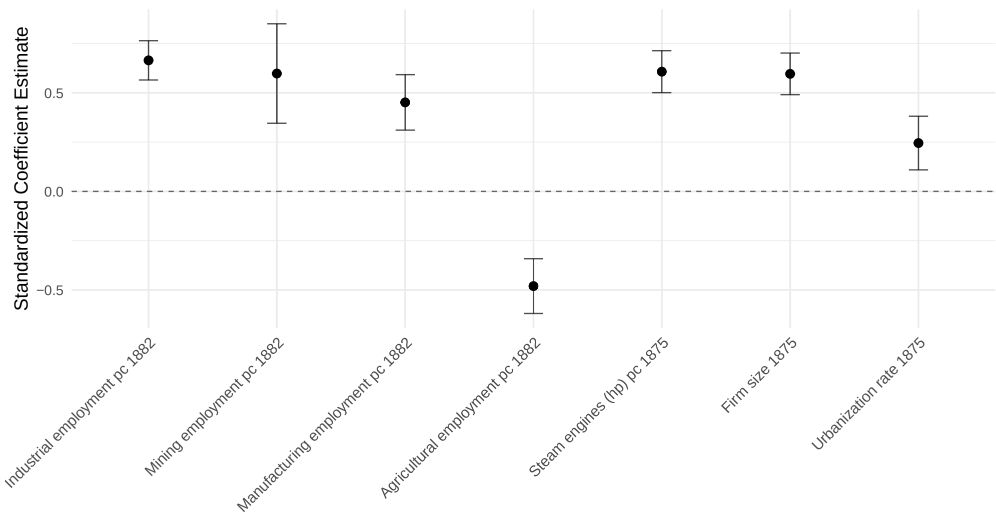


Figure 2: Coal deposits and industrialization in 1875 and 1882

Notes: The figure plots coefficients and 95% confidence intervals from cross-sectional OLS regressions of the dependent variables indicated below each plot on the negative (log) distance to coal deposits from the Carboniferous era in km. The unit of observation is a county ($N = 452$). All variables are logged and standardized to have a zero mean and unit SD. Results are conditional on a full set of geographic controls (slope, caloric suitability, distance to rivers, and distance to coast). Standard errors adjusted for heteroskedasticity. Corresponding estimates for each plot are presented in Table A8 in the Appendix.

strata serves as a reliable proxy for regional industrialization, we examine its correlation with various indicators of industrial activity in 1875 and 1882.[21] Figure 2 shows that closer proximity to coal deposits is associated with higher employment in industry (plot 1), mining (plot 2), and manufacturing (plot 3), and with lower employment in agriculture (plot 4). A one SD increase in proximity to coal is associated with a 0.67 SD increase in the industrial employment share and a 0.48 SD decrease in the agricultural share, suggesting that coal regions drew labor from agriculture into both mining and manufacturing.

The figure also supports several claims made in Section 2.3. Proximity to coal is associated with greater horsepower per capita in steam engines (plot 5), consistent with the role of transportation costs for coal in shaping technological adoption and with production linkage effects. Firms are larger (plot 6), suggesting that coal-led industrialization is associated with economies of scale. Urbanization rates are higher (plot 7), confirming earlier findings by Fernihough and O'Rourke (2021). We conclude that geographical variation in Carboniferous strata shaped the spatial distribution of industrial activity in Prussia, validating their use as a proxy for industrialization.

[21] To facilitate interpretation, we first log-transform and then standardize both the outcome and treatment variables. See Table A8 in the Appendix for regression estimates.

3.4 Geographic controls

We control for several potentially confounding geographic factors that may be correlated with the geographic concentration of carboniferous rock strata and that may have influenced the wages of unskilled labor in 19th-century Prussia. First, we control for variables that may have affected market access and thus wages and may be correlated with proximity to coal, including distance to the nearest river and distance to the nearest coast. Second, we include a caloric suitability index (Galor & Özak, 2016). This index reflects the potential agricultural production in 5'x5' grid cells (measured in calories), taking into account the crops available for cultivation. Third, we include the slope of the terrain to account for the fact that coal is most prevalent in mountainous regions that may generally differ in their economic activity. This may also account for the prevalence of water mills which was the main alternative technology to steam engines (Gutberlet, 2014) and may have determined the level of proto-industrial physical capital and thus wages (Ashraf et al., 2025).[22]

4 Industrialization and unskilled wages

In this section, we empirically investigate the effects of coal-driven industrialization on the wages of unskilled workers.

4.1 Empirical framework

To examine the relationship of interest, we start by estimating the following simple DiD model:

$$Wage_{it} = \alpha_i + \gamma_t + \beta(Coal_i \times I_t^{1850}) + \mu P_{it} + \delta(X_i' \times I_t^{1850}) + \epsilon_{it}, \tag{1}$$

where the dependent variable $Wage_{it}$ denotes the wage of unskilled seasonal workers in forestry i denoted in *Mark* during decade t (i.e., 1800–1809, 1810–1819, ..., 1875–1879). α_i are forestry-fixed effects that control for time-invariant characteristics. γ_t are time-fixed effects that account for common shocks affecting all forestries in a given decade t. The main explanatory variable $Coal_i$ is the negative of the distance from each forestry i to the nearest area with Carboniferous rock strata in (log+1) kilometers, interacted with an indicator variable I_t^{1850}, which equals one for all periods from 1850 onward. We choose 1850 since, as discussed in Section 2, the literature widely agrees that broad-based industrialization in Prussia began during this period (we present quantitative evidence for this below). The estimated β coefficient can be interpreted as the difference between coal and non-coal regions after 1850 compared to the difference before 1850. To facilitate interpretation, we sometimes define our explanatory variable $Coal_i$ as an indicator that takes the value one if a forestry is located within 30 kilometers around the forest administration and zero otherwise.[23] Our preferred specification includes an index P_{it} to account for regional differences in consumer prices

[22]The most important pre-industrial energy source in Prussia was water power. Water mills could only be built on rivers with a certain gradient. Regions suitable for the use of water mills may have acquired more proto-industrial capital and thus advantages in adapting new technologies.

[23]We show robustness to altering this cut-off in Table A12 in the Appendix.

that may develop heterogeneously over time and may be correlated with proximity to coal. To rule out confounding factors that may correlate with proximity to coal and may have influenced wages differently over the 19th century, we add a set of time-invariant geographical control variables interacted with the post-1850 indicator.

To study the dynamic effect of proximity to coal deposits on unskilled wages, we move on to adjusting Equation 1 to become the following event-study specification:

$$Wage_{it} = \alpha_i + \gamma_t + \sum_{t=1800}^{1879} \beta_t (Coal_i \times I_t) + \mu P_{it} + \sum_{t=1800}^{1879} \delta_t (X_i' \times I_t) + \epsilon_{it}, \tag{2}$$

where the main explanatory variable $Coal_i$ and all time-invariant geographical control variables are interacted with a full set of time-fixed effects I_t. All other variables are defined as in Equation 1. The estimated coefficients of interest, β_t, capture the effect of proximity to coal on wages in period t, relative to a baseline period. We choose the decade 1840–1849 as the baseline, as our data suggest that the relationship between proximity to coal and wages begins to change after the 1840s.[24] This quantitative assessment is based on two sets of exploratory regressions. The first set estimates Equation 2 without exploiting the panel structure of the data, that is, by estimating the model separately for each time period. The results, presented in Table 1, show how proximity to coal is associated with unskilled wages in each decade. We find no statistically significant relationship between coal proximity and wages in the first five decades, including the 1840s. Starting in the 1850s, however, forestries located closer to coal deposits exhibit significantly higher wages, with the size of the estimated coefficient increasing through the 1870s. By that time an increase in proximity to coal by one SD was associated with increase of wages by 7 *Pfennig*, equivalent to 18% of a SD.

The second approach exploits the panel structure of the data by estimating a series of simple DiD regressions using rolling sampling windows. Each regression restricts the sample to a 40-year window and includes an indicator variable equal to one for the later two decades, interacted with proximity to coal deposits. We begin with the window 1800–1839, then proceed to 1810–1849, and continue in this fashion up to 1840–1879. The results, presented in Table A9 in the Appendix, indicate that proximity to coal had no significant positive effect on wages in the earlier windows. Only when the comparison includes decades after the 1840s, specifically when comparing the two decades before and after 1850 or 1860, we observe significantly higher wages in regions closer to coal. We interpret this as evidence that coal-driven industrialization began to raise wages only from the 1850s onward. This finding further justifies our use of the 1840s as the reference period in the event-study analysis.

4.2 Main results

The main results estimated using the simple DiD framework are presented in Table 2. Column 1 uses the continuous version of the treatment variable, suggesting that an increase in proximity

[24] Choosing an earlier baseline decade yields qualitatively similar results, as pre-trends in wages are essentially flat until the 1840s.

Table 1: Industrialization and the return to unskilled labor (by decade)

Dependent variable:	Unskilled wages								
	1800-09	1810-19	1820-29	1830-39	1840-49	1850-59	1860-69	1870-74	1875-79
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coal (proximity)	0.004	0.022	0.006	0.018	0.018	0.042***	0.046***	0.062***	0.068***
	(0.016)	(0.015)	(0.012)	(0.012)	(0.012)	(0.014)	(0.017)	(0.020)	(0.016)
Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price Index	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	93	155	245	315	409	469	571	635	649
Adjusted R^2	0.10	0.07	0.08	0.05	0.13	0.15	0.27	0.31	0.40

Notes: The table shows results from estimating Equation 2 for each decade individually. The unit of observation is a firm (forestry). The dependent variable measures average wages of male seasonal forestry workers in *Mark* during a decade. Coal (proximity) measures the negative distance to coal deposits from the Carboniferous era in (log) kilometers. This variable is standardized with zero mean and unit SD. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Robust standard errors in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table 2: Industrialization and the return to unskilled labor (simple DiD)

Dependent variable:	Unskilled wages			
	Coal continuous	Coal dummy	County level	Extended sample
	(1)	(2)	(3)	(4)
Coal (proximity) × I^{1850}	0.056***		0.048***	0.076***
	(0.010)		(0.010)	(0.016)
Coal (30km dummy) × I^{1850}		0.114***		
		(0.029)		
Geo Controls	Yes	Yes	Yes	Yes
Price Index	Yes	Yes	Yes	Yes
Unit fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
Observations	3,541	3,541	2,928	6,098
Unit	Forestry	Forestry	County	County
Adjusted R^2	0.88	0.88	0.91	0.92

Notes: The table shows results from estimating Equation 1. The unit of observation is a forestry (columns 1–2) or a county (columns 3–4), observed in decadal intervals from the 1800s to the 1870s, extended with individual years until 1914 in column 4. The dependent variable measures average wages of male seasonal forest workers in *Mark* during a decade. The main explanatory variable in columns 1, 3, and 4 measure the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers. The main explanatory variable in column 2 is a dummy that is one if the unit is within 30 km of a coal deposit from the Carboniferous era. The explanatory variable is interacted with a dummy that is one from the 1850s. Geo: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the forestry level (1–2) or county level (3–4) in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

to coal by one SD leads to an increase in unskilled wages by 15 percent of a SD after 1850. The dichotomous version in column 2 suggests that wages in regions located within a 30-kilometer radius of the nearest coal deposit increased by about 11 *Pfennig* after 1849 compared to regions without nearby coal deposits. This implies an increase by 11 percent relative to the mean in the period. The results are qualitatively similar when aggregating to the county level in column 3. This aggregation allows us to extend the panel to include county-level data on the wages of ordinary day laborers for the years 1883, 1892, 1897, 1901, 1905, 1910, and 1914 in column 4. This extension confirms the main results and hints at larger increases in later periods.

We proceed by estimating Equation 2 as an event study, exploiting the full panel structure of the data. Results from this specification, presented in Figure 3a, offer two key insights. First, we observe that wage trends were approximately parallel prior to the 1850s. We can thus interpret coal-driven industrialization as a labor market shock that changed wage trends in regions near coal deposits after the 1840s but not in more distant areas. Second, we observe that once coal became an important input to industrialization its effect on wages continues to increase over time. If coal-driven industrialization represents a transitory shock, we would expect wage differences to diminish over time. However, the process of industrialization became increasingly coal-intensive (Gutberlet, 2014), thereby exerting a compounding effect on wages in the absence of strong forces equalizing wage differentials.

To further investigate the transitory nature of the industrialization shock, we again aggregate to the county level and extend our panel dataset with county-level data on the wages of ordinary day laborers from 1883 to 1914. Note that the absence of data on wood prices for this extension forces us to adopt a location-specific deflator that relies on local wheat prices. Results when estimating Equation 2 at the county level from 1800 to 1914 are presented in Figure 3b. We find that the wage differential between industrializing regions and their counterparts slightly declines in the 1890s but then continues to increase until the end of the sample period.[25] We note that the overall trend continues despite the concurrent reduction in coal transportation costs through railroad-infrastructure development and increasing labor mobility. This lack of convergence may be attributable to *Marshallian* agglomeration effects, as well as the steadily growing demand for labor from the local coal mining sector and its co-located up- and downstream industries (R. Eggert, 2001; Enrico, 2011).[26]

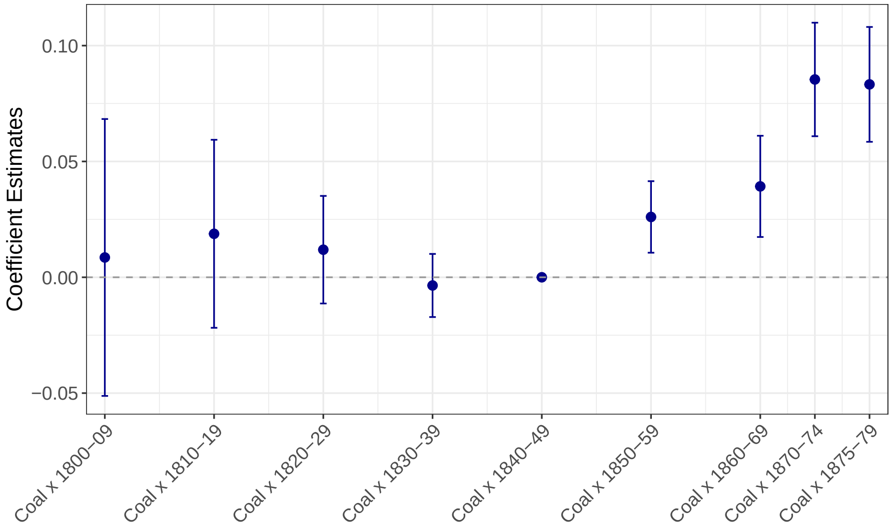
4.3 Robustness and endogeneity concerns

Our results so far show that unskilled wages in regions close to coal deposits increased once industrialization in Prussia accelerated. In the following, we conduct a series of tests to assess the robustness of these findings (Tables A10 to A15 in the Appendix).[27]

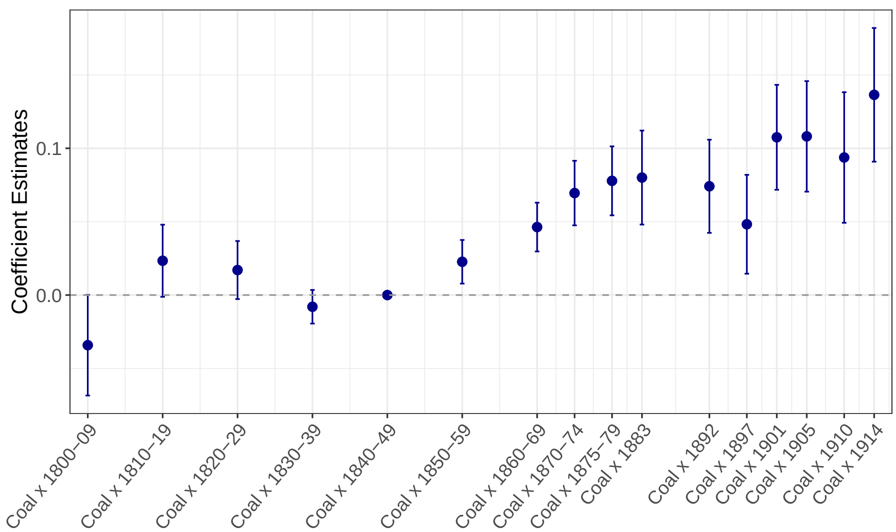
[25] The findings are confirmed when using either only urban or only rural wage data for the period 1883–1914, as shown in Figure A7 in the Appendix.

[26] Marshallian agglomeration effects arise when firms are located near one another, enabling the exchange of ideas and knowledge (knowledge spillovers), improving labor market efficiency (labor pooling), and reducing costs through shared suppliers or transport networks (input sharing) (Marshall, 1920).

[27] The first column of each table reproduces the baseline estimates from column 3 of Table A10 for reference.



(a) Firm-level results (1800-1879)



(b) County-level results (1800-1914)

Figure 3: Industrialization and the return to unskilled labor (event study)

Notes: Figures plot β_t coefficients estimated from Equation 2 with 95% confidence intervals. The omitted period is 1840–1849. The dependent variable measures average wages of male seasonal forestry workers in *Mark* during a decade until 1879. In Figure 3b these are extended using male day-laborer wages for individual years between 1883 and 1914. The main explanatory variables measure the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with time dummies. Results are conditional on a full set of geographic controls (slope, caloric suitability, distance to rivers, and distance to coast). The specification in Figure 3a uses forestries as the level of observation, controls for a wood-price index, and clusters standard errors are at the forestry level. Figure 3b uses counties as the level of observation, controls for wheat prices of the closest city in the same period, and clusters standard errors at the county level. Corresponding estimates for each plot are presented in columns 3 and 9 of Table A10 in the Appendix.

Specification robustness. Table A10 in the Appendix presents the first set of robustness checks on the event study results. We begin with more parsimonious specifications, where column 1 includes only fixed effects, column 2 adds the full set of geographic controls, and column 3 repeats the baseline results including the wood-price index. The estimated β_t coefficients remain largely unchanged when the price index is included, indicating that differential trends in consumer prices do not explain the observed wage increases. This further supports the interpretation that daily wages reflect broader labor market forces rather than just local variation in wood prices or the productivity of state forests. Column 4 uses a logarithmic version of the dependent variable to adjust for potential skewness. Column 5 aggregates the forestry-level data to the county level. Column 6 adjusts standard errors for spatial correlation following Conley (1999), whereas column 7 adopts the method to correct for spatial unit roots suggested by S. Becker et al. (2025). Column 8 uses the period 1800–09 as the omitted reference period. Column 9 restricts the analysis to a balanced panel. However, due to data scarcity in the first two decades, we drop these decades entirely and balance the panel with respect to the 1820s. The corresponding coefficient plot is presented in Figure A8 in the Appendix. None of these modifications change the pattern of results qualitatively. Column 10 extends the sample until 1914, as shown in Figure 3b. Some pre-1850 interaction terms are statistically significant, but their pattern points to declining wages in coal regions, reducing concerns about deviations from parallel trends.

Alternative measures of coal proximity. Tables A11–A13 in the Appendix test the robustness of our results to alternative measures of exposure to coal deposits. The corresponding maps are shown in Figure A9. In Table A11, column 2 measures proximity using the location of all coal mines in Prussia based on data from Fernihough and O'Rourke (2021). Column 3 refines these data by restricting the sample to hard-coal mines identified through our own research. Column 4 extends the measure to include Carboniferous rock strata located outside Prussia. Column 5 broadens the definition of the Carboniferous period to 354–292 million years ago, while column 6 defines coal regions not by rock age but by the dominant rock type.[28] Across all specifications, the estimated β_t coefficients remain stable in magnitude and significance.

