Industrialization, Returns, Inequality

Thilo N. H. Albers (HU Berlin)
Felix Kersting (HU Berlin)
Timo Stieglitz (HU Berlin)

Discussion Paper No. 462

November 22, 2023

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Georg Weizsäcker, Ph.D., Humboldt University Berlin, 10117 Berlin, Germany
info@rationality-and-competition.de

Industrialization, returns, inequality*

Thilo N. H. Albers[†] Felix Kersting[‡] Timo Stieglitz[§]

November 22, 2023

Abstract

How does revolutionary technological change impact wealth inequality? We turn to the mother of all technological shocks–the Industrial Revolution–and analyze its role for wealth concentration both empirically and theoretically. Based on a novel dataset on wealth shares at the level of Prussian counties, we provide causal evidence on the positive effect of industrialization on the top percentile's wealth share and the inequality among top fortunes. We show that this relationship between industrialization, wealth concentration, and tail fattening is consistent with both cross-country data on national wealth distributions and with a new individual-level dataset of Prussian millionaires. We disentangle the mechanisms underlying the observed wealth concentration and tail fattening by introducing a dynamic two-sector structure into an overlapping generations model with heterogeneous returns to capital. In particular, we study the role of sector-specific scale dependence, i.e. the positive correlation of rates of return and wealth in industry, and dynastic type dependence in returns, i.e., the gradual one-directional transition of wealth-holders from the low-return traditional to the high-return industrial sector. The simulations suggest that the combination of these two features explains about half of the total increase of the top-1% share, while the other half resulted from the general increase and higher dispersion of returns induced by the emerging industrial sector.

Keywords: Rates of return, wealth inequality, industrialization, technology, simulation.

JEL classification: D31, E21, N13, O14.

*We are grateful for advice and comments to Jess Benhabib, Alberto Bisin, James Fenske, Martin Holm, Kalle Kappner, Ulrich Pfister, Thomas Piketty, Monique Reiske, Felix Schaff, Eric Schneider, Tamas Vonyo, Jacob Weisdorf, Nikolaus Wolf; participants at the World Economic History Congress (Paris), the Congress for Social and Economic History (Vienna and Leipzig), the European Historical Economics Society Conference (Groningen), CRC TRR 190 meeting (Ohlstadt), the workshop "Growth, History and Development" (Odense), the summer school "Political Economy of Inequality" (Duisburg), the conference "Big counterfactuals of Macro-Political history" (Manchester), and the workshop "Unveiling Wealth and Income Inequalities" (Berlin); seminar participants in Berlin, Halle, London (LSE), Madrid, Mannheim, and Münster. Berta Llugany i Costa and Can Aycan provided excellent research assistance. Financial support by Deutsche Forschungsgemeinschaft through CRC TRR 190 (project number 280092119) is gratefully acknowledged.

[†]Humboldt University Berlin; thilo.nils.hendrik.albers@hu-berlin.de.

[‡]Humboldt University Berlin; f.kersting@hu-berlin.de

[§]Humboldt University Berlin; stiglitt@hu-berlin.de

1 Introduction

The effects of fundamental technological change on society have been subject to much public debate from the beginning of modern economic growth until today (Marx, 1867; Piketty, 2014; Acemoglu and Johnson, 2023). A repeatedly raised concern is that the arrival of new general purpose technologies may lead to a strong increase in inequality. On the one hand, they may substitute human labor and drive down wages, increasing wage inequality. On the other hand, wealth inequality may increase as those capitalists owning new technologies make extraordinary profits. However, as truly revolutionary technological change is a rare occurrence, our knowledge about its impact on wealth distributions remains limited. This study turns to the mother of all technological shocks–the Industrial Revolution–to shed light on this relationship. At the end of the Industrial Revolution, wealth concentration was at its historical peak. To what degree was the advent of new technology responsible for this peak? What were the precise mechanisms that led to this unparalleled concentration of wealth?

We confront these questions theoretically and empirically using Germany as our historical case. In particular, we make two contributions. First, levying novel regional wealth inequality data and employing an instrumental variable approach, we document the causal effect of industrialization on the top percentile's wealth share as well as the fatness of the wealth distribution's tail, i.e., the inequality among top fortunes. Second, we extend an overlapping-generations (OLG) model with heterogeneous returns by implementing a two-sector structure. This allows us to incorporate two features of industrialization: *dynastic type* dependence of returns, i.e., the one-directional transition of wealth-holders from the low-return traditional to the high-return industrial sector, and *scale* dependence, i.e., the positive correlation between rate of return and wealth to reflect the existence of economies of scale in the industrial sector. Under a realistic parametrization, our simulation allows us to replicate the observed changes in the wealth distribution. Based on a counterfactual analysis, we attribute almost half of the total simulated increase of 28pp in the top-1% share to the co-existence of scale and dynastic type dependence of capital returns. The other half resulted from the general increase and higher dispersion of returns induced by the emerging industrial sector. These results underscore the importance of understanding technological revolutions not only as a shock to the remuneration of labor, but also to that of capital. To this end, we also show that already relatively modest levels of capital return taxation, if implemented from the start of the Industrial Revolution, would have substantially slowed down wealth concentration.

Estimating the causal effect of industrialization on wealth inequality requires, first, variation in wealth inequality and, second, exogenous variation in industrialization. We overcome the identification challenge by using within-country variation in the outcome and independent variable of interest. In particular, we estimate sub-national top shares for 575 counties in Prussia based on newly digitized tax data for 1914. Throughout, we ensure that our county-level measures of wealth inequality are consistent with the national aggregates and modern concepts of national wealth accounting (as described by Piketty and Zucman, 2014). To establish a causal relationship between industrialization and wealth inequality, we use the distance to carboniferous strata as an instrumental variable for the level of industrialization as measured by the share of industrial employment (following De Pleijt et al. 2020 and Fernihough and O'Rourke 2021). In our preferred specification, a one standard deviation increase in industrial employment leads to a shift of more than 7 percentage points in the top-1% wealth share. We also document a positive relationship between industrialization and the fattening of the wealth distribution's tail.

These results from subnational data are consistent with evidence from national wealth distributions. Cross-country evidence suggests that industrialization was associated with an increase in the top percentile's wealth share and tail fattening. Using a novel decomposition, we show that indeed the former mainly originated in the latter. Put differently, much of the increase in the top percentile's wealth share was driven by the polarization at the top rather than the proportional increase of the concentration at the top relative to the bottom 90%.

The observed tail fattening is suggestive of a change in the composition of the economic elite as well as their ability to accumulate wealth. Using a novel individual-level dataset of Prussian millionaires detailing their income, wealth, and the source of their fortune, we document four stylized facts about the economic elite at the eve of World War I. First, industrialization led indeed to a replacement of the traditional elite. Among the richest 500 Prussians in 1908, around 35% were industrial entrepreneurs, whereas the once-dominant agricultural elite only made up around 20%. Second, returns between old and new wealth differed. While entrepreneurs earned a mean return of more than 6% on their capital, the landed agricultural elite earned a mere 4%. Additionally, industrial returns exhibited a higher variance. Third, industrialization implied the advent of economies of scale. Consequently, returns were scale dependent in industry, but not in agriculture. Fourth, we observe a bifurcation of the top tail: The tail for entrepreneurial fortunes was significantly fatter than that for agricultural ones.

Using the insights and parameters from the millionaires dataset, we then employ an overlapping generations model with heterogeneous returns to dissect the drivers of wealth concentration. In particular, we extend a standard model of this class (Benhabib et al., 2011) by implementing a two-sector structure. This allows for modeling two important aspects of structural change associated with the Industrial Revolution. First, dynastic type dependence—the irreversibility of the move of a dynasty's wealth to the new industrial sector—reflects the historically one-directional but relatively slow transition from the traditional to the new industrial sector. This feature amplifies the persistence of the heterogeneity of returns. Second, sector-specific scale dependence—the positive correlation between wealth and the rate of return on wealth—reflects the emergence of economies of scale in the new industrial sector in later stages of industrialization.[1] To gauge the relevance of dynastic type dependence, in particular in connection with scale dependence, we simulate alternative counterfactual scenarios lacking one or both of the features. We calibrate the model on the preindustrial distribution and shock it with the industrial returns observed in the millionaires data, hence not targeting the levels of wealth inequality attained at the eve of World War I.

The simulations suggest that only the combination and interaction of scale and dynastic type dependence of capital returns replicates the wealth concentration observed in the 19th century. Adding either dynastic type or scale dependence can only explain two thirds of the increase and turning off both features while allowing for the new higher industrial returns only explains half of it. Put simply, only half of the increase in wealth concentration is explained by the expanding availability of industrial rates of return per se, while the other half is explained by the specific nature of the capital return process. Additionally, we show that scale and dynastic type dependence underlie the emergence of modern fat tails, and that dynastic type dependence is responsible for the bifurcation between holders of old and new wealth observed in actual data. In a final set of simulations, we explore whether taxes can reign in the distributional consequences of Industrial Revolution-type technological change. Our results indicate that modest capital return taxes, if implemented at the start of the technological shock, can have large effects as they compress the variance of the return process. The common and generalizable insight from all simulations is that, next to the level and dispersion of capital returns, the distinct features of return processes are of first order importance for wealth concentration.

Our paper speaks to three strands of literature. First, we contribute to the literature on the long-run development of wealth inequality. In the tradition of the work by Piketty et al. (2006) on France, there now exist long-run series on wealth concentration for multiple countries stretching back to the 19th or early 20th century (Roine and Waldenström, 2009; Saez and Zucman, 2016; Alvaredo et al., 2018; Bengtsson et al., 2018, 2019; Albers et al., 2022). Additionally, another strand following Lindert (1986) and Van Zanden (1995) continues to produce new estimates of inequality in the 18th century and earlier (surveyed by Alfani, 2021). With the notable exception of Dray et al. (2023), long-run studies are typically descriptive and provide one series for a whole country over multiple centuries. Given the nature of these data—one data point per country per year—it is difficult to make causal statements and quantify the effect of industrialization or, in fact, any other drivers of wealth inequality. Relative to the existing literature, we highlight a new empirical avenue for the identification and quantification of drivers of wealth inequality by leveraging sub-national data. Additionally, through our theoretical framework, we define conditions that led to the

[1] Others have implemented scale dependence into this type of model and stressed its importance (e.g., Gabaix et al., 2016). As an innovation, we implement a sector-specific formulation of scale dependence.

unparalleled concentration of wealth at the eve of World War I—a central question with respect to the future evolution of wealth inequality (Piketty, 2015, p. 49).