Table A12 further tests sensitivity to alternative distance thresholds in the dichotomous version of $Coal_i \times I_t^{1850}$ from Equation 1. We vary the definition of proximity in intervals from 10 to 50 km. The estimated β coefficients decline gradually as the threshold increases, consistent with the notion that transport costs and agglomeration effects shape the spatial distribution of industrialization.

Finally, Table A13 examines potential spatial spillovers of coal proximity following the approach of Butts (2023). We estimate effects across distance bins, introducing indicators for forestries located within 0–30 km (our baseline), 30–60 km, and 60–90 km distance of coal deposits. The results indicate that the wage effects of coal proximity are confined to the first 0–30 km bin, suggesting that spillovers decay sharply with distance. Importantly, we find no evidence of negative wage effects in adjacent areas, implying that wage gains in coal-proximate regions did not come at the expense of neighboring counties.

Additional control variables and potential confounders. Table A14 in the Appendix

[28] In this specification, a region is classified as a hard-coal region if hard coal is one of the two most prevalent rock types (Asch & Bellenberg, 2005).

introduces additional controls to account for time-invariant regional characteristics that may have affected wage dynamics. Column 2 adds an indicator for regions west of the Elbe, which historically marked a divide in agricultural and labor institutions.[29] Column 3 adds an indicator for regions occupied by France during the Napoleonic period, capturing the legacy of liberal institutions implemented until 1815 (Acemoglu et al., 2011). Column 4 adds an indicator for forestries located in territories annexed by Prussia after 1866, where state forests, labor markets, and coal mines may have been subject to different institutional legacies before their incorporation.[30]

Columns 5 to 9 add geographic factors that may have shaped the ability of regions to adopt new technologies such as the steam engine, and account for labor mobility. Column 5 controls for the geographic distance to London to capture the diffusion of industrial technologies from Britain. Column 6 adds distance to Wittenberg to account for potential differences in human capital formation associated with the Protestant Reformation (S. O. Becker & Woessmann, 2009). Column 7 includes distance to the Polish border to ensure that wage differentials do not arise from increasing cross-border labor mobility. Column 8 controls for distance to the nearest county border to account for potential migration within Prussia, and column 9 adds distance to the Ruhr area to verify that the results are not entirely driven by industrialization in Germany's core coal and steel region. Across all specifications, the estimated β_t coefficients remain of similar magnitude and significance to the baseline estimates, suggesting that our results are not explained by omitted regional characteristics.

Alternative measures of consumer prices. Table A15 in the Appendix examines the sensitivity of our results to alternative measures of regional consumer prices. Our preferred index combines prices of five varieties of wood, two types of firewood (conifer and beech) and three types of timber (oak, pine, and spruce), reported in the same source as the wage data. Because firewood was an important substitute for coal during much of the nineteenth century, including it in the index may absorb part of the treatment effect, as lower firewood prices in coal-rich regions could themselves result from industrialization.[31] To address this concern, columns 3 and 4 construct separate indices based exclusively on either firewood or timber prices. Column 5 replaces the local wood price with the wheat price in the nearest market to capture changes in consumer prices unrelated to forestry products. Column 6 combines wood and wheat prices into a composite index. Across all specifications, the estimated β_t coefficients remain very similar to the baseline results, indicating that our findings are not driven by local price dynamics.

[29] In the 18th century, the eastern regions of Prussia differed from the western regions in terms of labor institutions and reliance on large-scale agriculture, with the Elbe forming a natural boundary (Tilly & Kopsidis, 2020).

[30] Through the annexations of October 1866, Prussia officially incorporated the Kingdom of Hanover, the Electorate of Hesse-Kassel, the Duchy of Nassau, the Free City of Frankfurt, and Schleswig-Holstein.

[31] Figure A10 in the Appendix presents event-study results using prices as dependent variables. Consistent with expectations, the price of firewood declined in closer proximity to coal after 1849, while the price of timber increased, reflecting higher construction activity due to urbanization and demand spillovers from mining activity in industrializing regions.

5 Industrialization and the demand for skilled labor

The results presented so far indicate a substantial increase in unskilled wages after 1849 in regions closer to coal deposits. In this section, we aim to develop a more comprehensive understanding of the impact of industrialization on labor incomes by examining its effects across different types of workers. For this purpose, we draw on an additional dataset containing approximately 3,000 household accounts from the German Empire. This dataset was compiled in a meta-study by H. K. Fischer (2011) and combines household-level budgets drawn from more than 150 historical sources. We include household budgets if they report the annual income of an adult male household member and provide information on his occupation, categorized at the 5-digit level according to the HISCO classification system for historical occupations, as applied by the author of the study. The data cover the period from 1859 to 1914 and include observations from all regions of the German Empire. Data frequency and regional coverage are illustrated in Figures A11–A13.

To analyze how incomes relate to coal proximity, we estimate the following pooled cross-sectional regression:

$$Income_{ipt} = +\alpha_p + \gamma_d + \beta Coal_i + \mu P_{id} + \delta X_i' + \lambda Z_{it}' + \epsilon_{iptd} \tag{3}$$

where the dependent variable, $Income_{iptd}$, denotes the (log) annual income in Marks of adult male household member i located in province p. This individual is observed in the data only once in year t. The key independent variable, $Coal_i$, is the (standardized) negative (log) distance in kilometers from the individual's location to the nearest Carboniferous rock strata. P_{id} denotes the average wheat price in the nearest market in decade d. The vector X_i' includes geographic control variables measured at the location of the individual. The vector Z_{it}' includes household-level control variables, comprising seven indicators for settlement size and an indicator for the presence of children in the household. The specification includes province-fixed effects, α_p, and decade-fixed effects, γ_d.

Table 3 presents the results from estimating Equation 3. Column 1 shows that labor incomes were generally higher in regions closer to coal. After adding controls (column 2), we find that a one SD increase in coal proximity is associated with an 10 percent increase in wages. Adding province-by-decade fixed effects does not change the overall results. To retain the largest possible sample, regressions include all observations from the German Empire, but column 4 confirms that the relationship also holds within Prussia.

In columns 5 and 6, we divide the sample into high-skilled (medium- and high-skilled) and low-skilled (unskilled and low-skilled) occupations, as derived from the HISCLASS classification that allows transforming the HISCO classification into a system of skill levels.[32] The results indicate that only low-skilled workers earned significantly higher labor incomes in regions closer to coal deposits, while there is no such relationship for high-skilled workers. Column 7 introduces an

[32] Occupations are classified into high-skilled (HISCLASS 1–4, 6–8) and low-skilled (HISCLASS 5, 9–12) categories. A finer distinction is not feasible, as the number of unskilled workers (HISCLASS 11 and 12) and high-skilled workers (HISCLASS 1 and 2) in the sample is small.

Table 3: Industrialization and the return to labor across different skill groups (pooled cross section)

Dependent variable:	Annual labor income								
	Geo controls	Additional controls	Province × Decade FE	Prussia sample	Low-skilled sample	High-Skilled sample	Interaction High-skilled=1	Mechanics sample	Interaction Non-routine=1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coal (proximity)	0.100***	0.098***	0.089***	0.117***	0.170***	0.008	0.153***	0.048**	0.045**
	(0.021)	(0.022)	(0.020)	(0.030)	(0.025)	(0.028)	(0.021)	(0.020)	(0.021)
Coal (proximity) × High-Skilled							-0.086***		
							(0.020)		
Coal (proximity) × Nonroutine									0.010
									(0.018)
High-Skilled							0.292***		
							(0.021)		
Nonroutine									0.067***
									(0.019)
Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
HH Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price Index	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Decade-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Province- × Decade-fixed effects	No	No	Yes	No	No	No	No	No	No
Observations	2,717	2,706	2,706	1,756	1,449	1,249	2,698	994	994
Adjusted R^2	0.49	0.54	0.63	0.52	0.59	0.54	0.58	0.47	0.48

Notes: The table shows results from estimating Equation 3. The unit of observation is an adult male recorded once between 1859 and 1914. The dependent variable captures the (log) annual labor income of the adult male in a given year. The main explanatory variable measures the negative (log) distance to coal deposits from the Carboniferous era in km. This variable is standardized with zero mean and unit SD. *High skilled* refers to medium- and high-skilled occupations, whereas *Low skilled* refers to unskilled and low-skilled occupation according to the HISCLASS scheme. *Mechanics* and *Non-routine* workers defined following Feldman and Van der Beek (2016). Geographic controls: slope, caloric suitability, distance to rivers, and distance to coast. Household controls (HH) controls: 7 settlement size indicators, indicator for presence of children in the household. Standard errors clustered at the county level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

interaction between coal proximity and an indicator for high-skilled occupations. The negative coefficient on the interaction term suggests that the wage gains associated with coal proximity were comparatively smaller for high-skilled workers.

The coarse high-/low-skilled classification may obscure important nuance in the types of skills demanded in industrializing regions. In particular, the high-skilled group includes many civil servants, teachers, priests, and similar occupations whose labor markets were not directly affected by industrial production. At the same time, this classification may overlook the importance of occupations that were essential because industrialization especially required workers that had the ability to install, operate, and maintain new equipment, i.e., the *tweakers*, in the spirit of Meisenzahl and Mokyr (2012). To account for this, we follow Cinnirella et al. (2025), De Pleijt et al. (2020), and Feldman and Van der Beek (2016) and focus on occupations that required mechanical skills in column 8. According to the HISCLASS scheme, roughly two thirds of these mechanics are classified as medium-skilled, and about one third as low-skilled. We find that among these mechanics, those working in closer proximity to coal earned significantly higher wages. However, the estimated coefficient is smaller than for the broader group of low-skilled workers, suggesting that wage gains were less pronounced for this group. Column 9 further distinguishes between mechanics performing more routine and more non-routine tasks. We find no statistically significant difference in wage gains between these two subgroups in proximity to coal.

Taken together, the results suggest that wage gains from coal proximity were concentrated in occupations requiring little formal training or skill. One may cautiously interpret this pattern as consistent with a "workshop-to-factory" hypothesis, according to which industrial production replaced skilled artisans with low-skilled machine operators (Acemoglu, 2002; Atack et al., 2004; Goldin & Katz, 1998; O'Rourke et al., 2013).

Table 4: Industrialization and industrial employment across different skill groups (cross section)

Dependent variable:	Manufacturing employment pc	Low-skilled Manufacturing employment p.c.	High-skilled Manufacturing employment p.c.	Low-skilled manufacturing employment share	Low-skilled non-Agricultrual employment share
	(1)	(2)	(3)	(4)	(5)
Coal (proximity)	0.452***	0.480***	0.290***	0.342***	0.530***
	(0.072)	(0.066)	(0.080)	(0.057)	(0.046)
Geo Controls	Yes	Yes	Yes	Yes	Yes
Observations	452	452	452	452	452
Adjusted R^2	0.20	0.22	0.15	0.21	0.26

Notes: The table shows results from cross-sectional OLS regressions of the dependent variables indicated in the column headers on the negative distance to coal deposits from the Carboniferous era in (log) kilometers. All dependent variables are measured in 1882. The unit of observation is a county. All variables are logged and standardized to have a zero mean and unit SD. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Robust standard errors in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

While the wage results suggest that coal-based industrialization primarily raised incomes among low-skilled workers, this interpretation relies on wages as a proxy for labor demand and does not account for employment quantities. To strengthen our interpretation, we complement the analysis with evidence on the occupational structure from the 1882 census. Because the census provides only county-level information for broadly defined industry classes, it does not allow us to classify individual workers using the HISCLASS scheme. However, based on the available occupational categories, we adopt a coarse classification: individuals recorded as self-employed or managers are classified as high-skilled, while white- and blue-collar employees are classified as low-skilled. Although this classification differs from the one used in the wage analysis, it may be particularly well suited to studying the 'workshop-to-factory' hypothesis, as artisans will fall into the high-skilled category by definition but will be classified as low-skilled once employed in a factory setting.

The results in columns 1–3 of Table 4 indicate that both high-skilled and low-skilled manufacturing workers constituted a larger share of the total population in regions closer to coal, consistent with higher levels of industrialization in these areas. However, the magnitude of the coal effect is substantially larger for low-skilled employment than for high-skilled employment, suggesting that coal-based industrialization was more strongly associated with low-skilled job creation. This pattern is confirmed when using the share of low-skilled workers in total manufacturing employment as the dependent variable (column 4). We find that regions closer to coal deposits employed a higher share of low-skilled workers within the manufacturing sector. We tentatively interpret the findings as further evidence that the demand for low-skilled labor was particularly high in coal-abundant regions. This pattern is further confirmed in column 5, where we extend the analysis to include low-skilled employment outside of agriculture.[33]

The comparatively smaller wage gains for high-skilled workers and mechanics, together with the greater prevalence of low-skilled employment in coal-rich regions, appear at odds with the view that industrialization increased the returns to human capital and formal skills. Moreover, earlier research has shown that Prussian industrialization was facilitated by its initially high levels of education (S. O. Becker et al., 2011; Cinnirella & Streb, 2017). However, we do not see our findings as contradicting this view, but rather as pointing to a more nuanced pattern. It is possible that educated or trained workers in industrializing regions earned their marginal product in lower-skilled occupations, with the higher wages observed in these jobs reflecting a premium on their human capital. The process of deskilling and mechanization may have displaced artisans into roles requiring less formal training, while at the same time raising wages in those roles to compensate for the skills they brought with them.

Having introduced wage data from multiple sources in Sections 4 and 5, we now offer a comparative perspective to assess the extent to which observed patterns generalize across datasets. Table A7 in the Appendix shows that unskilled workers consistently earned more in regions closer to coal across all sources. The results confirm that industrialization raised unskilled wages in a variety of settings, including both rural and urban areas. The standardized coefficients on coal proximity are remarkably similar when comparing forestry workers in 1875–1879 (column 1), ordinary day laborers in 1883 (column 2), rural day laborers (column 3), urban day laborers (column 4), and the annual labor income of low-skilled workers in the household accounts dataset (column 5).

6 *How* did industrialization affect labor incomes?

Having established that especially unskilled wages increased in response to coal-driven industrialization, we now turn to the channels through which industrialization affects labor incomes.

6.1 Potential mediators

Guided by theoretical considerations in the literature, we consider seven mediating factors through which industrialization may have affected unskilled wages. These can be broadly categorized as transformations on the firm side (technological change; organizational change; market integration) and changes on the worker side (sectoral change; deskilling). In addition to these core mechanisms, we also consider health degradation and migration as potential mediators of wage dynamics. We discuss each mediator's theoretical relevance, its relationship to proximity to coal, and its empirical operationalization below. See Appendix A.3 for detailed variable definitions.

Technological change. During industrialization production was fundamentally changed due to the adoption of new technologies that increased productivity in manufacturing. We expect such technological change to be most pronounced in regions close to coal deposits, which facilitated the adoption of steam engines and blast furnaces (De Pleijt et al., 2020; Franck & Galor, 2022).

[33] Note that the increase in the coefficient as compared to column 4 may be driven by the inclusion of mining.

The effect of technological change on wages hinges on two opposing forces: displacement and productivity effects (Acemoglu & Restrepo, 2020b). On the one hand, technologies can displace labor by mechanizing tasks formerly performed manually, thus reducing labor demand. On the other hand, increased productivity may encourage firms to expand production and increase labor demand. Some argue that early industrial technologies, particularly in textiles, were labor-saving and thus tended to depress wages (Allen, 2009b, 2021; Frey, 2019). Others view such technologies as augmenting labor and raising its productivity (De Pleijt et al., 2020; Goldin & Katz, 1998; Ridolfi et al., 2022). We capture technological change through the adoption of the general-purpose technology of the steam engine, using the change in total horsepower of steam engines per capita.

Organizational change. The transformation of production during industrialization was also characterized by the adoption of the factory system, with a shift from decentralized domestic work and small workshops to large-scale, centralized factory production. Larger industrial firms benefit from economies of scale and labor specialization, are more likely to rely on coal-powered technologies, and are more capital-intensive (Bartels et al., 2021; Pollard, 1965). These aspects arguably increase labor productivity and thus wages of unskilled workers. Furthermore, larger firms often face greater monitoring costs, prompting them to pay efficiency wages to motivate higher worker effort. Conversely, large firms might hold monopsony power allowing them to offer wages below competitive levels. We measure organizational change using changes in average firm size in the manufacturing sector.

Market integration. Another central aspect of industrialization was increasing market integration driven by improvements in transportation. Especially the emergence of railroads and steamboats massively reduced the cost of transportation and increased firm's access to markets. Market access is expected to be higher near coal fields, as railroad lines where particularly build to integrate coal-rich areas into the broader network. This may have affected unskilled wages through productivity effects and via economies of scale if larger firms require more basic forms of labor inputs, as discussed above. However, recent findings by Alvarez-Palau et al. (2025) suggest mixed effects of market access on different types of low-skilled labor during early industrialization in England, making theoretical predictions less clear. We measure changes in market integration using a market access index that aggregates the population of all counties, weighted by changes in transportation costs along least-cost routes.[34] We measure changes in market integration using a market access index that aggregates the population of all counties, weighted by changes in transportation costs along least-cost routes.

Sectoral change. Industrialization also entailed a shift in employment from agriculture to manufacturing. If labor productivity in the industrial sector exceeds productivity in the agricultural sector, such sectoral change should raise average wages W. A. Lewis (1954). We expect this reallocation to be more pronounced near coal deposits, where mining and coal-dependent industries provide employment opportunities and stimulate related sectors. We measure sectoral change as the change in the share of the population employed in manufacturing.

[34]See Appendix A.3 for further details.

Deskilling. An increase in the supply of low-skilled labor is generally expected to depress low-skilled wages. However, theories of directed technical change argue that abundant factors attract more innovation (Acemoglu, 2002; Habakkuk, 1962), implying that a larger supply of low-skilled workers may have encouraged the adoption of technologies that complement it, thereby raising the demand for low-skilled workers. In addition, local agglomeration effects can reinforce this mechanism. A larger local workforce can raise productivity through knowledge spillovers and specialization, potentially offsetting diminishing returns and even leading to an upward-sloping labor demand curve (Enrico, 2011). At the same time, the deskilling hypothesis posits that industrialization raised the demand for simpler, repetitive tasks that unskilled workers could perform efficiently. Such dynamics are likely to have been more pronounced in regions near coal deposits, where industrial production expanded more rapidly and with it the scope for unskilled tasks. We measure deskilling as the change in the share of the population employed in low-skilled occupations outside of agriculture.

Health degradation. Industrial jobs often entailed hazardous working conditions, pollution, and poor sanitation, particularly in fast-growing urban areas near coal fields (Hanlon, 2020). To attract workers under these conditions, firms may have paid compensating wage differentials. Such premiums would be more pronounced where living conditions were worse due to population pressures and industrial externalities. We proxy for local living conditions using infant mortality, a sensitive indicator of environmental and health-related risks, and interpret higher wages in high-mortality areas as potential compensating differentials.

Migration. Labor mobility is a natural response to regional wage differentials. While productivity-driven wage growth may attract labor inflows, such migration acts as an equalizing force, alleviating labor scarcity and moderating wage increases. Industrializing regions closer to coal deposits were likely to have experienced higher immigration, particularly of unskilled labor. We proxy migration by net population growth, defined as total population change net of births and deaths, expressed in per capita terms.

The description of the proposed mediators suggests that they are likely to be interdependent and not mutually exclusive. For example, while technological change is the mechanism most directly linked to coal abundance, other mediators may be related to coal only through their connection to technological change. As a result, the mediators are likely to be highly correlated, making it difficult to fully isolate their individual contributions. However, their combined effect can still be measured. The estimates presented below should therefore be interpreted with this in mind.

6.2 Mediation framework

To quantify each mechanisms contribution to explaining the positive effect of coal-driven industrialization on wages, we apply a mediation analysis. Figure 4 shows a representation of the model structure in path diagram notation. The idea of the mediation analysis is to decompose the total effect (TE) of coal-driven industrialization on wages into the indirect effect (IE), which works through the mechanisms, and the controlled direct effect (CDE), representing the (remaining) effect

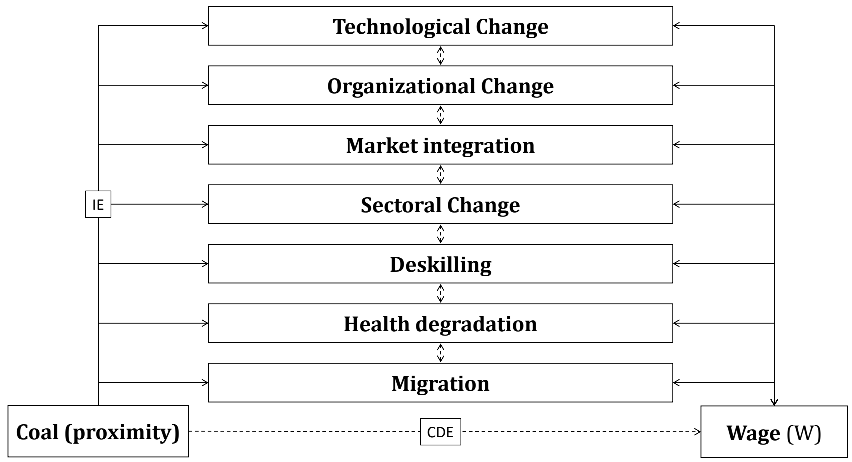


Figure 4: Mediation analysis path diagram

Notes: The figure illustrates the basic idea of the mediation analysis, which is to split the total effect (TE) of coal on wages into the controlled direct effect (CDE) and the indirect effect (IE) and to calculate the proportion mediated (PM)=IE/TE.

of coal on wages that is not captured by the mechanisms. The proportion mediated (PM) by each mediator is then obtained by dividing the IE by the TE. To obtain the decomposed effects, we run the following two-step model using county-level panel data.[35]

First-step regressions: In the first step, we estimate a series of panel regressions in which each of the seven potential mediators is regressed separately on proximity to coal deposits and a set of control variables:

$$M_{it,m} = \alpha_i + \gamma_t + \beta_m(Coal_i \times I_t^{1850}) + \mu P_{it} + \delta(X_i' \times I_t^{1850}) + \epsilon_{it,m}, \tag{4}$$

where $M_{it,m}$ denotes mediator m in county i and period t, and all other variables are defined as in Equation (1). Due to data constraints, the panel is limited to two periods: 1849, just before the estimated increase in unskilled wages in response to industrialization, and 1875/1882, toward the end of our wage panel. The coefficient β_m captures the differential change in mediator m as a function of proximity to coal.[36]

Second-step regressions: In the second step, we estimate a panel regression of unskilled wages on proximity to coal deposits and the full vector of mediators:

$$Wage_{it} = \alpha_i + \gamma_t + \beta(Coal_i \times I_t^{1850}) + \delta_m M_{it}' + \mu P_{it} + \delta(X_i' \times I_t^{1850}) + \varepsilon_{it}, \tag{5}$$

[35]It is common practice in mechanism analyses to employ so-called horse-race regressions, which correspond to the second-step regression of our model. However, a horse-race regression alone cannot establish whether the mediator is indeed affected by the treatment. Confirming this is essential since mediation cannot occur without a treatment effect on the mediator.

[36]To simplify interpretation, both the treatment and mediator variables are standardized to have mean zero and unit variance. We restrict the sample to counties with non-missing observations for all relevant variables. Missing wage data is the primary source of attrition.

where $Wage_{it}$ is the average wage of unskilled forestry workers within 30 kilometers around the centroid of a county i in period t (i.e., 1850–1859 and 1875–1879). The coefficient β captures the controlled direct effect (CDE) of proximity to coal on wages, conditional on mediators and controls. The vector δ_m contains the indirect effects of the mediators on wages. All other variables are defined as in Equation (4).

6.3 Mediation results

Table 5: Transmission channels linking industrialization to labor incomes - first step

Dependent variable:	Steam engines p.c.	Firm size	Market access	Manufacturing employment p.c.	Low-skilled employment p.c.	Infant mortality	Migration rate
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coal (proximity) $\times I^{1850}$	0.306***	0.035	0.182***	0.259***	0.534***	0.095**	0.029
	(0.033)	(0.067)	(0.054)	(0.072)	(0.092)	(0.044)	(0.093)
Geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price index	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	550	550	550	550	550	550	550
Adjusted R^2	0.97	0.80	0.58	0.19	0.27	0.88	0.14

Notes: The table shows results from estimating Equation 4. The unit of observation is a county, observed at two points in time (1849 and 1875/1882). Dependent variables are indicated in column heads. Coal (proximity) measures the negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with a dummy that is one in the second period. All variables are logged and standardized to have a zero mean and unit SD. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the county level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

We report results of the first-step regressions in Table 5. Each mediator is consecutively regressed on proximity to coal deposits and controls. The estimates indicate that regions closer to coalfields experienced stronger technological change, as measured by increases in per capita horsepower of steam engines (column 1); higher market integration, proxied by changes in market access (column 3); greater sectoral change, measured by growth in manufacturing employment per capita (column 4); more deskilling, as proxied by rising per capita employment of low-skilled workers outside agriculture (column 5); and more intense health degradation, as reflected in rising infant mortality rates (column 6). By contrast, we find no evidence that regions closer to coal had significantly larger firms (column 2) or higher net migration rates (column 7). A comparison of standardized coefficients across columns suggests that technological change and deskilling are the two mediators most strongly associated with proximity to coal.

Table 6 reports the second-step regressions, in which we examine to what extent the direct effect of coal proximity on unskilled wages is mediated by the various channels. We start out with repeating the reduced form results in column 1. A one SD increase in proximity to coal deposits is associated with an increase in unskilled wages by 8 *Pfennig*, or 0.21 SD. Columns 2 through 8 sequentially add each mediator to the specification. We find that technological change, market integration, sectoral change, deskilling, and health degradation are all positively associated with

Table 6: Transmission channels linking industrialization to labor incomes - second step

Dependent variable:	Unskilled wages 1850–59 to 1875–79									to 1883
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coal (proximity) $\times I^{1850}$	0.079***	0.045***	0.079***	0.074***	0.069***	0.060***	0.076***	0.079***	0.039***	0.012
	(0.013)	(0.014)	(0.013)	(0.012)	(0.013)	(0.013)	(0.013)	(0.013)	(0.014)	(0.021)
Steam engines (hp) p.c.		0.114***							0.072**	0.164***
		(0.025)							(0.029)	(0.033)
Firm size			0.003							
			(0.013)							
Market access				0.031**					0.016	0.011
				(0.013)					(0.013)	(0.017)
Manufacturing employment p.c.					0.039***				0.011	-0.025
					(0.012)				(0.014)	(0.017)
Low-skilled employment p.c.						0.037***			0.021**	0.052***
						(0.009)			(0.010)	(0.013)
Infant mortality rate							0.032**		0.015	-0.019
							(0.016)		(0.015)	(0.021)
Net migration rate								0.003		
								(0.009)		
Geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price index	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	550	550	550	550	550	550	550	550	550	550
Adjusted R^2	0.90	0.91	0.90	0.91	0.91	0.91	0.90	0.90	0.91	0.87

Notes: The table shows results from estimating Equation 5. The unit of observation is a county, observed at two points in time (1849 and 1875/1882). The dependent variable measures average wages of male seasonal forestry workers in *Mark* in 1850–59 and 1875–79, except for column 10 were second period wages are measured in 1883. Coal (proximity) measures the negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with a dummy that is one in the second period. All explanatory variables are logged and standardized to have a zero mean and unit SD. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the county level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

unskilled wages. Among these, technological change appears to be the most important transmission channel. Including this mediator reduces the direct effect of coal by ca. 44% (column 2). This finding suggests that the productivity effect of steam engines outweighs their labor-displacement effect. Deskilling also emerges as an important channel: Roughly 25% of the coal effect operates through increased employment of low-skilled workers in non-agricultural sectors (column 6). This finding is consistent with Marshallian externalities, whereby agglomeration economies and localized knowledge spillovers raise productivity and wages even as labor supply increases (Enrico, 2011; Moretti, 2010).[37] It is also consistent with the interpretation that industrialization attracted low-skilled workers into manufacturing and related industries, where demand expanded faster than supply and thus pushed up wages.

The other mediators market access, sectoral change, and health degradation contribute only modestly to explaining the coal effect. As shown in column 9, these channels are also not robust determinants of unskilled wages once technological change and deskilling are accounted for. In this specification, only the latter two remain significantly associated with unskilled wages, while the direct effect of coal is reduced by approximately half. Notably, the deskilling channel remains relevant even when controlling for technological change, suggesting that the rising demand for low-skilled workers is not solely driven by the adoption of new technologies, but also reflects the broader transformation of production associated with industrialization.

One potential concern is that, in the second period of our panel dataset, deskilling is measured in

[37]Note that this pattern emerges despite controlling for changes in local consumer prices.

1882, slightly after the wage outcome period of 1875–1879, raising the possibility of reverse causality. As a robustness check, we re-estimate the model using unskilled wages in 1883 as the outcome variable. The results, reported in column 10, confirm the main pattern. In this specification, the inclusion of mediators renders the coal coefficient statistically insignificant and small (see Table A16 in the Appendix for the full set of results). This pattern is consistent with the notion of complete mediation, in which the effect of coal-driven industrialization on unskilled wages operates almost entirely through the mediators technological change and deskilling.

While the effects of individual mediators cannot be cleanly isolated, we assess their joint contribution in Columns 9 and 10 using the approach by Jérolon et al. (2021).[38] According to this, the mediators collectively account for 51 percent of the total effect in Column 9 and 85 percent in Column 10, implying controlled direct effects of 49 percent and 15 percent, respectively.

In sum, the results of our mediation analysis indicate that technological change and deskilling are the primary channels through which coal-driven industrialization affected the wages of unskilled workers. This suggest that the productivity effects outweigh any displacement effects associated with the adoption of steam engines. We acknowledge, however, that displacement effects may have been more pronounced for other technologies not captured in our data. While other mechanisms such as firm size or sectoral change, appear to have played a more limited role, they may still have mattered in ways not fully captured by our measures.

7 Conclusion

This paper examines how coal-driven industrialization affected the wages of workers in 19th-century Prussia. Using novel panel data and a difference-in-differences design that exploits spatial variation in coal proximity, we show that unskilled wages rose persistently in industrializing regions from the mid-19th century onward. This finding is confirmed using individual-level data from household accounts and further supported by much smaller wage gains among high-skilled workers and mechanical occupations. Mediation analysis indicates that wage increases for unskilled workers in close proximity to coal were primarily driven by technological change and deskilling rather than through organizational change, market integration, or migration.

These findings contribute to our understanding of how industrialization shaped the distribution of income during the transition to modern growth. They suggest that the early stages of Prussian industrialization were characterized by labor-augmenting rather than labor-saving technological change, which increased the demand for low-skilled workers. These results highlight that the distributional consequences of industrialization depended crucially on the nature of technological change and its interaction with local resource endowments. In contrast to the conventional view that industrialization initially widened inequality through capital deepening and skill-biased technological progress, coal-based industrialization in Prussia appears to have raised the incomes of unskilled workers and thereby compressed wage differentials.

[38] This quasi-Bayesian approach is suited for settings such as ours where multiple mediators are correlated to one another but not causally ordered.

References

Acemoglu, D. (2002). Directed Technical Change. *Review of Economic Studies*, *69*(4), 781–809.

Acemoglu, D. (2010). When does labor scarcity encourage innovation? *Journal of Political Economy*, *118*(6), 1037–1078.

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

Acemoglu, D., & Restrepo, P. (2020a). Robots and jobs: Evidence from us labor markets. *Journal of Political Economy*, *128*(6), 2188–2244.

Acemoglu, D., & Restrepo, P. (2020b). The wrong kind of AI? Artificial intelligence and the future of labour demand. *Cambridge Journal of Regions, Economy and Society*, *13*(1), 25–35.

Allcott, H., & Keniston, D. (2018). Dutch disease or agglomeration? the local economic effects of natural resource booms in modern america. *Review of Economic Studies*, *85*(2), 695–731.

Allen, R. C. (2009a). *The British industrial revolution in global perspective*. Cambridge University Press.

Allen, R. C. (2009b). Engels' pause: Technical change, capital accumulation, and inequality in the british industrial revolution. *Explorations in Economic History*, *46*(4), 418–435.

Allen, R. C. (2015). The high wage economy and the industrial revolution: A restatement. *The Economic History Review*, *68*(1), 1–22.

Allen, R. C. (2021, November). *The interplay among wages, technology and globalisation: The labour market and inequality, 1620–2020* (tech. rep.). The IFS.

Alvarez-Palau, E. J., Bogart, D., Satchell, M., & Shaw-Taylor, L. (2025). Transport and urban growth in the first industrial revolution. *The Economic Journal*, *135*(668), 1191–1228.

Andersson, D., Karadja, M., & Prawitz, E. (2022). Mass migration and technological change. *Journal of the European Economic Association*, *20*(4), 1859–1896.

Aragón, F. M., & Rud, J. P. (2013). Natural resources and local communities: Evidence from a peruvian gold mine. *American Economic Journal: Economic Policy*, *5*(2), 1–25.

Asch, K., & Bellenberg, S. (2005). *The 1: 5 million international geological map of Europe and adjacent areas (IGME 5000)[cartographic material]*. Bundesanstalt für Geowissenschaften und Rohstoffe.

Ashraf, Q. H., Cinnirella, F., Galor, O., Gershman, B., & Hornung, E. (2025). Structural Change, Elite Capitalism, and the Emergence of Labour Emancipation. *Review of Economic Studies*, *2*(92), 808–836.

Atack, J., Bateman, F., & Margo, R. A. (2004). Skill Intensity and Rising Wage Dispersion in Nineteenth-Century American Manufacturing. *The Journal of Economic History*, *64*(1), 172–192.

Atack, J., Margo, R. A., & Rhode, P. W. (2019). "Automation" of Manufacturing in the Late Nineteenth Century: The Hand and Machine Labor Study. *Journal of Economic Perspectives*, *33*(2), 51–70.

Ayres, R. U., & Warr, B. (2009, February). *The Economic Growth Engine: How Energy and Work Drive Material Prosperity*. Edward Elgar Publishing.

Banken, R. (1993). Die Diffusion der Dampfmaschine in Preußen um 1830. *Jahrbuch für Wirtschaftsgeschichte / Economic History Yearbook*, *34*(2), 219–248.

Bartels, C., Kersting, F., & Wolf, N. (2021, March). Testing Marx. Income Inequality, Concentration, and Socialism in Late 19th Century Germany.

Bartik, A. W., Currie, J., Greenstone, M., & Knittel, C. R. (2019). The local economic and welfare consequences of hydraulic fracturing. *American Economic Journal: Applied Economics*, *11*(4), 105–155.

Becker, S., Boll, P. D., & Voth, H.-J. (2025). *Spatial unit roots in regressions: A practitioner's guide and a stata package* (CESifo Working Paper Series No. 11651). CESifo.

Becker, S. O., Cinnirella, F., Hornung, E., & Woessmann, L. (2014). iPEHD—The ifo Prussian Economic History Database. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, *47*(2), 57–66.

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

Becker, S. O., & Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History*. *Quarterly Journal of Economics*, *124*(2), 531–596.

Berbée, P., Braun, S. T., & Franke, R. (2024). Reversing fortunes of German regions, 1926–2019: Boon and bane of early industrialization? *Journal of Economic Growth*.

Black, D., McKinnish, T., & Sanders, S. (2005). The economic impact of the coal boom and bust. *Economic Journal*, *115*(503), 449–476.

Broadberry, S. N., Campbell, B. M. S., Klein, A., Overton, M., & Leeuwen, B. v. (2015). *British economic growth, 1270-1870*. Cambridge University Press.

Butts, K. (2023). *Difference-in-differences estimation with spatial spillovers* (tech. rep.). University of Colorado Boulder, mimeo.

Caprettini, B., & Voth, H.-J. (2020). Rage against the Machines: Labor-Saving Technology and Unrest in Industrializing England. *American Economic Review: Insights*, *2*(3), 305–320.

Caselli, F., & Michaels, G. (2013). Do oil windfalls improve living standards? evidence from brazil. *American Economic Journal: Applied Economics*, *5*(1), 208–238.

Châtel & Dollfus. (1931). *Atlas Châtel et Dollfus: Les Houillères Européennes*. Société de Documentation Industrielle.

Cinnirella, F. (2008). Optimists or pessimists? A reconsideration of nutritional status in Britain, 1740-1865. *European Review of Economic History*, *12*(3), 325–354.

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

Cinnirella, F., Hornung, E., & Koschnick, J. (2025). Flow of ideas: Economic societies and the rise of useful knowledge. *The Economic Journal*, *135*(669), 1496–1535.

Cinnirella, F., & Streb, J. (2017). The role of human capital and innovation in economic development: Evidence from post-Malthusian Prussia. *Journal of Economic Growth*, *22*(2), 193–227.

Clark, G. (2007). *A Farewell to Alms: A Brief Economic History of the World*. Princeton University Press.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics*, *92*(1), 1–45.

Crafts, N. (2005). Market potential in british regions, 1871–1931. *Regional Studies*, *39*(9), 1159–1166.

De Pleijt, A., Nuvolari, A., & Weisdorf, J. (2020). Human Capital Formation During the First Industrial Revolution: Evidence from the use of Steam Engines. *Journal of the European Economic Association*, *18*(2), 829–889.

De Pleijt, A. M., & Weisdorf, J. L. (2017). Human capital formation from occupations: The 'deskilling hypothesis' revisited. *Cliometrica*, *11*, 1–30.

Eggert, R. (2001). Mining and economic sustainability: National economies and local communities. *A Study Prepared for the Mining, Minerals, and Sustainable Development Project, Colorado School of Mines*.