Second, we complement a lively literature exploring the distributional consequences of the Industrial Revolution that has focused on the wage, functional, and personal income distributions. Motivated by the deskilling debate (Acemoglu, 2002) and the role of human capital in unified growth theory (Galor and Moav, 2004, 2006), much of the existing empirical work circles around the question of wage inequality, i.e., whether the Industrial Revolution was skill-saving or skill-demanding (Atack et al., 2004; De Pleijt et al., 2020; Franck and Galor, 2022; Ridolfi et al., 2022). Theoretical and conceptional work focuses on the impact of rural-urban migration on the functional and personal income distribution (Lewis, 1954; Kuznets, 1955; Allen, 2009), political economy implications (Bartels et al., forthcoming), and the impact of human capital on the distribution (Galor and Moav, 2004, 2006). We add to this literature by focusing on wealth rather than income inequality and, in particular, the role of heterogeneous returns as the main determinant of 19th century wealth concentration.

Finally, we contribute to the theoretical macroeconomic literature on wealth inequality by introducing a sectoral structure and a new channel. Recent models have overcome the difficulty to explain the consistently higher wealth inequality with features of the income distribution by turning to the rate of return channel as the main determinant of wealth concentration at the top (Benhabib et al., 2011, 2019, 2021; Piketty and Zucman, 2015; Gabaix et al., 2016; Hubmer et al., 2021; Moll et al., 2022; Gaillard and Wangner, 2022). While our model's building blocks are from Benhabib et al. (2011), the focus of our paper is most closely related to the work by Moll et al. (2022) on automation and returns. We build on these models by adding a two-sector structure that differs from a portfolio approach as our agents exclusively hold their wealth in one of these sectors. The mean and dispersion of rates of return differ across sectors as does the scale dependence of returns. This allows us to introduce dynastic type dependence as the one-directional and gradual transition from one sector to the other. It differs substantially from and complements the 'classical' modeling of type dependence, in which households draw a fixed rate at birth. While industrialization in the 19th century is our case in point, it is conceivable that this channel matters in other time periods. As farms in developing countries or businesses in now-developed countries remain family-run over many generations, wealth is slow to be transferred to emerging high-risk high-return sectors. According to the dynastic type dependence channel, this affects the shape of the wealth distribution, in particular at the top.

The remainder of the paper is organized as follows. Section 2 presents the empirical evidence. Section 3 explores the mechanisms through the lenses of a dynastic wealth model. Section 4 concludes.

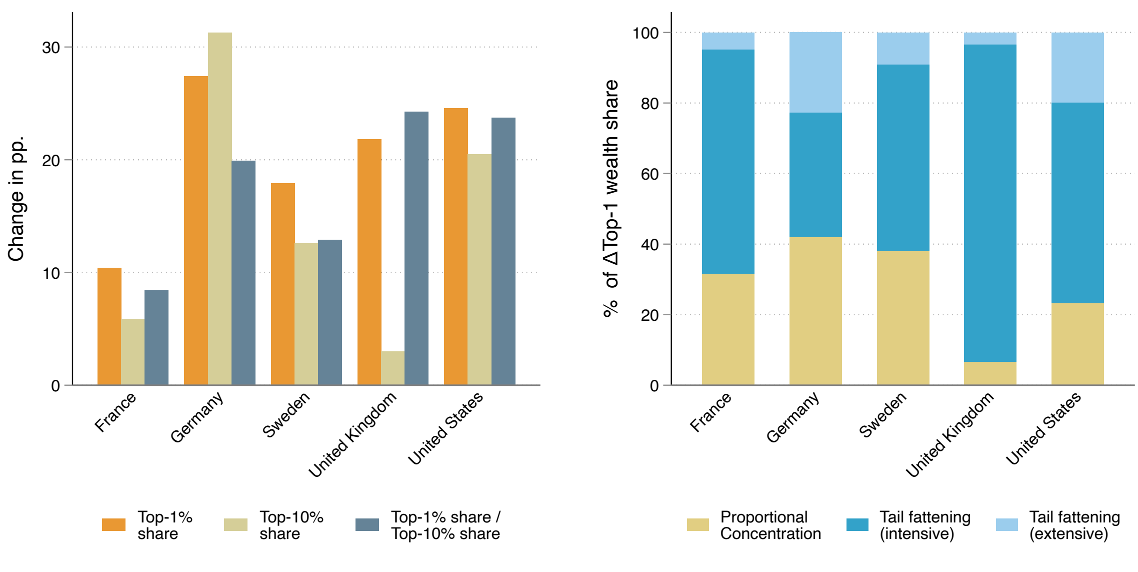
2 Empirical evidence from the country to the entrepreneur

Leveraging three kinds of data, this section draws out the relationship between the Industrial Revolution and wealth inequality. First, we document that industrialization was associated with an increase in wealth inequality across countries and that this increase originated in the polarization among the wealthy, i.e., the fattening of the wealth distribution's tail (Section 2.1). Second, we explore the implied transformation of the economic elite through the lens of a new individual-level millionaires dataset, showing that the new elite was comprised of industrialists (Section 2.2). Third, using subnational data, we provide causal evidence on the effect of industrialization on tail fattening and wealth inequality (Section 2.3).

2.1 Cross-country evidence on tail fattening during industrialization

As the most common analytical unit for the study of inequality is the nation, it is useful to draw out and characterize common trends across countries. Due to the paucity of pre-industrial estimates, the sample is necessarily small and the data quality varies across countries (see Appendix A.1 for a discussion). These caveats notwithstanding, they produce two consistent stylized facts about the association of industrialization and wealth inequality.

Figure 1: The Industrial Revolution and wealth concentration



(a) Industrialization and changing top shares (b) Fattening tails of the wealth distribution

Notes: The graph shows changes between a pre-industrial estimate and an estimate close to 1900. The precise reference years across countries vary subject to data availability, quality, and consistency. Appendix A.1 provides details and discusses the quality of the sources.

The first insight pertains to the change in the top percentile's wealth share: While pre-industrial top-1% wealth shares varied between 17% (US) and 44% (France) across the sample of countries, the orange bars in Figure 1(a) document a large increase in the top-1% share during the period of industrialization across all of them. At the eve of World War I, Germany and the United States had caught up with and surpassed, respectively, British aggregate productivity levels (Broadberry, 1998; Broadberry and Burhop, 2007). All three countries display changes of more than 20pp since the beginning of industrialization. France and Sweden lagged behind in industrialization at this point (Broadberry et al., 2010) and the respective increases in the top-1% shares were more modest: 10pp and 18pp. In sum, industrialization was associated with a surge in wealth inequality and more industrialized countries had experienced a larger increase in the top percentile's wealth share.

The second insight pertains to the nature of this shift in wealth concentration. It is well known that the distribution of households at the top of the wealth distribution follows the Pareto distribution. As long as the distribution's central parameter on tail fatness–the Pareto α–stays constant, we would expect that the increase in the top decile's wealth share is proportional to the increase of the share owned by the top percentile. Yet, the khaki bars in Figure 1(a) indicate that the share increases are not even equal in absolute terms for all but one country (Germany). The blue bars confirm that the share of the top-1% grew disproportionately to that of the 90th-99th percentiles for all countries, leading to an increase in the expression $\frac{\text{Top-1\% share}}{\text{Top-10\% share}}$. As the increase in this share-in-share is directly linked to a decrease in the Pareto exponent α (Jones and Kim, 2018), it indicates tail fattening. Put differently, a part of the increase in the top percentile's wealth share came at the expense of the share of the 90-99th percentiles, not of the share of the bottom 90%.

The following decomposition (see Appendix A.1 for the derivation) solidifies this impression in a more structural manner by exploiting the mechanical relationship between changes in top shares ($\Delta s^{Top_1} = s_1^{\text{Top 1}} - s_0^{\text{Top 1}}$ and $\Delta s^{\text{Top 10}} = s_1^{\text{Top 10}} - s_0^{\text{Top 10}}$) and shares-in-shares ($\gamma_1 - \gamma_0$, where $\gamma_t = s_t^{\text{Top 1}} / s_t^{\text{Top 10}}$) between two points in time ($t = 0, 1$). In particular, the change in the top percentile's wealth share can be rewritten as the sum of three parts:

$$\Delta s^{\text{Top 1}} = \underbrace{\gamma_0 \Delta s^{\text{Top 10}}}_{\substack{\text{proportional} \\ \text{wealth concentration}}} + \underbrace{(\gamma_1 - \gamma_0) s_0^{\text{Top 10}}}_{\substack{\text{tail fattening} \\ \text{(intensive margin)}}} + \underbrace{(\gamma_1 - \gamma_0) \Delta s^{\text{Top 10}}}_{\substack{\text{tail fattening} \\ \text{(extensive margin)}}}, \tag{1}$$

The first term on the RHS of equation (1) allocates the increase in the top-10% share proportionally to the top-1%, holding tail fatness constant. If there is no tail fattening ($\gamma_1 = \gamma_0$), the RHS collapses to this term. By contrast, the middle term on the RHS isolates the effect of tail fattening at the intensive margin, i.e., assuming a constant top-10% share. If there is no increase in the top-10% share ($\Delta s^{\text{Top 10}} = 0$), the RHS collapses to this term. In this case, the increase in the top-1% share is entirely due to a reshuffling of wealth shares within the top-10%, from the 90th-99th percentiles to the top-1%. Last, the right term reflects the additional effect of tail fattening at the extensive margin, i.e., if the top-10% share increases.

Figure 1(b) reports the results of the decomposition. Across countries, on average 28% of the increase in the top-1% share originates in a proportional increase of wealth concentration, whereas the intensive and extensive margin of tail fattening make up for 60% and 12%, respectively. This implies that the wealth concentration observed during industrialization did not foremost originate in an existing elite becoming richer relative to the rest of society. Instead, a new much smaller elite transpired that could accumulate much more capital relative to other affluent parts of society.

In sum, despite the uncertainty of historical estimates, the assembled cross-country data consistently show two stylized facts across the surveyed countries. First, industrialization was associated with an increase in the top percentile's wealth share. Second, the increase in tail fatness and its large contribution to the change in the top percentile's share reflect a fundamental change in capital accumulation at the top. Who were the rich in 1900 and how had they been able to amass so much more capital than the rich in 1800?