Eggert, U. (1883). Die bewegung der holzpreise und tagelohn-sätze in den preussischen staatsforsten von 1800 bis 1879. In E. Blenck (Ed.), *Zeitschrift des königlich preussischen statistischen bureaus* (pp. 1–44, Vol. 23). Verl. d. Königl. Statist. Bureaus.
Enrico, M. (2011). Local labor markets. In D. Card & O. Ashenfelter (Eds.), *Handbook of labor economics* (pp. 1237–1313, Vol. 4). Elsevier.
Ericsson, J., & Molinder, J. (2020). Economic growth and the development of real wages: Swedish construction workers' wages in comparative perspective, 1831–1900. *The Journal of Economic History*, *80*(3), 813–852.
Esposito, E., & Abramson, S. F. (2021). The European coal curse. *Journal of Economic Growth*, *26*(1), 77–112.
Federico, G., Schulze, M.-S., & Volckart, O. (2021). European Goods Market Integration in the Very Long Run: From the Black Death to the First World War. *The Journal of Economic History*, *81*(1), 276–308.
Feinstein, C. H. (1998). Pessimism Perpetuated: Real Wages and the Standard of Living in Britain during and after the Industrial Revolution. *The Journal of Economic History*, *58*(3), 625–658.
Feldman, N. E., & Van der Beek, K. (2016). Skill choice and skill complementarity in eighteenth century england. *Explorations in Economic History*, *59*, 94–113.
Fernihough, A., & O'Rourke, K. H. (2021). Coal and the European Industrial Revolution. *The Economic Journal*, *131*(635), 1135–1149.
Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: New 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315.
Fischer, H. K. (2011, December). *Konsum im Kaiserreich: Eine statistisch-analytische Untersuchung privater Haushalte im wilhelminischen Deutschland*. Akademie Verlag.
Fischer, W. (2011). Germany's mining production statistics from 1850 to 1914. Statistik der Bergbauproduktion Deutschlands 1850–1914.
Franck, R. (2024). Labor scarcity, technology adoption and innovation: Evidence from the cholera pandemics in 19th century france. *Journal of Economic Growth*, *29*(4), 543–583.
Franck, R., & Galor, O. (2021). Flowers of evil? industrialization and long run development. *Journal of Monetary Economics*, *117*, 108–128.
Franck, R., & Galor, O. (2022). Technology-Skill Complementarity in Early Phases of Industrialisation. *The Economic Journal*, *132*(642), 618–643.
Fremdling, R. (1979). Modernisierung und Wachstum der Schwerindustrie in Deutschland, 1830-1860. *Geschichte und Gesellschaft*, *5*(H. 2), 201–227.
Frey, C. B. (2019, December). *The Technology Trap: Capital, Labor, and Power in the Age of Automation*. Princeton University Press.
Fritzsche, M., & Wolf, N. (2023). *Fickle fossils. economic growth, coal and the european oil invasion, 1900-2015* (tech. rep.). CESifo Working Paper.
Fuest, C., Peichl, A., & Siegloch, S. (2018). Do Higher Corporate Taxes Reduce Wages? Micro Evidence from Germany. *American Economic Review*, *108*(2), 393–418.
Gallardo-Albarrán, D., & De Jong, H. (2021). Optimism or pessimism? a composite view on english living standards during the industrial revolution. *European Review of Economic History*, *25*(1), 1–19.
Galloway, P. (2007). Galloway prussia database 1861 to 1914.
Galor, O. (2005). Chapter 4 From Stagnation to Growth: Unified Growth Theory. In *Handbook of Economic Growth* (pp. 171–293, Vol. 1). Elsevier.
Galor, O. (2011). *Unified growth theory*. Princeton University Press.

Galor, O., & Özak, Ö. (2016). The Agricultural Origins of Time Preference. *American Economic Review*, *106*(10), 3064–3103.
Goldin, C., & Katz, L. F. (1998). The Origins of Technology-Skill Complementarity. *The Quarterly Journal of Economics*, *113*(3), 693–732.
Gutberlet, T. (2012). Cheap coal, market access, and industry location in Germany 1846 to 1882. *University of Arizona*.
Gutberlet, T. (2014). Mechanization and the spatial distribution of industries in the German Empire, 1875 to 1907. *The Economic History Review*, *67*(2), 463–491.
Habakkuk, H. J. (1962). *American and British technology in the nineteenth century: The search for labour saving inventions*. Cambridge University Press.
Hanlon, W. W. (2020). Coal Smoke, City Growth, and the Costs of the Industrial Revolution. *The Economic Journal*, *130*(626), 462–488.
Henning, F.-W. (1973). Die Industrialisierung in Deutschland 1800 bis 1914, Paderborn, F.
Herzig, T., & Ott, H. (1986). *Statistik der öffentlichen Elektrizitätsversorgung Deutschland 1890–1913* (Vol. 1). Scripta Mercaturae Verlag.
Hicks, J. (1932). *The theory of wages*. Macmillan.
Hoffmann, W. G. (1963). The Take-off in Germany. In W. W. Rostow (Ed.), *The Economics of Take-Off into Sustained Growth* (pp. 95–118). Palgrave Macmillan UK.
Holtfrerich, C.-L. (2005). Quantitative Economic History of Coal Mining in the Ruhr Area in the 19th Century.
Hornbeck, R., & Naidu, S. (2014). When the levee breaks: Black migration and economic development in the american south. *American Economic Review*, *104*(3), 963–990.
Humphries, J., & Weisdorf, J. (2019). Unreal wages? real income and economic growth in england, 1260–1850. *The Economic Journal*, *129*(623), 2867–2887.
Hunt, E. H. (1986). Industrialization and Regional Inequality: Wages in Britain, 1760–1914. *The Journal of Economic History*, *46*(4), 935–966.
IEG. (2010). Ieg-maps: Server for digital historical maps [Leibniz Institute of European History, Mainz, Germany. http://www.ieg-maps.uni-mainz.de/].
Jacobsen, G. D., & Parker, D. P. (2014). The economic aftermath of resource booms: Evidence from boomtowns in the american west. *Economic Journal*, *126*(593), 1092–1128.
Jacobsen, G. D., Parker, D. P., & Winikoff, J. B. (2023). Are resource booms a blessing or a curse? evidence from people (not places). *Journal of Human Resources*, *58*(2), 393–420.
Jérolon, A., Baglietto, L., Birmelé, E., Alarcon, F., & Perduca, V. (2021). Causal mediation analysis in presence of multiple mediators uncausally related. *The International Journal of Biostatistics*, *17*(2), 191–221.
Johnson, N. D., & Koyama, M. (2017). Jewish communities and city growth in preindustrial europe. *Journal of Development Economics*, *127*, 339–354.
Kander, A., & Stern, D. I. (2014). Economic growth and the transition from traditional to modern energy in Sweden. *Energy Economics*, *46*, 56–65.
Kaufhold, K. H., & Sachse, W. (2011). Industry of Prussia before 1850. Vol. 1: Mining, Metallurgy, and saltworks.
Kelly, M., Mokyr, J., & Ó Gráda, C. (2023). The mechanics of the Industrial Revolution. *Journal of Political Economy*, *131*(1), 59–94.
Königlich Preussisches Statistisches Bureau. (1861–1934). *Preussische statistik* (E. Kühn, Ed.; Vol. 1–305). Verl. d. Königl. Statist. Bureaus.
Königlich Preussisches Statistisches Bureau. (1861–1914). *Zeitschrift des königlich preussischen statistischen bureaus/landesamts* (E. Blenck, Ed.; Vol. 1–54). Verl. d. Königl. Statist. Bureaus.

Kopsidis, M., & Bromley, D. W. (2017). Explaining German Economic Modernization: The French Revolution, Prussian Reforms, and the Inevitable Continuity of Change. *Annales. Histoire, Sciences Sociales*, *72*(4), 729–766.

Lewis, E. (2011). Immigration, skill mix, and capital–skill complementarity. *Quarterly Journal of Economics*, *126*(2), 1029–1069.

Lewis, W. A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, *22*(2), 139–191.

Lindert, P. H., & Williamson, J. G. (1983). English Workers' Living Standards During the Industrial Revolution: A New Look. *The Economic History Review*, *36*(1), 1–25.

Malm, A. (2016). *Fossil capital: The rise of steam power and the roots of global warming*. Verso.

Marshall, A. (1920). *The Principles of Economics* (8th ed.). Macmillan.

Meisenzahl, R. R., & Mokyr, J. (2012). The rate and direction of invention in the british industrial revolution: Incentives and institutions. In J. Lerner & S. Stern (Eds.), *The rate and direction of inventive activity revisited* (pp. 443–479). University of Chicago Press.

Meitzen, A. (1868). *Der boden und die landwirthschaftlichen verhältnisse des preussischen staates* (Vol. 1-4). Berlin: Verlag von Paul Parey.

Michaels, G. (2011). The long term consequences of resource-based specialisation. *The Economic Journal*, *121*(551), 31–57.

Mokyr, J., Sarid, A., & van der Beek, K. (2022). The wheels of change: Technology adoption, millwrights, and persistence in britain's industrialization. *Economic Journal*, *132*(645), 1894–1926.

Mokyr, J., Vickers, C., & Ziebarth, N. L. (2015). The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different? *Journal of Economic Perspectives*, *29*(3), 31–50.

Moretti, E. (2010). Local Multipliers. *American Economic Review*, *100*(2), 373–377.

Morris, M., Kaplinsky, R., & Kaplan, D. (2012). "One thing leads to another"—Commodities, linkages and industrial development. *Resources Policy*, *37*(4), 408–416.

Neuhaus, G. (1904). Die ortsüblichen Tagelöhne gewöhnlicher Tagearbeiter in Preußen 1892 und 1901. *Zeitschrift des Königlich Preussischen Statistischen Buraus*, *44*, 310–346.

Nicholas, S., & Steckel, R. H. (1991). Heights and Living Standards of English Workers During the Early Years of Industrializations, 1770–1815. *The Journal of Economic History*, *51*(4), 937–957.

Nuvolari, A. (2002). The 'machine breakers' and the industrial revolution. *Journal of European Economic History*, *31*(2), 393–426.

O'Rourke, K. H., Rahman, A. S., & Taylor, A. M. (2013). Luddites, the industrial revolution, and the demographic transition. *Journal of Economic Growth*, *18*(4), 373–409.

Pfister, U. (2018). Real Wages in Germany during the First Phase of Industrialization, 1850-1889. *Jahrbuch für Wirtschaftsgeschichte / Economic History Yearbook*, *59*(2), 567–596.

Pierenkemper, T. (2009). The Rise and Fall of the 'Normalarbeitsverhaeltnis' in Germany. *IZA DP No. 4068*.

Pollard, S. (1965). *The genesis of modern management: A study of the industrial revolution in great britain*. Edward Arnold.

Pomeranz, K. (2000). *The Great Divergence: China, Europe, and the Making of the Modern World Economy*. Princeton University Press.

Ridolfi, L., Salvo, C., & Weisdorf, J. (2022). *The impact of mechanisation on wages and employment: Evidence from the diffusion of steam power* (tech. rep.). CEPR Discussion Papers 17389.

San, S. (2023). Labor supply and directed technical change: Evidence from the termination of the bracero program in 1964. *American Economic Journal: Applied Economics*, *15*(1), 136–163.

Schmitz, J. (1888). *Übersicht der für die sämtlichen deutschen Bundesstaaten in Gemäßheit des § 8 des Reichsgesetzes betreffend die Krankenversicherung der Arbeiter vom 15. Juni 1883 festgestellten ortsüblichen Tagelöhne gewerblicher Tagearbeiter* (2nd ed.). Heuser's Verlag.

Schneider, M. (1989). Kleine Geschichte der Gewerkschaften. *Ihre Entwicklung in Deutschland von den Anfängen bis heute. Bonn.*

Schulz, T. (1911). *Die Entwickelung des deutschen Steinkohlenhandels unter besonderer Berücksichtigung von Ober-und Niederschlesien* [PhD Thesis]. Niederschles. Dr. u. Verlagsanstalt.

Sparhawk, W. N. (1938). *Forests and employment in Germany.* US Department of Agriculture.

Statistisches Bureau zu Berlin. (1851–1855). *Tabellen und amtliche nachrichten über den preussischen staat für das jahr 1849* (Vol. 1–6b). A. W. Hayn.

Tilly, R. H., & Kopsidis, M. (2020). *From old regime to industrial state: A history of German industrialization from the eighteenth century to World War I.* London : University of Chicago Press.

Treitschke, K. (1928). *Die wirtschaftlichen und sozialen Verhältnisse der preußischen Staatsforstarbeiter*. Enckehaus.

van der Ploeg, F. (2011). Natural resources: Curse or blessing? *Journal of Economic Literature*, *49*(2), 366–420.

Voth, H.-J. (2004, January). Living standards and the urban environment. In R. Floud & P. Johnson (Eds.), *The Cambridge Economic History of Modern Britain* (1st ed., pp. 268–294). Cambridge University Press.

Voth, H.-J., Caprettini, B., & Trew, A. (2023). *Fighting for growth: Labor scarcity and technological progress during the british industrial revolution* (tech. rep.). CEPR Discussion Papers 17881.

Wrigley, E. A. (2013). Energy and the English Industrial Revolution. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *371*(1986), 20110568.

Zeira, J. (1998). Workers, machines, and economic growth. *Quarterly Journal of Economics*, *113*(4), 1091–1117.

Zentralblatt. (1892–1914). *Zentralblatt für das Deutsche Reich* (R. d. Innern, Ed.; Vol. 20-42). Carl Lehmanns Verlag.

A Appendix

A.1 Additional Tables

A.1.1 Descriptive statistics

Table A1: Descriptive statistics: forestry-level dataset

Variable	N	Mean	SD	Min	Max
Forestry-level: wages and prices					
Unskilled wages	3541	1.02	0.38	0.2	2.7
Spruce	1178	12.5	5.22	0.33	32
Pine	2287	12.1	4.63	1.14	32
Beech	2511	4.58	2.2	0.07	19.2
Conifer	2653	2.91	1.42	0.17	9
Oak	2505	20.2	8.05	3	83.3
Wheat price (nearest city)	3541	86.9	22	40	160
County-level: coal and controls					
Coal (proximity)	667	-172	142	-587	0
Coal (30km dummy)	667	0.08	0.27	0	1
Distance to navigable river	667	23	21	0.03	145
Distance to coast	667	178	107	0.07	461
Caloric suitability index	667	1682	313	527	2265
Slope	667	1.2	1.1	0.01	5.4
Distance to London	667	905	300	432	1555
Distance to Wittenberg	667	305	163	25	746
Distance to Poland	667	389	243	1.6	832
Distance to county border	667	4.2	3	0.03	18
Distance to Ruhr area	667	382	288	0	1037
West Elbe	667	0.52	0.5	0	1
French occupation	667	0.53	0.5	0	1
Prussia 1849	667	0.61	0.49	0	1

Notes: This table presents summary statistics for variables used in the main DiD and event-study analysis.

Table A2: Descriptive statistics: household account dataset

Variable	N	Mean	SD	Min	Max
Annual income	2706	1710	1991	5	24720
Annual income (Low-skilled)	1457	1161	550	6	4402
Annual income (High-skilled)	1257	2342	2729	5	24720
Annual income (Mechanics)	994	1474	594	29.5	4980
High-Skilled	2698	0.46	0.5	0	1
Coal (proximity)	2706	-96	105	-514	0
Children	2706	0.91	0.29	0	1
Location size class	2706				
... Large city (>100,000 inhabitants)	1424	0.53			
... Medium-sized town (20,000 - 100,000 inhabitants)	392	0.14			
... Small town (5,000 - 20,000 inhabitants)	256	0.09			
... Rural town (2,000 - 5,000 inhabitants)	66	0.02			
... Spa town	4	0			
... Village (2,000 - 5,000 inhabitants)	309	0.11			
... Village (<2,000 inhabitants)	255	0.09			
Wheat price (nearest city)	2706	173	31	66	219

Notes: This table presents summary statistics for variables used in the analysis of household account data in Table 3.

Table A3: Descriptive statistics: county-level cross-sectional dataset

Variable	N	Mean	SD	Min	Max
Coal (proximity)	452	-147	142	-599	0
Industrial employment pc 1882	452	0.12	0.065	0.021	0.63
Manufacturing employment pc 1882	452	0.11	0.057	0.02	0.63
Mining employment pc 1882	452	0.011	0.033	0	0.36
Agricultural employment pc 1882	452	0.2	0.075	0.00092	0.33
Steam engines (hp) pc 1875	452	0.0096	0.011	0.000071	0.073
Firm size 1875	452	28	25	3	245
Urbanization rate 1875	452	0.28	0.22	0	1
Low-skilled manufacturing employment pc 1882	452	0.064	0.04	0.0097	0.37
High-skilled manufacturing employment pc 1882	452	0.042	0.021	0.011	0.26
Low-skilled manufacturing employment share 1882	452	0.59	0.09	0.31	0.81
Low-skilled non-Agricultural employment share 1882	452	0.45	0.15	0.13	0.86

Notes: This table presents summary statistics for cross-sectional estimates at the county level.

Table A4: Descriptive statistics: mediation analysis dataset

Variable	N	Mean	SD	Min	Max
Variables measured in ca. 1849					
Coal (proximity)	275	-150	152	-599	0
Unskilled Wages 1850-59	275	0.87	0.2	0.5	1.5
Steam engines (hp) pc 1849	275	0	0.01	0	0.1
Firm size 1849	275	7.3	8.2	0.77	57
Market access 1849	275	21	5.8	10	45
Manufacturing employment p.c. 1849	275	0.02	0.03	0	0.32
Low-skilled employment pc 1849	275	0.16	0.05	0.06	0.5
Infant mortality rate 1849	275	0.16	0.03	0.09	0.28
Migration rate 1846–49	275	-0.03	0.1	-0.53	0.59
Variables measured in ca. 1875/1882					
Coal (proximity)	275	-150	152	-599	0
Unskilled Wages 1874-75	275	1.3	0.3	0.78	2.2
Unskilled Wages 1883	275	1.3	0.33	0.8	2.4
Steam engines (hp) pc 1875	275	0.01	0.01	0	0.07
Firm size 1875	275	26	23	3.8	245
Market access 1875	275	27	9.4	9.2	69
Manufacturing employment pc 1875	275	0.03	0.03	0	0.19
Low-skilled employment pc 1882	275	0.14	0.06	0.05	0.36
Infant mortality rate 1875	275	0.25	0.05	0.15	0.44
Migration rate 1867-75	275	-0.07	0.07	-0.27	0.34

Notes: This table presents summary statistics for mediation analysis at the county level.

A.1.2 Missing wage data

Table A5: Industrialization and non-missing wage data (by decade)

Dependent variable:	Unskilled wage available (dummy)								
	1800-09	1810-19	1820-29	1830-39	1840-49	1850-59	1860-69	1870-74	1875-79
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coal (proximity)	0.009	0.047**	0.101***	0.123***	0.139***	0.125***	0.047***	0.020**	0.010
	(0.016)	(0.022)	(0.024)	(0.022)	(0.021)	(0.020)	(0.015)	(0.008)	(0.006)
Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	667	667	667	667	667	667	667	667	667
Adjusted R^2	-0.002	0.01	0.04	0.05	0.06	0.05	0.01	0.005	0.003

Notes: The table reports results from estimating linear probability models in cross-sections of forestries by decade. The dependent variables are indicators assuming the value one if data on wages for day laborers is non-missing in a given decade. Coal (proximity) measures the negative distance to coal deposits from the Carboniferous era in (log) kilometers. This variable is standardized with zero mean and unit SD. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Robust standard errors in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A6: Share of non-missing wage data available

Year	Coal	Non-Coal
1800–1809	11.76	17.05
1810–1809	33.33	25.65
1820–1829	58.82	37.01
1830–1839	76.47	47.73
1840–1849	88.24	60.39
1850–1859	94.12	69.81
1860–1869	96.08	87.50
1870–1874	100.00	95.13
1875–1879	100.00	97.24

Notes: The table shows the share of non-missing wage observations by decade for forestries located within 30 km of coal deposits from the Carboniferous era (Coal) and those located farther away (Non-Coal).

A.1.3 Comparing results across various wage measures

Table A7: Comparing results across various wage measures (cross section)

Dependent variable:	Wages of:				
	Forestry workers 1875–79	Ordinary day laborers 1883 (weighted average)	Ordinary day laborers 1883 (urban)	Ordinary day laborers 1883 (rural)	HH labor income (low-skilled) 1863–1912
	(1)	(2)	(3)	(4)	(5)
Coal (proximity)	0.279***	0.363***	0.383***	0.321***	0.271***
	(0.060)	(0.051)	(0.078)	(0.054)	(0.040)
Geo Controls	Yes	Yes	Yes	Yes	Yes
HH controls	No	No	No	No	Yes
Price Index	Yes	Yes	Yes	Yes	Yes
Decade-fixed effects	No	No	No	No	Yes
Province-fixed effects	No	No	No	No	Yes
Observations	398	452	218	427	1,449
Adjusted R^2	0.29	0.45	0.34	0.47	0.59

Notes: The table reports results from cross-sectional OLS regressions of the dependent variables indicated in the column headers on the negative distance to coal deposits from the Carboniferous era in (log) kilometers. The unit of observation is a forestry (1), county (2–4) or an individual (5). All variables are standardized to have a zero mean and unit SD. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Robust standard errors in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level..

A.1.4 Coal as proxy for industrialization

Table A8: Coal deposits and industrial outcomes 1875-82 (cross-section)

Dependent variable:	Industrial employment pc 1882 (1)	Mining employment pc 1882 (2)	Manufacturing employment pc 1882 (3)	Agricultural employment pc 1882 (4)	Steam engines (hp) pc 1875 (5)	Firm size 1875 (6)	Urbanization rate 1875 (7)
Coal (proximity)	0.665***	0.598***	0.452***	-0.480***	0.607***	0.596***	0.245***
	(0.051)	(0.129)	(0.072)	(0.071)	(0.054)	(0.054)	(0.069)
Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	452	452	452	452	452	452	452
Adjusted R^2	0.41	0.26	0.20	0.25	0.29	0.38	0.10

Notes: The table reports results from cross-sectional OLS regressions of the dependent variables indicated in the column headers on the negative distance to coal deposits from the Carboniferous era in (log) kilometers. The unit of observation is a county. All variables are logged and standardized to have a zero mean and unit SD. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Robust standard errors in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

A.1.5 Rolling Estimates

Table A9: Industrialization and the return to unskilled labor (simple DiD with alternative time cut-offs)

Dependent variable:	Unskilled wages				
	Post 1819	Post 1829	Post 1839	Post 1849	Post 1859
	(1)	(2)	(3)	(4)	(5)
Coal (proximity) × Post	-0.021*	-0.008	0.012	0.034***	0.055***
	(0.012)	(0.007)	(0.008)	(0.009)	(0.008)
Geo Controls	Yes	Yes	Yes	Yes	Yes
Price Index	Yes	Yes	Yes	Yes	Yes
Unit-fixed effects	Yes	Yes	Yes	Yes	Yes
Year-fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	808	1,124	1,438	1,764	2,733
Adjusted R^2	0.81	0.82	0.84	0.88	0.87

Notes: The table shows results from estimating Equation 1. The unit of observation is a forestry. Each column is confined to 40 year windows in which outcomes are observed in decadal intervals. The dependent variable measures average wages of male seasonal forest workers in *Mark* during a decade. Coal (proximity) measures the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with a dummy that is one after the year indicated in the column header. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the forestry level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

A.1.6 Robustness checks on main results

Table A10: Specification robustness (event-study regressions)

Dependent variable:	Unskilled wages									
	Without controls	Geo controls	Baseline	Wage (log)	County level	Conley SE	Spatial Inf	Reference period 1800-09	Balanced sample	Extended sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coal (proximity) × 1800-09	-0.005	-0.021	0.009	0.030	-0.016	0.009	0.009			-0.034*
	(0.016)	(0.022)	(0.030)	(0.045)	(0.010)	(0.034)	(0.028)			(0.018)
Coal (proximity) × 1810-19	0.007	0.004	0.019	0.034	0.013	0.019	0.019	0.010		0.023*
	(0.013)	(0.016)	(0.021)	(0.023)	(0.011)	(0.019)	(0.016)	(0.025)		(0.013)
Coal (proximity) × 1820-29	0.010	0.011	0.012	0.012	0.005	0.012	0.012	0.003	0.000	0.017*
	(0.009)	(0.012)	(0.012)	(0.013)	(0.009)	(0.012)	(0.012)	(0.028)	(0.010)	(0.010)
Coal (proximity) × 1830-39	-0.008	-0.006	-0.004	-0.004	-0.009	-0.004	-0.004	-0.012	-0.007	-0.008
	(0.006)	(0.007)	(0.007)	(0.008)	(0.006)	(0.008)	(0.008)	(0.028)	(0.008)	(0.006)
Coal (proximity) × 1840-49								-0.009		
								(0.030)		
Coal (proximity) × 1850-59	0.008	0.025***	0.026***	0.021***	0.021***	0.026*	0.026*	0.017	0.010	0.023***
	(0.006)	(0.008)	(0.008)	(0.007)	(0.007)	(0.014)	(0.011)	(0.030)	(0.008)	(0.008)
Coal (proximity) × 1860-69	0.014	0.040***	0.039***	0.022**	0.033***	0.039*	0.039	0.031	0.021*	0.046***
	(0.009)	(0.011)	(0.011)	(0.010)	(0.009)	(0.021)	(0.022)	(0.030)	(0.012)	(0.008)
Coal (proximity) × 1870-74	0.062***	0.083***	0.085***	0.045***	0.064***	0.085***	0.085***	0.077**	0.068***	0.070***
	(0.010)	(0.012)	(0.013)	(0.008)	(0.012)	(0.021)	(0.021)	(0.033)	(0.018)	(0.011)
Coal (proximity) × 1875-79	0.068***	0.081***	0.083***	0.040***	0.068***	0.083***	0.083***	0.075**	0.072***	0.078***
	(0.010)	(0.013)	(0.013)	(0.010)	(0.012)	(0.020)	(0.010)	(0.030)	(0.017)	(0.012)
Coal (proximity) × 1883										0.080***
										(0.016)
Coal (proximity) × 1892										0.074***
										(0.016)
Coal (proximity) × 1897										0.048***
										(0.017)
Coal (proximity) × 1901										0.107***
										(0.018)
Coal (proximity) × 1905										0.108***
										(0.019)
Coal (proximity) × 1910										0.094***
										(0.023)
Coal (proximity) × 1914										0.136***
										(0.023)
Geo Controls	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Wood Price Index	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unit fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,647	3,647	3,541	3,541	2,928	3,541	3,541	3,541	1,610	6,098
Adjusted R^2	0.88	0.89	0.89	0.90	0.92	0.89	0.89	0.89	0.88	0.93
Unit	Forestry	Forestry	Forestry	Forestry	County	Forestry	Forestry	Forestry	Forestry	County

Notes: The table shows results from estimating Equation 2. The unit of observation is a forestry (columns 1–4, 6–9) or a county (columns 5, 9), observed in decadal intervals from the 1800s to the 1870s, extended with individual years until 1914 in column 10. The omitted period is 1840—1849, except for column 8. The dependent variable measures average wages of male seasonal forest workers in *Mark* during a decade. Coal (proximity) measures the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with time dummies. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors, clustered at the forestry level (1–4, 8, 9), the county level (5,10), or adjusted for spatial autocorrelation following Conley (1999) in (6) and S. Becker et al. (2025) in (7), in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A11: Alternative measures of coal proximity (event-study regressions)

Dependent variable:	Unskilled wages					
	Baseline	Coal mines	Hard coal mines	Coal Outside Prussia	Different age definition	Main rock type definition
	(1)	(2)	(3)	(4)	(5)	(6)
Coal (proximity) × 1800-09	0.009	0.010	0.022	0.009	-0.039	0.014
	(0.030)	(0.016)	(0.020)	(0.030)	(0.027)	(0.019)
Coal (proximity) × 1810-19	0.019	-0.006	0.020	0.017	-0.004	0.004
	(0.021)	(0.015)	(0.018)	(0.020)	(0.020)	(0.015)
Coal (proximity) × 1820-29	0.012	-0.007	0.006	0.012	0.000	0.004
	(0.012)	(0.014)	(0.012)	(0.012)	(0.013)	(0.011)
Coal (proximity) × 1830-39	-0.004	-0.006	-0.008	-0.003	0.003	0.000
	(0.007)	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)
Coal (proximity) × 1850-59	0.026***	0.007	0.025***	0.024***	0.031***	0.025***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)	(0.007)
Coal (proximity) × 1860-69	0.039***	0.007	0.035***	0.038***	0.046***	0.039***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.008)
Coal (proximity) × 1870-74	0.085***	0.046***	0.054***	0.086***	0.113***	0.105***
	(0.013)	(0.013)	(0.013)	(0.012)	(0.014)	(0.012)
Coal (proximity) × 1875-79	0.083***	0.046***	0.046***	0.086***	0.111***	0.110***
	(0.013)	(0.014)	(0.014)	(0.013)	(0.013)	(0.011)
Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes
Price Index	Yes	Yes	Yes	Yes	Yes	Yes
Unit fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,541	3,541	3,541	3,541	3,541	3,541
Adjusted R^2	0.89	0.89	0.89	0.89	0.89	0.89

Notes: The table shows results from estimating Equation 2. The unit of observation is a forestry, observed in decadal intervals from the 1800s to the 1870s. The omitted period is 1840—1849. The dependent variable measures average wages of male seasonal forest workers in *Mark* during a decade. Coal (proximity) measures the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with time dummies. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the forestry level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A12: Alternative distance cut-offs (simple DiD)

Dependent variable:	Unskilled wages				
	10 km	20 km	30 km	40 km	50 km
	(1)	(2)	(3)	(4)	(5)
Coal (10km dummy) × post1849	0.159***				
	(0.041)				
Coal (20km dummy) × post1849		0.126***			
		(0.038)			
Coal (30km dummy) × post1849			0.114***		
			(0.029)		
Coal (40km dummy) × post1849				0.079***	
				(0.023)	
Coal (50km dummy) × post1849					0.066***
					(0.020)
Geo Controls	Yes	Yes	Yes	Yes	Yes
Price Index	Yes	Yes	Yes	Yes	Yes
Unit fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Observations	3,541	3,541	3,541	3,541	3,541
Adjusted R^2	0.88	0.88	0.88	0.88	0.88

Notes: The table shows results from estimating Equation 1. The unit of observation is a forestry, observed in decadal intervals from the 1800s to the 1870s. The dependent variable measures average wages of male seasonal forest workers in *Mark* during a decade. The main explanatory variable is an interaction between a dummy indicating whether the forestry is located within the number of kilometers from a Carboniferous-era coal deposit, as specified in the column header, interacted with a dummy that is one from the 1850s. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the forestry level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A13: Spillovers (simple DiD)

Dependent variable:	Unskilled wages		
	Baseline	Spillover 1	Spillover 2
	(1)	(2)	(3)
Coal (30km dummy) × Post 1849	0.114***	0.109***	0.126***
	(0.029)	(0.029)	(0.032)
Coal (30-60km dummy) x Post 1849		-0.020	-0.003
		(0.020)	(0.023)
Coal (60-90km dummy) x Post 1849			0.037
			(0.025)
Geo Controls	Yes	Yes	Yes
Price Index	Yes	Yes	Yes
Unit fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	3,541	3,541	3,541
Adjusted R^2	0.88	0.88	0.88

Notes: The table shows results from estimating Equation 1. The unit of observation is a forestry. The dependent variable measures average wages of male seasonal forest workers in *Mark*, observed in decadal intervals from the 1800s to the 1870s. The main explanatory variables are interactions between dummies indicating whether the forestry is located within the number of kilometers from a Carboniferous-era coal deposit specified in the variable label, interacted with a dummy that is one from the 1850s. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the forestry level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A14: Additional control variables and potential confounders (event-study regressions)

Dependent variable:	Unskilled wages								
	Baseline	West Elbe	French Occupation	Prussia 1849	Distance London	Distance Wittenberg	Distance Poland	Distance County Border	Distance Ruhr Area
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coal (proximity) × 1800-09	0.009	0.003	0.009	0.011	0.008	0.012	0.016	0.004	0.018
	(0.030)	(0.032)	(0.033)	(0.033)	(0.034)	(0.035)	(0.037)	(0.033)	(0.033)
Coal (proximity) × 1810-19	0.019	0.020	0.023	0.017	0.029	0.021	0.032	0.012	0.034
	(0.021)	(0.022)	(0.022)	(0.022)	(0.023)	(0.023)	(0.023)	(0.020)	(0.027)
Coal (proximity) × 1820-29	0.012	0.010	0.012	0.012	0.011	0.013	0.014	0.008	0.019
	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.013)	(0.012)	(0.012)	(0.013)
Coal (proximity) × 1830-39	-0.004	-0.004	-0.002	-0.003	-0.003	-0.004	-0.002	-0.007	-0.004
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)	(0.008)
Coal (proximity) × 1850-59	0.026***	0.023***	0.024***	0.028***	0.019**	0.029***	0.024***	0.028***	0.025***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
Coal (proximity) × 1860-69	0.039***	0.033***	0.033***	0.045***	0.030***	0.043***	0.037***	0.041***	0.036***
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.012)	(0.013)
Coal (proximity) × 1870-74	0.085***	0.070***	0.068***	0.098***	0.050***	0.080***	0.057***	0.087***	0.058***
	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.014)	(0.012)	(0.014)	(0.014)
Coal (proximity) × 1875-79	0.083***	0.060***	0.059***	0.100***	0.043***	0.077***	0.051***	0.083***	0.053***
	(0.013)	(0.011)	(0.011)	(0.012)	(0.012)	(0.014)	(0.012)	(0.013)	(0.015)
Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price Index	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Unit fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,541	3,541	3,541	3,541	3,541	3,541	3,541	3,209	3,541
Adjusted R^2	0.89	0.89	0.90	0.89	0.90	0.89	0.90	0.89	0.89

Notes: The table shows results from estimating Equation 2. The unit of observation is a forestry, observed in decadal intervals from the 1800s to the 1870s. The omitted period is 1840—1849. The dependent variable measures average wages of male seasonal forest workers in *Mark* during a decade. Coal (proximity) measures the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with time dummies. Each column adds to the specification the control variable indicated in the column head, interacted with time dummies. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the forestry level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

Table A15: Alternative measures of consumer prices (event-study regressions)

Dependent variable:	Unskilled wages					
	Baseline	Without price index	Timber price index	Firewood price index	Wheat price nearest market	Commodity price index
	(1)	(2)	(3)	(4)	(5)	(6)
Coal (proximity) × 1800-09	0.009	-0.021	0.055*	0.010	-0.025	-0.021
	(0.030)	(0.022)	(0.032)	(0.031)	(0.021)	(0.022)
Coal (proximity) × 1810-19	0.019	0.004	0.055**	0.019	0.009	0.005
	(0.021)	(0.016)	(0.025)	(0.021)	(0.017)	(0.016)
Coal (proximity) × 1820-29	0.012	0.011	0.018	0.013	0.010	0.011
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Coal (proximity) × 1830-39	-0.004	-0.006	-0.005	-0.003	-0.005	-0.006
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Coal (proximity) × 1850-59	0.026***	0.025***	0.029***	0.025***	0.021**	0.025***
	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
Coal (proximity) × 1860-69	0.039***	0.040***	0.042***	0.039***	0.047***	0.041***
	(0.011)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)
Coal (proximity) × 1870-74	0.085***	0.083***	0.087***	0.086***	0.085***	0.084***
	(0.013)	(0.012)	(0.013)	(0.013)	(0.012)	(0.012)
Coal (proximity) × 1875-79	0.083***	0.081***	0.084***	0.082***	0.086***	0.081***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.012)	(0.013)
Wood price index	0.002					
	(0.011)					
Timber price index			0.000			
			(0.010)			
Firewood price index				0.006		
				(0.010)		
Geo Controls	Yes	Yes	Yes	Yes	Yes	Yes
Unit fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,541	3,647	3,395	3,485	3,647	3,647
Adjusted R^2	0.89	0.89	0.89	0.89	0.89	0.89

Notes: The table shows results from estimating Equation 2. The unit of observation is a forestry, observed in decadal intervals from the 1800s to the 1870s. The omitted period is 1840–1849. The dependent variable measures average wages of male seasonal forest workers in *Mark* during a decade. Coal (proximity) measures the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with time dummies. Each column alters the measure to account for regional and intertemporal variation in consumer prices indicated in the column head. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the forestry level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

A.1.7 Robustness checks on mediation analysis

Table A16: Transmission channels linking industrialization to labor incomes (1849 and 1883) - second step

Dependent variable:	Unskilled wages 1850–59 to 1883								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coal (proximity) $\times I^{1850}$	0.084***	0.023	0.085***	0.077***	0.075***	0.049**	0.083***	0.083***	0.016
	(0.021)	(0.022)	(0.020)	(0.020)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)
Steam engines (hp) p.c.		0.199***							0.133***
		(0.032)							(0.038)
Firm size			-0.039**						-0.034*
			(0.016)						(0.020)
Market access				0.041**					0.011
				(0.017)					(0.017)
Manufacturing employment p.c.					0.036**				-0.003
					(0.016)				(0.021)
Low-skilled employment p.c.						0.066***			0.051***
						(0.012)			(0.013)
Infant mortality rate							0.016		
							(0.023)		
Net migration rate								0.019	
								(0.013)	
Geo controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Price index	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
County-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time-fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	550	550	550	550	550	550	550	550	550
Adjusted R^2	0.84	0.86	0.85	0.85	0.85	0.86	0.84	0.85	0.87

Notes: The table shows results from estimating Equation 5. The unit of observation is a county, observed at two points in time (1849 and 1875/1882). The dependent variable measures average wages of male seasonal forestry workers in *Mark* in 1850–59 and 1883. Coal (proximity) measures the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with a dummy that is one in the second period from 1850. Geo controls: slope, caloric suitability, distance to rivers, and distance to coast. Standard errors clustered at the county level in parenthesis. *** denotes statistical significance at the 1% level, ** at the 5% level, and * at the 10% level.

A.2 Additional Figures

A.2.1 Descriptive statistics

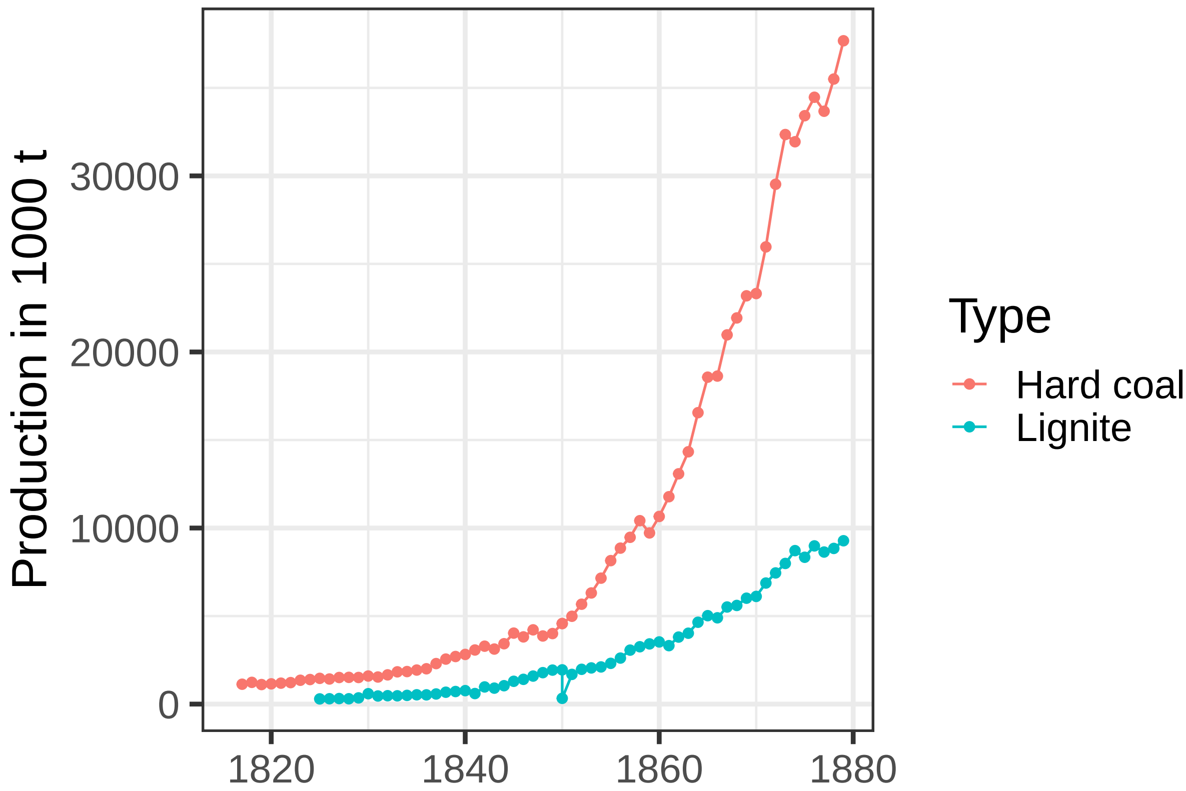


Figure A1: Coal production in Prussia 1817 to 1879

Notes: The figure shows the annual coal production from hard coal and lignite in Prussia in 1,000 t from 1817 to 1879. Data sources: W. Fischer, 2011; Holtfrerich, 2005; Kaufhold and Sachse, 2011.