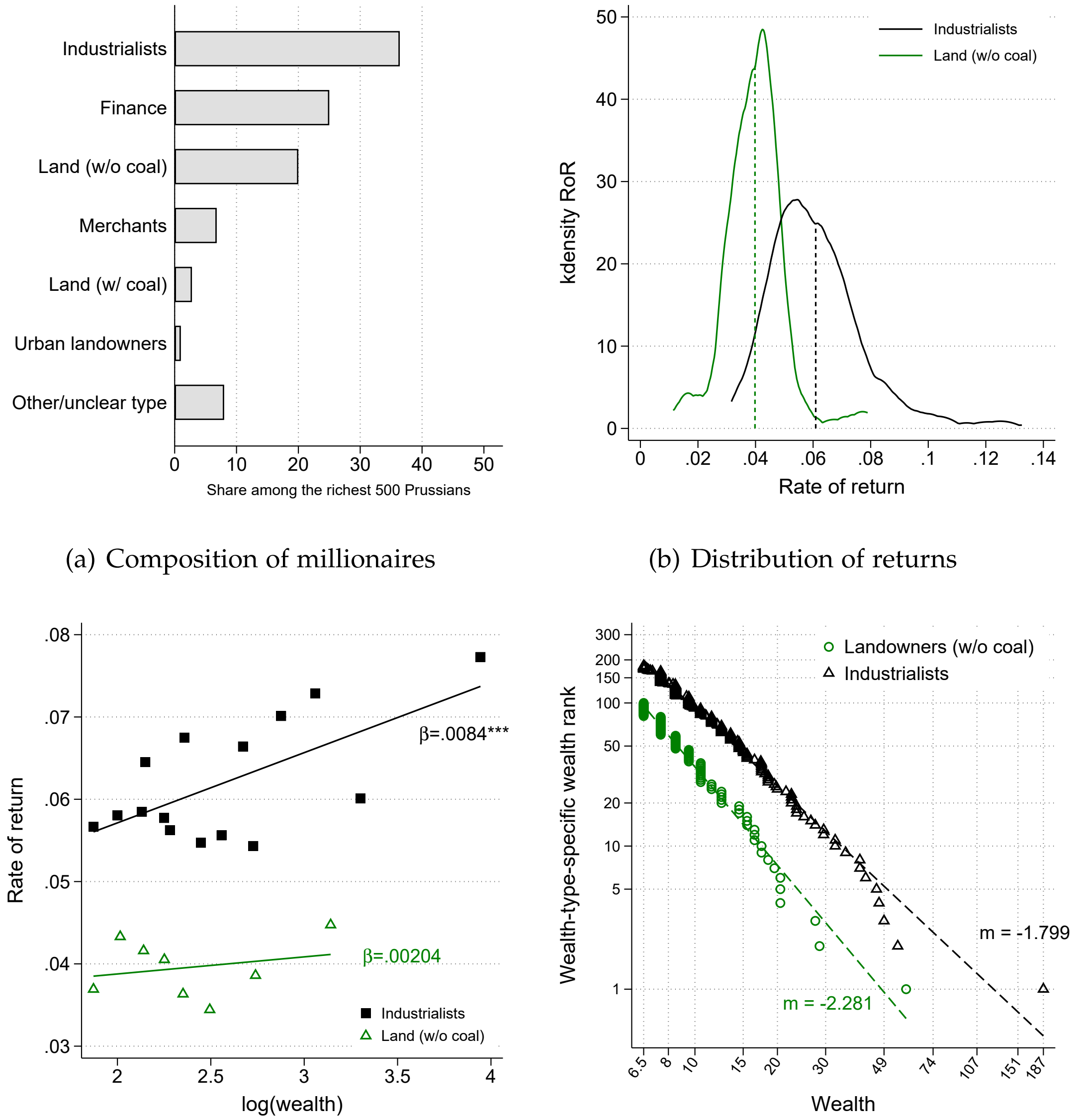
2.2 The returns to industrialization

To confront these questions, we rely on a novel dataset on the universe of Prussian millionaires. Its source is a historical curiosity. The Prussian bureaucrat Rudolf Martin (1912) had strong views on the threat of plutocracy and thought that the size and the ownership of the largest fortunes should be public knowledge. It is almost certain that he had access to the wealth tax returns of the Prussian state when he composed his book of all 8,000 Prussian millionaires detailing their wealth and income—a Forbes list for 1908, but with much better underlying data. For the richest 500 of them, the source also contains biographical data detailed enough to classify the origin of the fortune. This allows us to document the structure of this economic elite as well as the level and properties of the returns to their fortunes.[2]

Figure 2(a) provides the first key fact about this economic elite. By 1908, more than 35% of the richest 500 Prussians were industrialists. These included some well-known names such as Krupp (steel), Siemens (electrical engineering), and Bayer (dye/chemical industry). In contrast, the traditional old money—that is, owners of large agricultural estates that did not happen to have coal deposits on their grounds—only represented a mere 20% (relative to an employment share in agriculture of about 35% in 1907). The only two other significant groups were merchants and financiers, which fall between the dichotomy of old and new money. Often from affluent backgrounds, their owners and founders benefited from the new opportunities opened up by the Industrial Revolution: Serving a new consumption culture (De Vries, 1994) and bankrolling German industry (Tilly 1996, p. 114f and Wehler 2008a, p. 637f). While no comparable millionaires list exists for pre-industrial Germany, it is clear that the old elite looked different: It was composed of patricians in cities (bankers, merchants, owners of urban land) and owners of large agricultural estates outside of them (Wehler, 2008a, chapter 3). Entrepreneurs had replaced a considerable part of this century-lasting old elite by the eve of World War I.

[2]Appendix A.2 provides details on the data as well as a validation comparing it with the actual tabulated statistics of the wealth tax. The only comparable dataset provided by Cummins (2022) applies (homogeneous) macroeconomic returns data to the known portfolio structure.

Figure 2: The structure of the economic elite and the return to capital in 1908



(a) Composition of millionaires

(b) Distribution of returns

(c) Scale dependence in returns

(d) Top tails by fortune type

Notes: The figures are based on the wealth and income of the 500 richest Prussians in 1908. See Appendix A.2 for details on the data.

Key to understanding these shifts among the economic elite is the return to capital governing wealth accumulation. To this end, Figure 2(b) depicts the mean and distribution of the return to capital in our millionaire sample by wealth type. Typically, the return to capital is defined as $RoR_{it} = \frac{P_t - P_{t-1}}{P_{t-1}} + Y_{it}$. It reflects both the capital gain component $\frac{P_t - P_{t-1}}{P_{t-1}}$, i.e., changes in price P of assets held in individual's i portfolio, and the cash flow component Y_{it}. Unfortunately, our millionaires data do not allow us to calculate the return on capital this way. Since the income I and wealth W refer to the same year, we approximate the cash flow component of the return to capital by $Y_i = \frac{I_{i,1908}}{W_{i,1908} - I_{i,1908}}$.[3] For the agricultural returns this

[3]Implicitly, this makes the simplifying assumption that consumption in 1908 was negligible.

approximation is sufficient: Capital gains do not play a large role in agriculture at this time and the average of four percent depicted in Figure 2(b) is consistent with an independent estimate by Pfister (2019). For industrial returns, capital gains were more important. The data by Jordà et al. (2019) suggest that the total return on equity for the period 1880-1914 was 6.5% in Germany, i.e., half a percentage point higher than in our millionaires dataset. If we took this difference into account, the difference in the mean return for the industrialists and landowners in Figure 2(b) would be even more gaping.

Figure 2(b) also highlights the difference in the second moment of the return to capital. The difference in the cross-sectional standard deviation between the returns to agricultural and industrial wealth (1% vs. 1.6%) is substantial and would even widen if capital gains were properly accounted for. It reflects a higher risk of industrial capital and carries importance for the wealth distribution against the backdrop of intergenerational wealth models. If some firm owner repeatedly earns a higher return, ceteris paribus this will lead to a more concentrated wealth distribution. An example of such an entrepreneur in industrializing Germany was Carl Ziese, an engineer who owned a military shipbuilding company. Between 1896 and 1908 he earned an annual return of close to 16% (including capital gains), putting him at the end of the right tail of the distribution depicted in Figure 2(b).[4]

Such high returns were also possible because industrialization changed profoundly the relationship between the invested capital and the per-unit cost. In farming, there were no economies of scale until long into the 20th century: A small-scale farm would not be substantially less productive than a large-scale farm. This, of course, held also true for other members of the traditional sector such as craftsmen that had a relatively high social status, but are not found in the millionaires list as their wealth remained modest. The epitome of industrialization was the replacement of these craftsmen with large factories. According to Chandler (1994, p. 26), the "critical entrepreneurial act was not the invention—or even the initial commercialization—of a new or greatly improved product or process. Instead, it was the construction of a plant of the optimal size required to exploit fully the economies of scale or those of scope, or both." These economies of scale mattered as they led to oligopolistic market structures (Chandler, 1994). Figure 2(c) shows that the absence and presence of economies of scale for agricultural and industrial capital, respectively, are indeed reflected in the millionaires' returns. Each dot in the graph represents 10 millionaires and the colors represent the different wealth types. For the landowning elite, the small positive slope is not significant. For entrepreneurs, higher wealth leads to higher returns: Moving by a log-point in wealth (e.g., from a fortune of 7.4m Marks to one of 20.1m Marks) increases the return by about .84 percentage points. As the richest 500 millionaires represent a very selective sample, we document the positive relationship between average firm size

[4]While Martin's (1912) millionaires list generally provides only a snapshot of wealth and income for 1908, he details the wealth level for some back to 1895 or 1896. This allows us to calculate the annualized return between 1896 and 1908 for Ziese.

and capital per worker as well as firm size and non-housing capital return at the city level in Appendix Figure B.1. In line with Chandler (1994), cities with larger firms have higher capital per worker and higher non-housing returns. Industrial returns exhibited strong scale dependence, reflecting the presence of economies of scale.

Finally, differences between agricultural and industrial fortunes appear to affect the shape of the wealth distribution's tail. Figure 2(d) plots the absolute wealth of each millionaire against the corresponding wealth rank on a logarithmic scale. The wealth rank is type-specific, i.e., it refers to the position among the landowning or industrial elite of the respective individual. First, it is obvious that the curve of the industrialists is shifted to the right on the axis, indicating that at the same specific wealth rank, industrialists are substantially richer. Second, the absolute value of the fitted curve's slope–the Pareto coefficient–differs substantially across the two wealth types. The smaller Pareto coefficient (in absolute terms) for industrial fortunes indicates a higher concentration among them vis-à-vis those of the landowning elite. It also lies within the range of estimates of the Pareto coefficient for modern industrial economies of 1.3-1.8 (Saez and Zucman 2020, p. 9 and Vermeulen 2018, p. 380). Consistent with our cross-country evidence, this points to a potential relationship between industrialization and the emergence of modern tail fatness in wealth distributions.

In sum, four stylized facts emerge on the relationship of industrialization and the structure and distribution of top fortunes. First, the landowning elite had been replaced by an entrepreneurial one at the eve of World War I. Second, returns on capital show a difference in mean and variance depending on the type of wealth. Industrial returns are both on average higher and more dispersed. Third, industrial returns are increasing in the scale of the fortune while agricultural returns are not. Finally, industrial fortunes are more heavily concentrated than agricultural ones at the top.