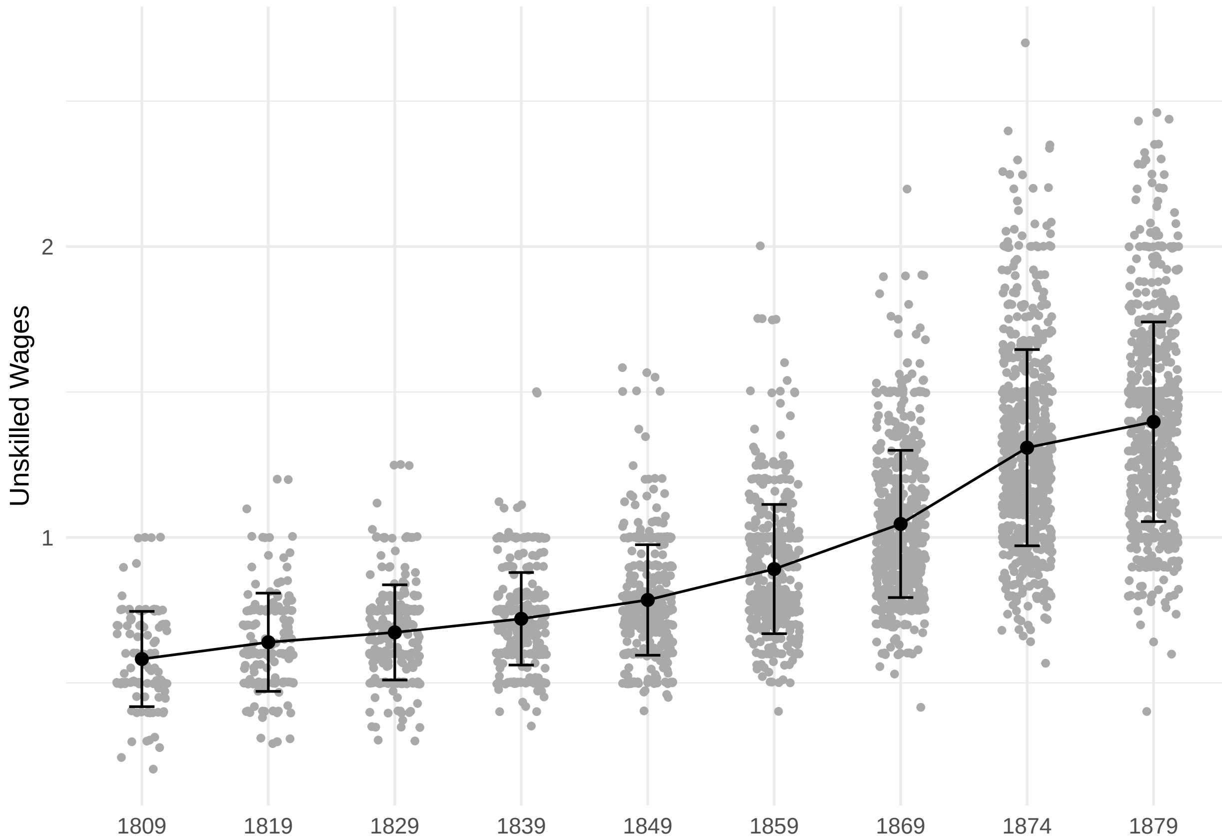


Figure A2: Nominal wages of day laborers over time.

Notes: This figures shows the distribution and mean nominal wages of day laborers employed in Prussia's state forestries in *Mark* from 1800 to 1879.

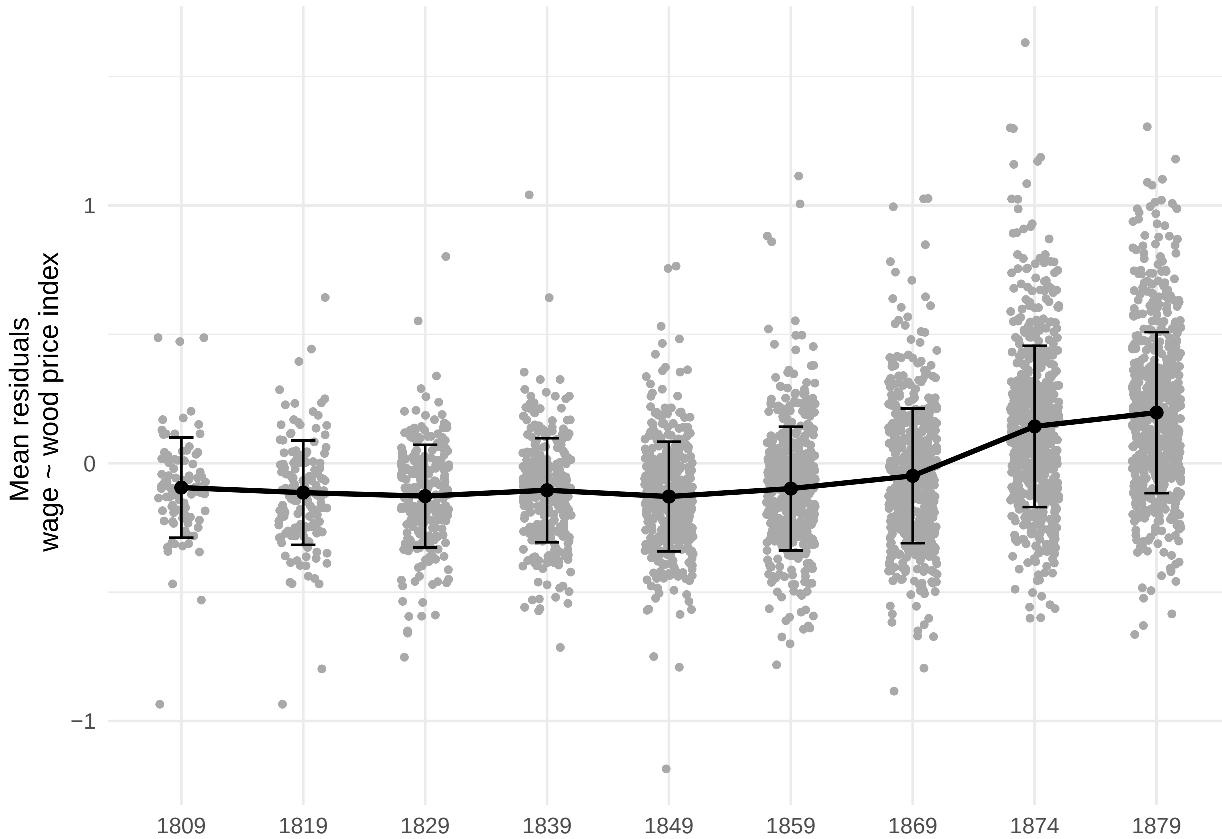


Figure A3: Real wages of day laborers over time.

Notes: This figure shows the distribution of residualized nominal wages of day laborers employed in Prussia's state forestries in *Mark* from 1800 to 1879. Residuals are obtained from a regression of nominal wages on the wood price index. Residuals capture deviations of nominal wages from the price-predicted trend and serve as a proxy for real wages.

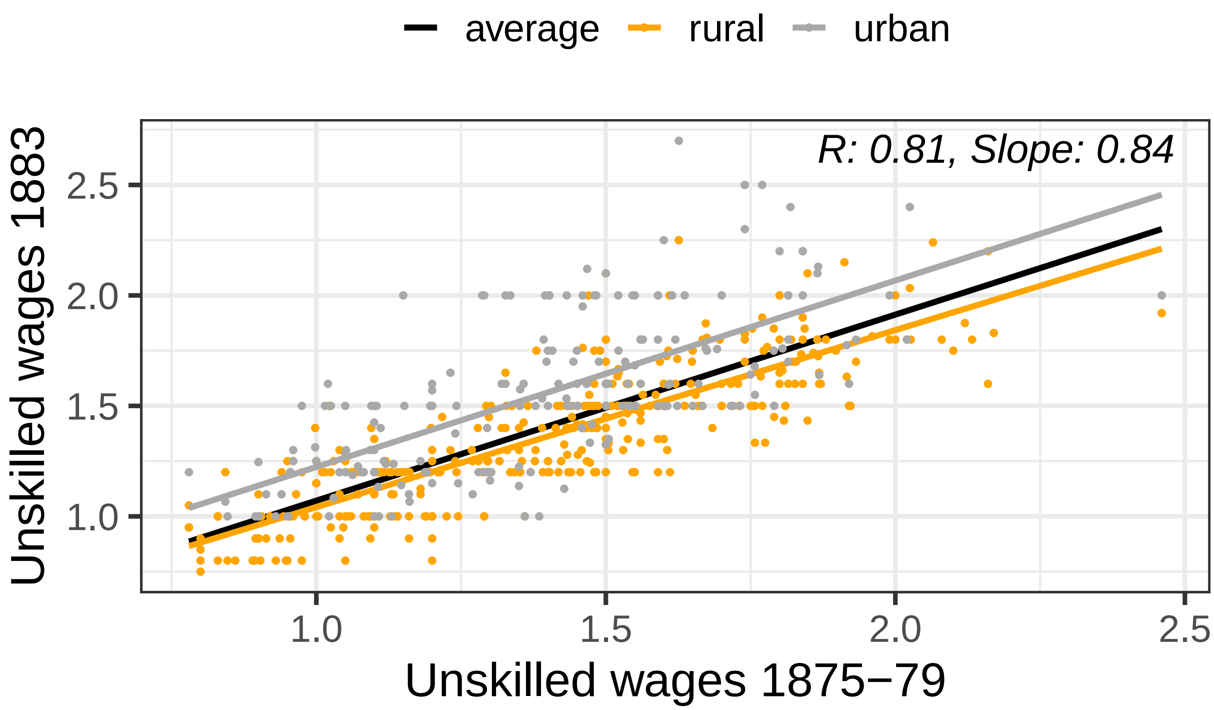


Figure A4: Forestry day-laborer wages (1875–79) and ordinary day-laborer wages (1883)

Notes: The scatter plot shows the correlation between forestry day-laborer wages in 1875–79 from U. Eggert, 1883 and ordinary male day-laborer wages in 1883 from Schmitz, 1888 in rural (R: 0.84, Slope: 0.8) and urban areas (R: 0.7, Slope: 0.84), as well as a weighted average (R: 0.79, Slope: 0.78). All variables are measured at the county level.

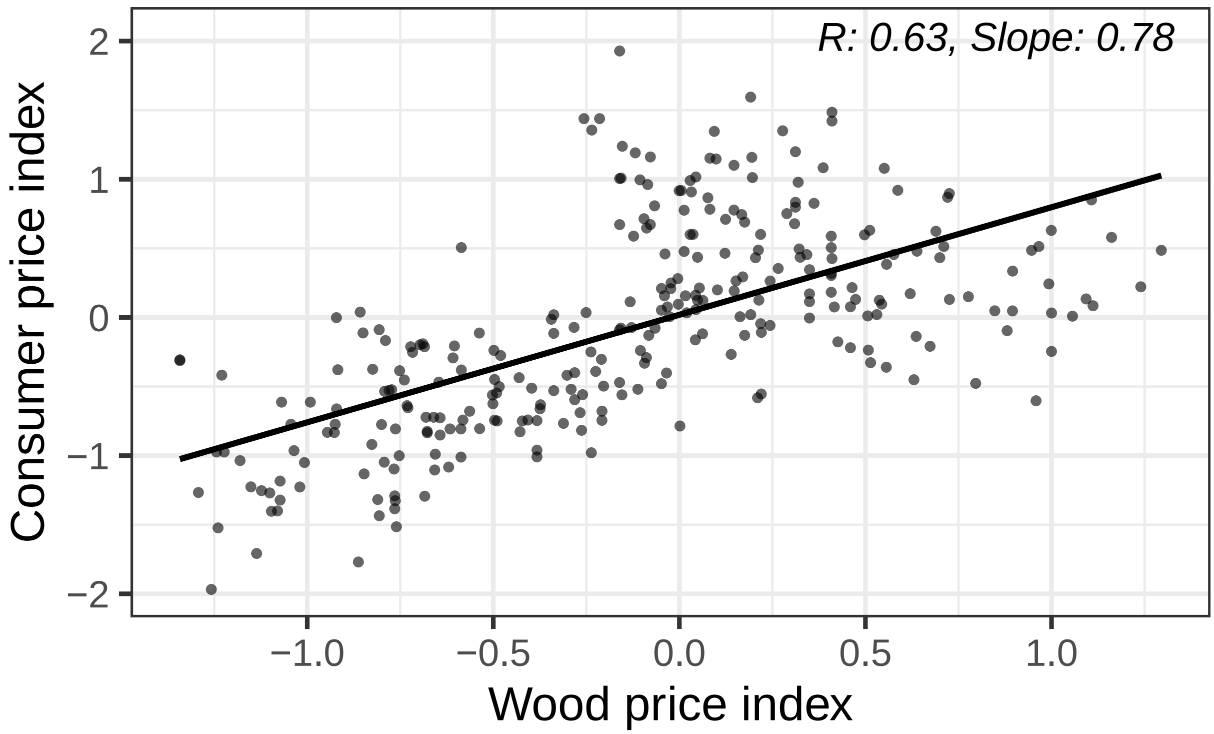


Figure A5: Wood-price index (1830–69) and consumer-price index (1837–1860)

Notes: The scatter plot shows the correlation between the wood-price index (conifer, beech, oak, pine and spruce) averaged over the period 1830–69 and the consumer-price index (wheat, rye, barley, oats, rape, potatoes, straw, hay) based on average market in the period 1837–1860. Variables are measured at the county level.

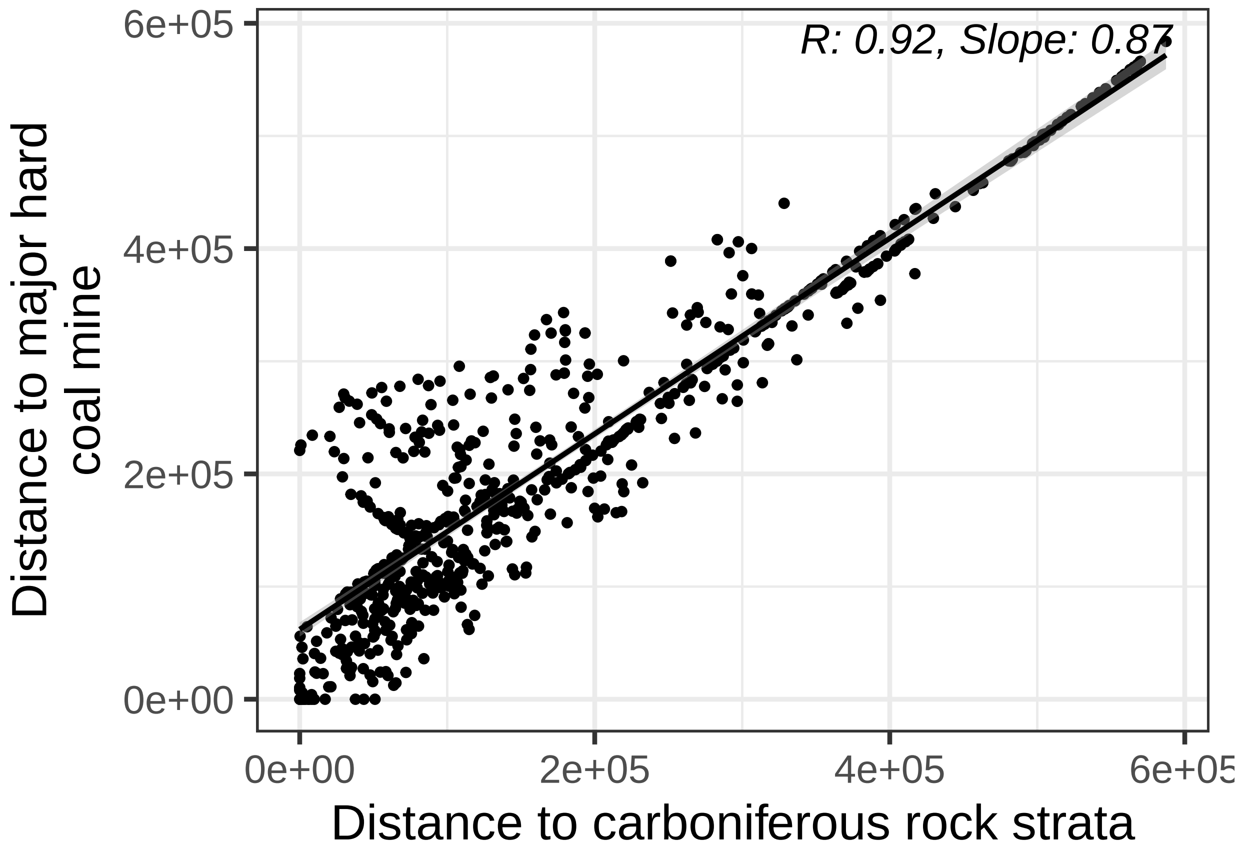
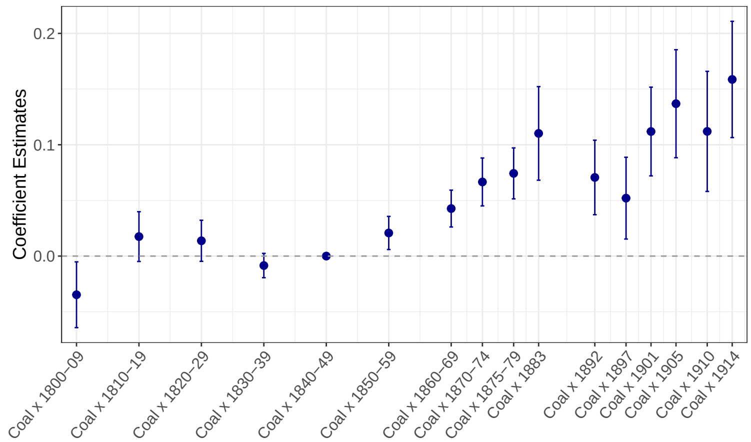


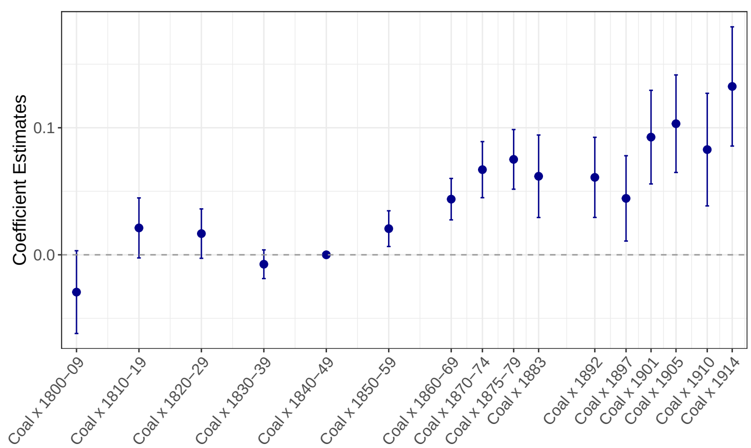
Figure A6: Distance to nearest major coalfield and nearest carboniferous rock strata

Notes: The scatter plot shows the correlation between the distance of a forestry to the nearest major hard coalfield and the nearest Carboniferous rock strata in Prussia. Data source: (Asch & Bellenberg, 2005; Fernihough & O'Rourke, 2021).