2.3 Causal evidence from the geography of wealth concentration

The observed bifurcation between agricultural and industrial wealth also reflects the spatial consequences of the Industrial Revolution. Assuming economies of scale in manufacturing, the standard two-sector New Economic Geography model (Krugman, 1991) predicts that industrialization leads to the spatial concentration of industrial economic activity. Such dualism was indeed typical for European industrialization patterns and Germany was no exception (Tilly and Kopsidis, 2020): A continuum of regions between the 'old' agricultural economy and the 'modern' capitalist economy existed within the same country. In the following, we exploit this insight to provide causal estimates of the effect of industrialization on wealth inequality by leveraging novel data on county-level wealth distributions.

Data on local wealth distributions The Prussian state levied its wealth tax on households—the so-called *Ergänzungssteuer* or 'additional tax'—for the first time in 1895. With minor exceptions for hard-to-assess items such as art, the law's net wealth definition encapsulated all types of assets net of liabilities. Assets were valued either at market prices or capitalized values.[5] Generally, the way wealth was assessed conforms with contemporary standards laid out by the recent wealth inequality literature (Piketty et al., 2006; Piketty and Zucman, 2014). All households with a net wealth larger than 6,000 Marks had to file a tax return and hence as much as 13% of all Prussian households were assessed. Households that were not assessed for the wealth tax owned mainly savings, and reliable contemporary estimates for those exist (Albers et al., 2022).

The tax was levied at the place of the household's residence. Hence, the household's location rather than that of the household's interest-bearing assets mattered to the taxman. Fortunately, the spatial separation between households' wealth and their primary residences was rather the exception than the rule for three reasons. First, the tax itself was only .05%, limiting the incentives for tax dodging through moving the wealth abroad (Albers et al., 2022). Second, both small-scale farmers and owners of large estates held their agricultural assets at their place of residence. By the eve of World War I, these made up around 37% of Prussian net wealth (Albers et al., 2022). Third, the stock market capitalization of the German corporate sector was still quite limited (Kuvshinov and Zimmermann, 2022). Most companies were held in closely-held legal forms and their owners often lived close to the company's headquarters. Even if they had the legal form of stock companies (*Aktiengesellschaft*), they were often not publicly traded. The richest German is a case in point. Bertha Krupp von Bohlen und Halbach had inherited all shares of the eponymous industrial firm and lived in the family villa next to its first plant (Martin, 1912).[6] It is fair to assume that, for the most part, wealth was taxed where it was generated.

Since the Prussian statisticians created detailed tabulations of the tax returns and wealth levels at the regional level, it is possible to create sub-national wealth distributions that accord exactly to the same standards that underlie the national distributional estimate by Albers et al. (2022). For 1914, the Prussian statistical office created such tabulated wealth tax data even at the county level. These allow us to calculate the total net wealth of those assessed for the wealth tax. Akin to the strategy of contemporaries, we approximate the wealth of those not assessed by exploiting the regional distribution of small saving accounts. We detail all sources and steps involved in generating these data in Appendix A.3. Based on the tabulated wealth data and wealth totals at the county level, we calculate top shares with the generalized Pareto interpolation method (Blanchet et al., 2022).

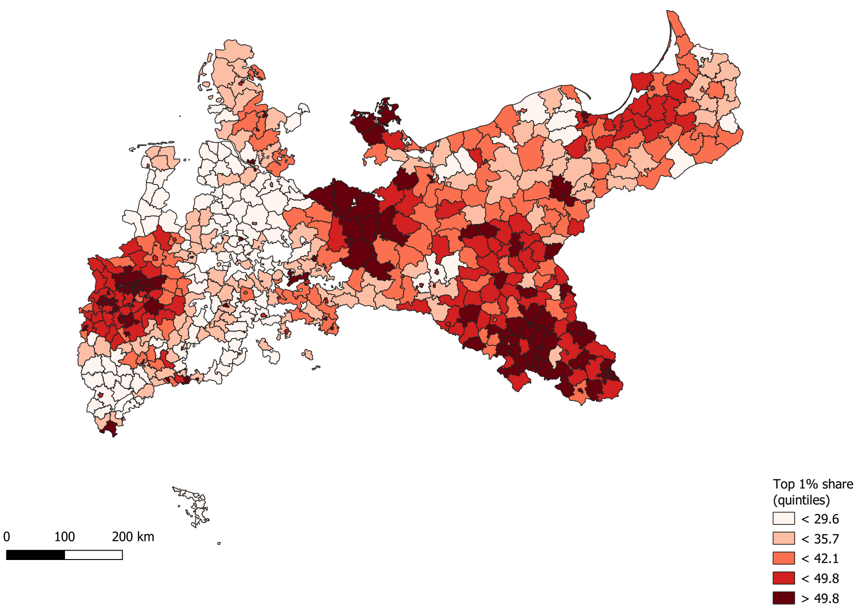
Figure 3 maps the county-level estimates for the top-1% wealth share. Relating these to historical patterns of land rather than our variable of interest (industrialization) provides a

[5]Capitalization was used in the case of rarely traded assets, most importantly for agricultural land in areas where no comparable lot had been on the market for a long time (see the appendix of Albers et al., 2022).

[6]This is not to say that entrepreneurs were not geographically mobile—many indeed were (Donges and Selgert, 2021). However, it appears from the biographies of millionaires that they typically resided in the place where they built their company.

useful way to assess the validity of our estimates. At the macro level, the pattern of wealth concentration across rural counties is consistent with the regional differences in farming that go back to before the agrarian reforms in the early 19th century. West of the river Elbe, peasants enjoyed more property rights and estates were generally smaller, whereas in the Eastern provinces—the home of the infamous Junkers—land ownership was more concentrated (Cinnirella and Hornung, 2016; Tilly and Kopsidis, 2020). At the micro level, we take this validation a step further. We collect data on land concentration at the county level and correlate it with the top percentiles' share in net wealth. The correlation is strong and positive, even conditional on district fixed effects and controls (Appendix Table A 3). As a second empirical plausibility check, we relate our estimates to the millionaires data. We uncover strong correlations with both the number and wealth of millionaires, again conditional on a large set of controls and fixed effects (Appendix Table A 4). These validations suggest that our top-1% estimates reflect historical facts and are accurate representations of local wealth inequality.