A.2.2 Robustness checks on main results



(a) urban



(b) rural

Figure A7: Urban and rural wages 1883–1914 (event study results)

Notes: Figures plot β_t coefficients from Equation 2, with 95% confidence intervals. The omitted period is 1840–1849. The dependent variable measures average wages of male seasonal forestry workers in *Mark* during a decade until 1879, extended using male day-laborer wages for individual years between 1883 and 1914. In Figure A7a this extension pertains only to wages paid in urban locations, whereas in A7b it pertains to wages paid in rural locations. The main explanatory variables measure the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with time dummies. Results are conditional a full set of geographic controls (slope, caloric suitability, distance to rivers, and distance to coast). Standard errors are clustered at the county level.

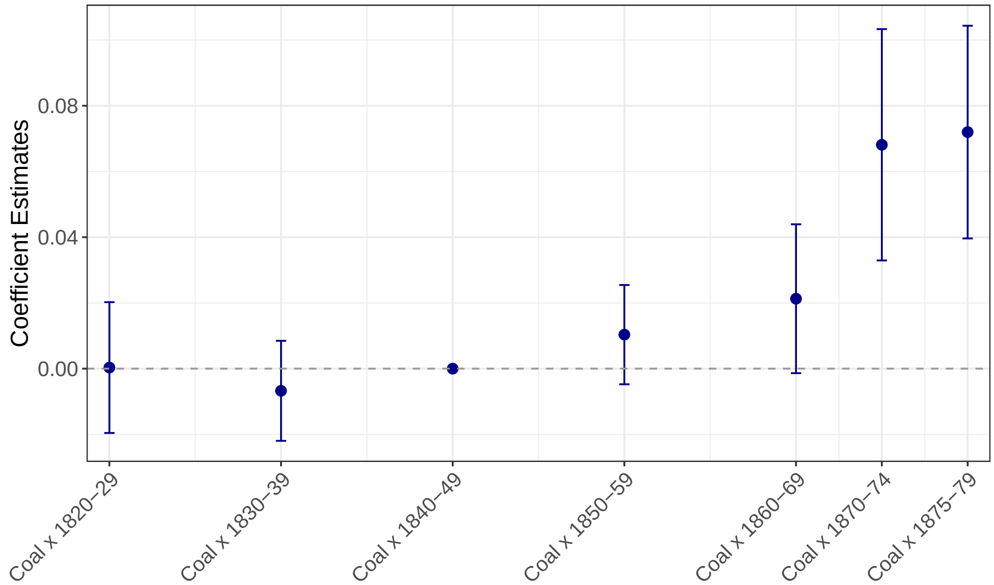
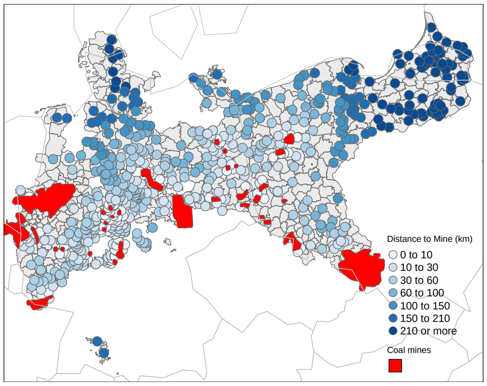


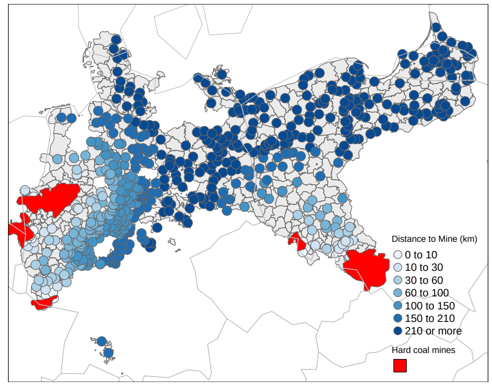
Figure A8: Balanced sample (event study results)

Notes: Figure plots β_t coefficients estimated from Equation 2 with 95% confidence intervals using a balanced subsample including all decades from 1820 to 1879. The omitted period is 1840–1849. The dependent variable measures average wages of male seasonal forestry workers in *Mark* during a decade until 1879. The main explanatory variables measure the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with time dummies. Results are conditional a full set of geographic controls (slope, caloric suitability, distance to rivers, and distance to coast).

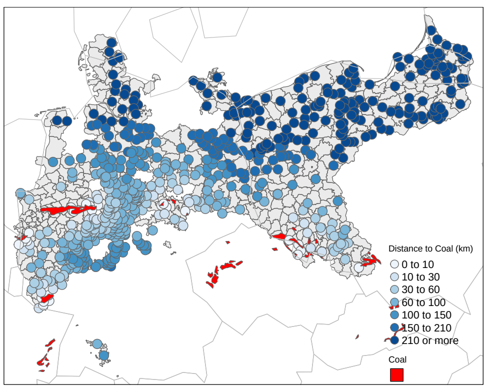
A.2.3 Sensitivity to different measures of coal resources



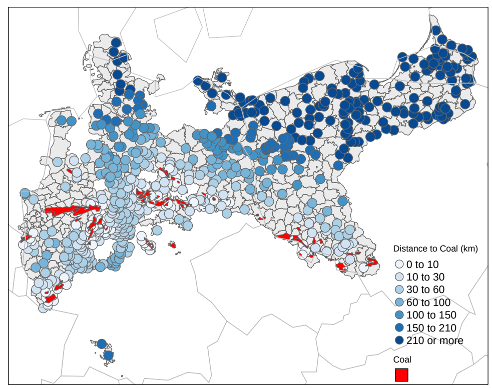
(a) Coal mines in Prussia



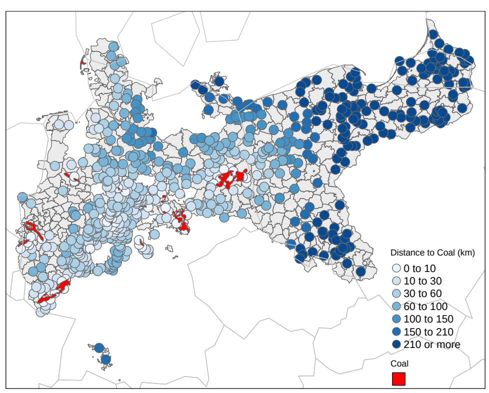
(b) Hard coal mines in Prussia



(c) Incl. Carbon. outside Prussia



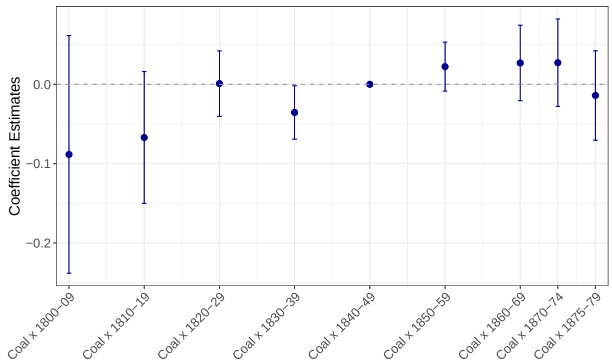
(d) Wider Carbon. period (354–250)



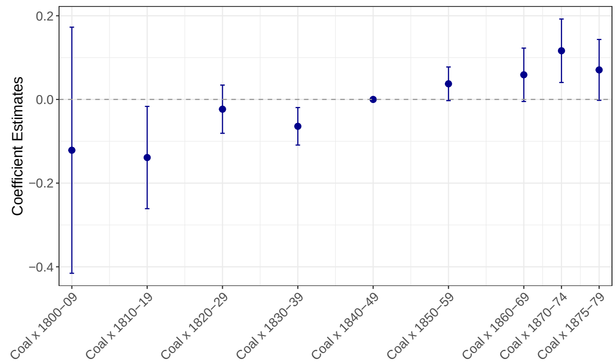
(e) Most prevalent rock types

Figure A9: Alternative versions of locating coal deposits

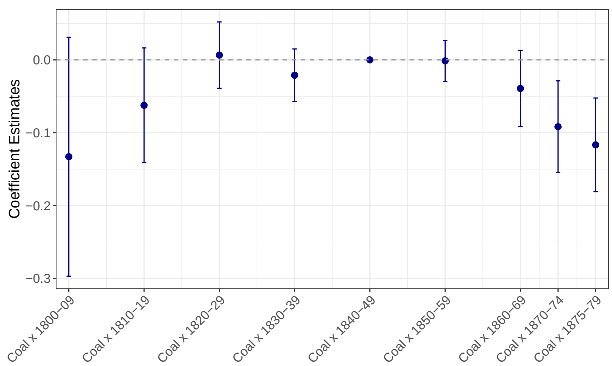
Notes: The maps show the location of state forestries (blue circles) and coal deposits (red polygons). Figure A9a shows locations of coal fields based on Fernihough and O'Rourke (2021), whereas A9b shows only hard coal fields. A9c shows locations of Carboniferous rock strata, including those outside of Prussia; A9d uses a wide definition of the Carboniferous period from 354 to 292 million years ago; A9e defines hard coal not by rock age but by the main rock type classified.



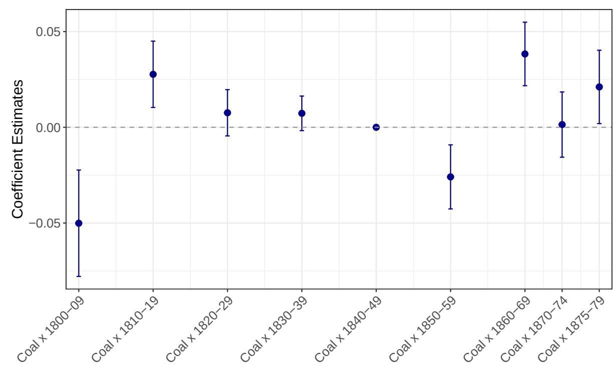
(a) Wood price



(b) Timber price



(c) Firewood price



(d) Wheat price

Figure A10: Coal-driven industrialization and consumer prices (event study results)

Notes: Figures plot β_t coefficients from Equation 2, with 95% confidence intervals. The omitted period is 1840–1849. The dependent variables measure the average price of consumer goods in *Mark* per decade. The main explanatory variable is the (standardized) negative distance to coal deposits from the Carboniferous era in (log) kilometers, interacted with time indicators. All specifications include geographic controls (slope, caloric suitability, distance to rivers, and distance to the coast). The dependent variable in Figure A10a is an index of five wood types (conifer, beech, oak, pine, and spruce); A10b restricts this index to timber (conifer and beech); and A10c to firewood (oak, pine, and spruce). A10d uses the average wheat price in the nearest market. Standard errors are clustered at the forestry level.

A.2.4 Descriptive statistics for household account data

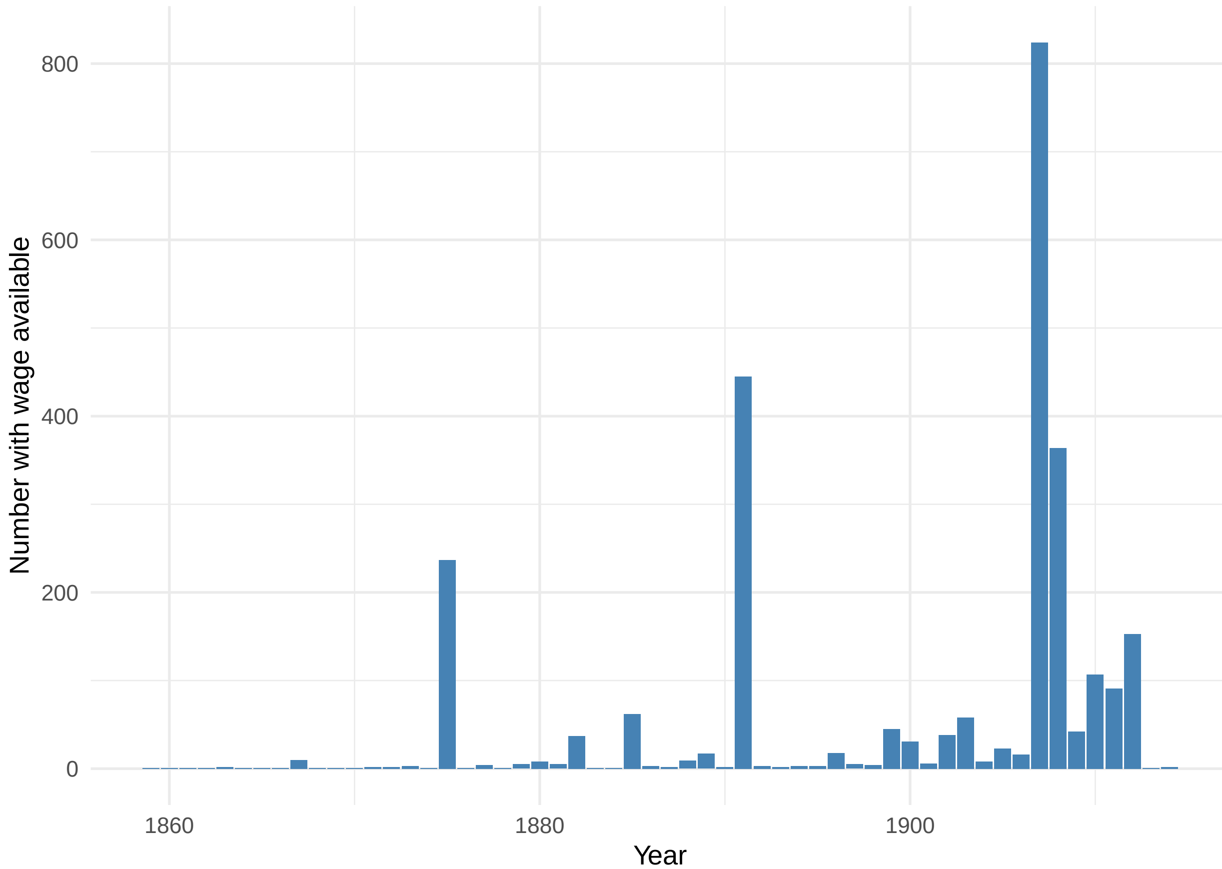


Figure A11: Frequency of male income data observations by year.

Notes: Histogram indicates the number of available observations with male annual income data in the household accounts dataset from H. K. Fischer (2011) by year in the German Empire.

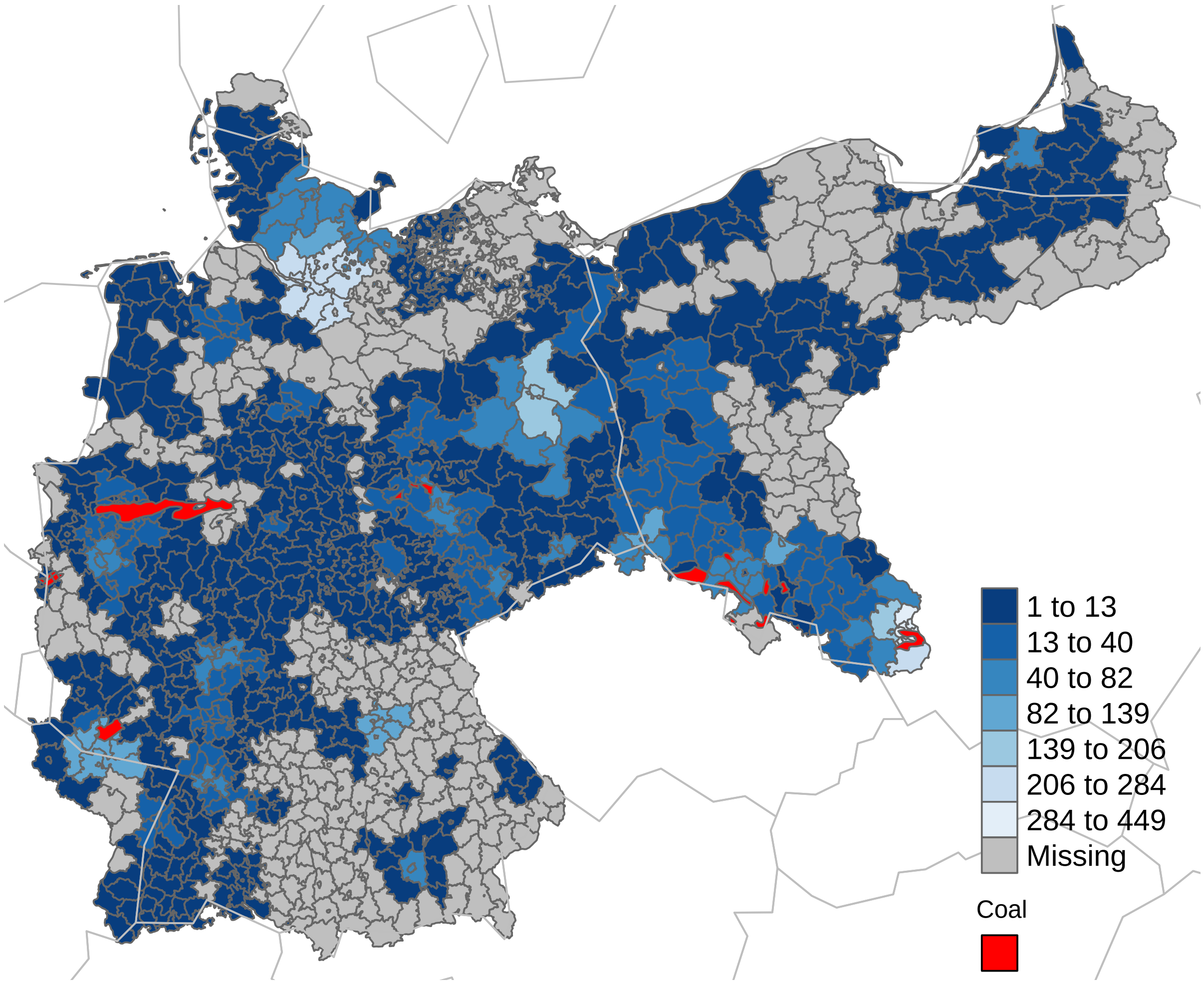
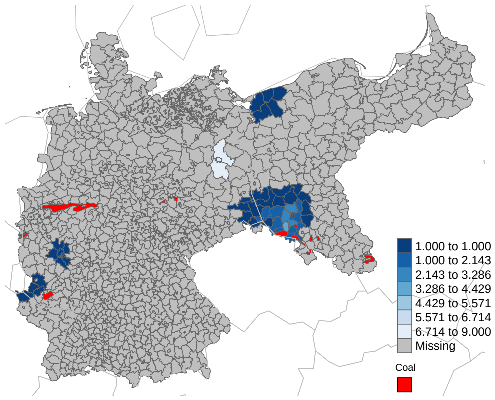
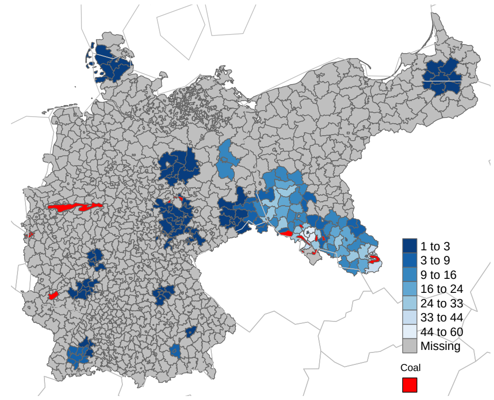


Figure A12: Spatial coverage of male income data observations by county.

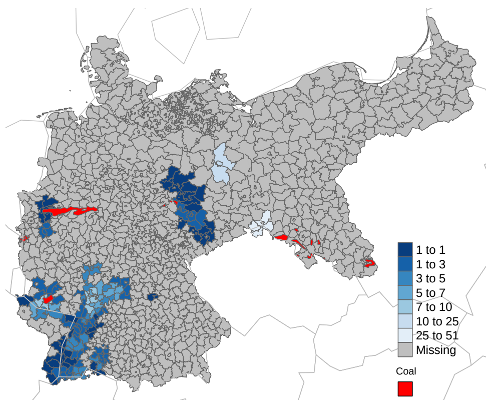
Notes: The map indicates the number of available observations with male annual income data in the household accounts dataset from H. K. Fischer (2011) across counties in the German Empire.



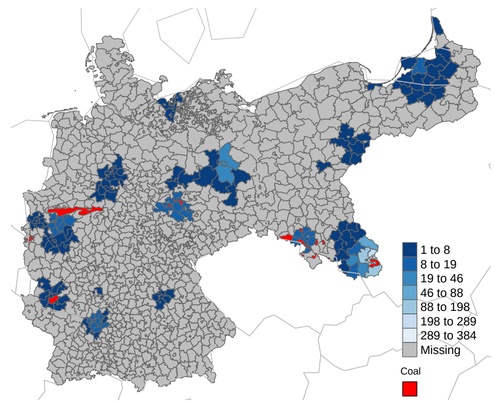
(a) 1861–1870



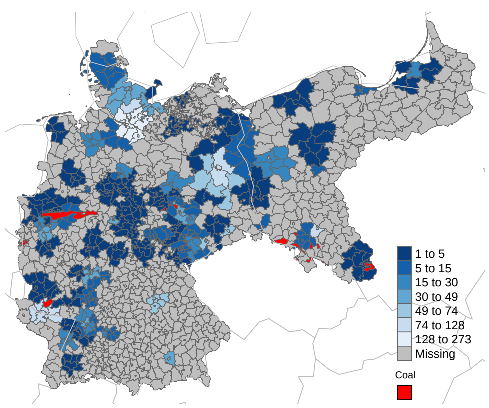
(b) 1871–1880



(c) 1881–1890



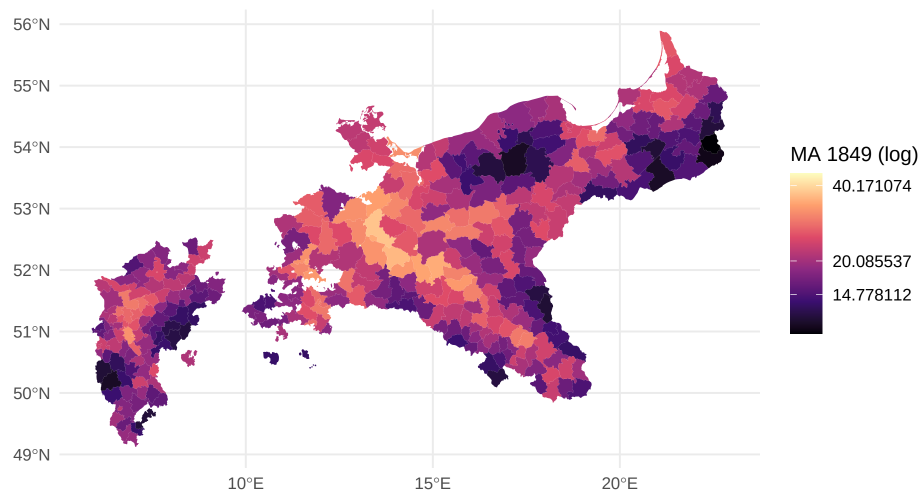
(d) 1891–1900



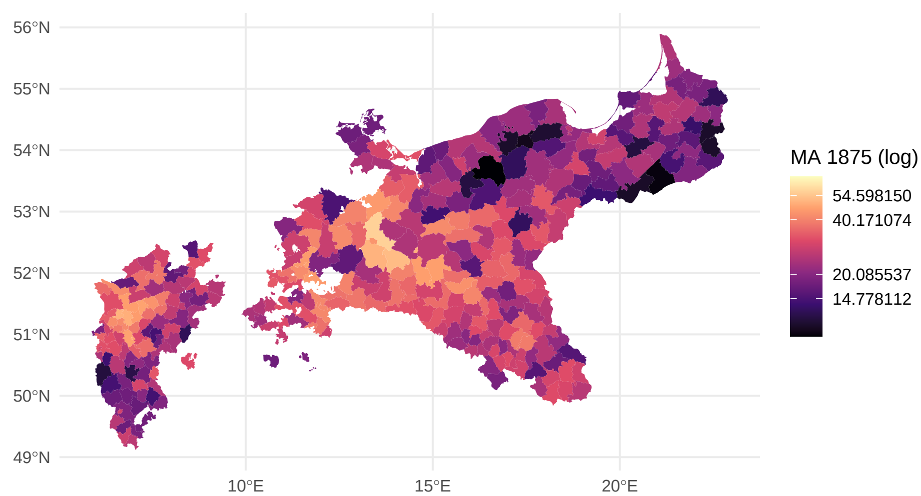
(e) 1901-1910

Figure A13: Spatial coverage of male income data observations by county and decade.

Notes: The maps indicate the number of available observations with male annual income data in the household accounts dataset from H. K. Fischer (2011) across counties in the German Empire by decade as indicated in the figure captions.



(a) Market access 1849



(b) Market access 1875

Figure A14: Market access 1849 and 1875

Notes: The maps show regional variation in market access across Prussian counties in 1849 and 1875

.

A.3 Variable definitions and data sources

A.3.1 Outcome variables

Unskilled wages 1800–1879: Average daily wage in *Mark* of unskilled male 'seasonal fill' workers (day laborers) employed in public state forestries during the time periods 1800–1809, 1810–1819, 1820–1829, 1830–1839, 1840–1849, 1850–1859, 1860–1869, 1870–1874, and 1875–1879. Data reported in U. Eggert (1883), transcribed for this project.

Unskilled wages 1883–1914: Average daily wages in *Mark* for manual day laborers in a county. Data reported in Schmitz, 1888 for the year 1883 and in Zentralblatt, 1892–1914 for the years 1892 to 1914. Data transcribed for this project, except for 1892 and 1901, which were taken from S. O. Becker et al. (2014).

A.3.2 Explanatory variables

Carboniferous rock strata (continuous): The distance in (log) kilometers from a forestry or county's centroid to the nearest rock strata formed during the carboniferous age within Prussia. We use dating most closely resembling the Pennsylvanian period from 320 to 292 million years ago. The variable is truncated at zero, i.e. at the point where a forestry or county's centroid lies directly on carboniferous rock strata. Data source: Asch and Bellenberg, 2005.

Carboniferous rock strata (indicator): Binary variable that assumes a value of one for forestries or counties that had access to coal deposits within Prussia, i.e., those in which rock strata formed during the carboniferous age (Pennsylvanian period from 320 to 292 million years ago) is located within 30 kilometers of the forestry or county's geodesic centroid, zero otherwise. Data source: Asch and Bellenberg, 2005.

Coalfield (continuous): The distance in (log) kilometers from a forest administration or county's centroid to the nearest hard coalfield using geospatial data on the location of major nineteenth-century coal fields in Europe from (Châtel & Dollfus, 1931) and digitized by Fernihough and O'Rourke, 2021. We manually classified each site based on historical sources of mine locations, retaining only hard coal fields. The variable is truncated at zero, i.e. at the point where a forestry or county's centroid lies directly on coalfields.

A.3.3 Consumer prices

Wood price 1800–1879: The average price of wood (conifer, beech, oak, pine and spruce) in *Mark* charged in a public state forestry during the time periods 1800–1809, 1810–1819, 1820–1829, 1830–1839, 1840–1849, 1850–1859, 1860–1869, 1870–1874, and 1875–1879. Data reported in U.

Eggert (1883), transcribed for this project.

Firewood price 1800–1879: The average price of firewood (conifer and beech) in *Mark* per stacked cubic meter (*Raummeter*) charged in a public state forestry during the time periods 1800–1809, 1810–1819, 1820–1829, 1830–1839, 1840–1849, 1850–1859, 1860–1869, 1870–1874, and 1875–1879. Data reported in U. Eggert (1883), transcribed for this project.

Timber price 1800–1879: The average price of timber (oak, pine and spruce) in *Mark* per solid cubic meter (*Festmeter*) charged in a public state forestry during the time periods 1800–1809, 1810–1819, 1820–1829, 1830–1839, 1840–1849, 1850–1859, 1860–1869, 1870–1874, and 1875–1879. Data reported in U. Eggert (1883), transcribed for this project.

Wheat price 1800–1879: The average annual price of wheat in the nearest market to the forestries administration or the county's centroid, respectively. Annual data averaged over time periods relevant for wage data: 1800–1809, 1810–1819, 1820–1829, 1830–1839, 1840–1849, 1850–1859, 1860–1869, 1870–1874, and 1875–1879. Data on wheat prices are available for 54 markets in Prussia in the period 1800 to 1879 from (Federico et al., 2021).

Wheat price 1883–1910: The average monthly price of wheat in the nearest market to the forestries administration or the county's centroid, respectively. Monthly data averaged over time periods relevant for wage data: 1883, 1892, 1897, 1901, 1905, 1910. Data on wheat prices for ca. 165 markets in Prussia in the period 1883 to 1910. Data reported in Königlich Preussisches Statistisches Bureau (1861–1914, vols. 24, 33, 38, 42, 46, 51), transcribed for this project.

Consumer price index 1837–1860: 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*. Data reported in Meitzen (1868, vol. 4), obtained from Cinnirella and Hornung (2016).

A.3.4 Geographical controls

Slope: Average slope of the terrain (in degrees) within a 1.5 km buffer around each forestry or within the county borders, derived from geospatial elevation data at a resolution of 30 arc seconds (approximately 1 km $\times$ 1 km). Slope is calculated using the eight nearest neighboring cells, based on the WorldClim version 2.1 dataset (Fick & Hijmans, 2017). This measure captures local terrain steepness, with higher values indicating steeper gradients.

Distance to coast: The distance in kilometers from a forest administration or county's centroid to the nearest coast, constructed using a map of all European coastlines, provided by the European Environment Agency (EEA).

Distance to navigable river: The distance in kilometers from a forest administration or county's centroid to the nearest navigable river, constructed using a map of all waterways (*Schiffahrtsstraßen*) in the German Customs Union (*Zollverein*) in 1850, provided by the Leibniz Institute of European History at the University of Mainz (IEG, 2010).

Caloric suitability index: The average caloric suitability at within XY km of a forest administration or within a county, as constructed following (Galor & Özak, 2016). This reflects the potential agricultural production (measured in calories) across $5' \times 5'$ grid cells, taking into account the crops available for cultivation.

A.3.5 Household accounts data

Annual labor income: The (log) annual income from labor of adult male household members. Based on individual-level data from approximately 3,000 household accounts in the German Empire (1859–1914), compiled in a meta-study by H. K. Fischer (2011).

Low-skilled sample: The (log) annual income from labor of adult male household members employed in low-skilled occupations, classified according to HISCLASS codes 5, 9, 10, 11, and 12. Classification based on the coding provided by H. K. Fischer (2011).

High-skilled sample: The (log) annual income from labor of adult male household members employed in high-skilled occupations, defined by HISCLASS codes 1, 2, 3, 4, 6, 7, and 8. Classification based on the coding provided by H. K. Fischer (2011).

Mechanics sample: The (log) annual income from labor of adult male household members working in mechanically trained occupations. Occupational classification is based on Feldman and Van der Beek (2016, Table A.1) and adapted with the assistance of ChatGPT-4o.

Non-routine=1: The (log) annual income from labor of adult male household members performing non-routine tasks within mechanically trained occupations. Classification follows Feldman and Van der Beek (2016, Table A.1), with coding support provided by ChatGPT-4o.

A.3.6 Mediators

Steam engines p.c. 1849: The total number of horsepower from steam engines used in the manufacturing sector in 1849, divided by the county's total population. Constructed using county-level data reported in Statistisches Bureau zu Berlin (1851–1855, vol. 6a), obtained from S. O. Becker et al. (2014).

Steam engines p.c. 1875: The total number of horsepower from steam engines used in the manufacturing sector in 1875, divided by the county's total population. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 40), transcribed for this project.

Firm size 1849: The total number of employees across all factories in the manufacturing sector in 1849, divided by the number of all factories. Constructed using county-level data reported in Statistisches Bureau zu Berlin (1851–1855, vol. 6a), obtained from S. O. Becker et al. (2014).

Firm size 1875: The total number of employees across firms in the manufacturing sector with more than five employees (*Grossbetriebe*) in 1875, divided by the total number of such factories. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 40), transcribed for this project.

Market access 1849 and 1875:
The population of all counties in 1849 and 1871, weighted by trade costs along the least-cost transport paths between county i and each county j $j \neq i$ at time $t \in \{1849, 1875\}$. Calculated according to Equation 6:

$$\text{MA}_{it} = \sum_{\substack{j=1 \\ j \neq i}}^{N} \frac{\text{POP}_{jt}}{(\text{COST}_{ijt})^d}. \tag{6}$$

Here, MA_{it} denotes the market access of county i in period t, N is the total number of counties, and POP_{jt} is the population of county j at time t. COST_{ijt} represents the least-cost path distance (in kilometers) between counties i and j, based on the prevailing transport costs in year t. The distance decay parameter d, which governs how quickly access declines with distance, is set to $d = 1$, as done by Crafts (2005) and Johnson and Koyama (2017).
Least-cost distances are computed between all pairs of county centroids using the `gdistance` package in R, which implements a transition-cost model. The algorithm evaluates the cumulative cost of movement across a raster surface by selecting the lowest-cost path from each cell to its eight neighboring cells. Travel frictions are assigned across a 1 km $\times$ 1 km raster grid, with uniform portage costs in areas without infrastructure. Frictions are reduced along roads, railways, rivers, and coastlines, based on relative transport cost estimates from Gutberlet (2012), summarized in Table A17.[39]
Figure A14 presents maps showing regional variation in market access across Prussian counties in 1849 and 1875.

[39] Each road, railway, and river segment is buffered by 150 m. Any raster cell intersecting a buffered transport line is assigned the corresponding mode's transport cost. Transfers between land and sea are permitted only at designated ports (Memel, Königsberg, Elbing, Danzig, Stettin, Greifswald, Altona, Emden, Bremen, and Lübeck), each defined by a 5 km port gate.

Table A17: Freight rates by modes of transportation and year

Mode	1846	1875
Sea	1.61	1.21
River	4.50	1.55
Rail	11.20	3.42
Road	40.00	40.00
Portage	50.00	50.00

Notes: National average freight rates for different modes of transportation and transshipment measured in *Pfennig* (100 Pfennig = 1 Mark) per ton kilometer. Source: Gutberlet (2012).

Manufacturing employment p.c. 1849: The total number of workers employed in the manufacturing sector in 1849, divided by the total population. Constructed using county-level data reported in Statistisches Bureau zu Berlin (1851–1855, vol. 5, 6a), obtained from S. O. Becker et al. (2014).

Manufacturing employment p.c. 1875: The total number of workers employed in the manufacturing sector in 1875, divided by the total population. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 40), transcribed for this project.

Low-skilled employment p.c. in 1849: The total number of low-skilled workers employed in non-agricultural occupations in 1849, divided by the total population. Occupations are classified according to our own coding using the HISCLASS scheme. We identify low-skilled workers as those in classes 5 (Lower clerical and sales personnel), 9 (Lower-skilled workers), and 11 (Unskilled workers). Constructed using county-level data reported in Statistisches Bureau zu Berlin (1851–1855, vol. 5, 6a), obtained from S. O. Becker et al. (2014).

Low-skilled employment p.c. in 1882: The total number of low-skilled workers employed in non-agricultural occupations in 1882, divided by the total population in 1880. The original sources report three relevant occupational categories of which we classify white-collar or blue-collar workers as low-skilled, whereas we classify self-employed or managers as high-skilled. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 76), obtained from Galloway (2007).

Infant mortality 1849: The total number of deaths of infants under one year of age divided by total live births in 1849. Constructed using county-level data reported in Statistisches Bureau zu

Berlin (1851–1855, vol. 2), obtained from S. O. Becker et al. (2014).

Infant mortality 1875: The total number of deaths of infants under one year of age divided by total live births in 1875. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 42), obtained from Galloway (2007).

Migration rate 1849. The total population growth between 1846 and 1849, net of natural population change, approximated by three times the birth-death differential in 1849. This variable is expressed in per capita terms by dividing by the average county population over the period. Net migration between 1846 and 1849 is calculated as:

$$\text{Migration Rate}_i^{46:49} = \frac{\left(\text{Pop}_{i,1849} - \text{Pop}_{i,1846}\right) - 3 \cdot \left(\text{Births}_{i,1849} - \text{Deaths}_{i,1849}\right)}{\frac{1}{2} \cdot \left(\text{Pop}_{i,1846} + \text{Pop}_{i,1849}\right)}$$

Constructed using county-level data reported by Statistisches Bureau zu Berlin (1851–1855, vol. 2), obtained from S. O. Becker et al. (2014).

Migration rate 1875. The total population growth between 1867 and 1875, net of natural population change due to births and deaths in the same period. This variable is expressed in per capita terms by dividing by the average county population over the period. Net migration between 1867 and 1875 is calculated as:

$$\text{Migration Rate}_i^{67:75} = \frac{\left(\text{Pop}_{i,1875} - \text{Pop}_{i,1867}\right) - \sum_{t=1867}^{1875}\left(\text{Births}_{it} - \text{Deaths}_{it}\right)}{\frac{1}{2} \cdot \left(\text{Pop}_{i,1867} + \text{Pop}_{i,1875}\right)}$$

Constructed using county-level data obtained from Galloway (2007).

A.3.7 Other outcomes

Industrial employment p.c. 1882: The total number of workers employed in the industrial sector (including mining) in 1882, divided by the total population 1880. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 76), obtained from Galloway (2007).

Mining employment p.c. 1882: The total number of workers employed in mining in 1882, divided by the total population 1880. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 76), obtained from Galloway (2007).

Manufacturing employment p.c. 1882: The total number of workers employed in the manufacturing sector in 1882, divided by the total population 1880. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 76), obtained from Galloway (2007).

Agricultural employment p.c. 1882: The total number of workers employed in the agricultural sector in 1882, divided by the total population 1880. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 76), obtained from Galloway (2007).

Urbanization rate 1875: The total number of inhabitants in cities with city rights within a county divided by the county's population in 1875. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 42), obtained from Galloway (2007).
Low-skilled Manufacturing employment p.c. 1882: The total number of low-skilled workers (white-collar or blue-collar workers) employed in the manufacturing sector in 1882, divided by the total population 1880. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 76), obtained from Galloway (2007).

High-skilled Manufacturing employment p.c. 1882: The total number of skilled workers (self-employed or managers) employed in the manufacturing sector in 1882, divided by the total population 1880. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 76), obtained from Galloway (2007).

Low-skilled Manufacturing employment share 1882: The total number of low-skilled workers employed in the manufacturing sector in 1882, divided by the total manufacturing workforce in 1882. Constructed using county-level data reported in Königlich Preussisches Statistisches Bureau (1861–1934, vol. 76), obtained from Galloway (2007).