Figure 3: Top-1% wealth share in Prussian counties, 1914



Notes: Own calculations. For a detailed data description, see text and Appendix A.3.

The geographical pattern in Figure 3 also speaks to the role of industrialization in shaping the wealth distribution. Wealth was particularly concentrated in Germany's industrial heartland, the Rhine area in the West with its cluster of heavy industries and some of the newer industries such as chemicals and machinery as well as in its Eastern counterpart (Silesia). The coexistence of high wealth concentration and industrial capitalism also extends to the Berlin metropolitan area, Germany's financial and political center of the time as well as home to a burgeoning startup sector (chemicals and machinery). Finally, highly industrial areas in the province of Saxony with a more diverse mix of heavy and light industries appear to have lower levels of wealth concentration than their counterparts in Silesia or the Rhineland.[7] Economies of scale could play an important role in this pattern as these were most relevant in industries requiring large fixed-cost investments.

Identification problems and empirical strategy To quantify the effect of industrialization on wealth concentration, we collect county-level data on industrial employment shares (Appendix A.4) and relate them to our distributional estimates. However, a central concern for studying the effect of inequality and industrialization in our cross-sectional setting—unfortunately no comparable distributional data for the preindustrial period exist—is endogeneity. First, ex ante it is plausible that unobserved characteristics from before the start of industrialization affected both wealth inequality and industrialization, including, for example, the presence of merchants and the level of financialization at the county level. These could have varied substantially, especially given the preindustrial differences in urbanization across Prussia (Tilly and Kopsidis, 2020). Second, there are strong theoretical priors to suggest the presence of reverse causality. For the initial stages of industrialization, unified growth theory posits a positive effect of inequality on industrialization since the saving rate is a positive function of wealth and income (Galor and Moav, 2004). If this is indeed a relevant channel in the context of industrialization, a simple OLS estimate would be biased.

Fortunately, the recent literature on the effect of the Industrial Revolution on a variety of outcomes provides an insightful solution through a credibly exogenous and relevant instrument (Fernihough and O'Rourke, 2021; De Pleijt et al., 2020). Coal, which is crucial for the energy supply during industrialization, is often found in strata from the carboniferous age. Thus, we expect a strong negative correlation between the distance to carboniferous strata and the share of industrial employment. Crucially, the location of this strata is exogenous *and* has little relevance for economic life before the Industrial Revolution across Europe as shown by Fernihough and O'Rourke (2021). This makes it a suitable instrument for industrialization levels. Specifically, we calculate the least distance from a county's centroid to carboniferous strata using a geological map provided by the German Federal Institute for Geosciences and Natural Resources (Asch, 2005).

We then run the following 2SLS-model:

$$WC_c = \alpha + \beta \hat{Ind}_c + \delta_x X_c + \epsilon_c \quad (2)$$

$$Ind_c = \eta + \gamma Carbon_c + \rho_x X_c + v_c, \quad (3)$$

where Ind_c corresponds to the employment share in industry in county c in 1907 and WC_c corresponds to the measure of wealth concentration in county c—the top-1% wealth share in our main specification. *Carbon*, the distance to carboniferous strata, is the instrumental variable. The vector of controls X_c consists of arguably exogenous geographic variables that

[7] See Appendix Figure B.2 for a map of industrialization levels. The description of industry clusters is based on our own calculations using the occupational census of 1907 (Appendix A.4).

affect economic development. Following the literature using the same instrument (Fernihough and O'Rourke, 2021; De Pleijt et al., 2020), we include distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation (see Appendix A.4 for the corresponding sources and calculations). To account for unobservable regional differences, we add fixed effects for 36 districts, which constitute the closest administrative layer above counties.

Ideally, of course, we would add data on the level of wealth inequality before the start of the Industrial Revolution, which would allow us to estimate a difference-in-differences model. Unfortunately, these data do not exist. However, differences in wealth inequality before industrialization would only be a concern for our empirical strategy if these differences were correlated with our instrument. We can exclude this possibility for the two plausible candidates. Pre-industrial urbanization proxies potential differences in inequality between urban and non-urban counties. Land inequality in 1816 is a direct measure of inequality. Neither of them shows a significant correlation with the instrument (Appendix Table B.1). The result for urbanization is consistent with results for a broader European sample (Fernihough and O'Rourke, 2021, Figure 5).

Results We proceed with the presentation of results in two steps. First, we document the effect of industrialization on the wealth holdings of the top-percentile relative to the rest of the distribution. Second, we shift the focus on the concentration within the top decile, i.e., the fattening of the wealth distribution's tail.

Table 1: The effect of industrialization on wealth inequality

Dep. var.: Top 1% share	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Employment share industry	0.307***	0.276***	0.251***	0.288***	0.269***	0.439***
	(0.050)	(0.031)	(0.028)	(0.104)	(0.094)	(0.093)
Mean dependent variable	39.48	39.48	39.48	39.48	39.48	39.48
SD dependent variable	11.06	11.06	11.06	11.06	11.06	11.06
District FE		✓	✓		✓	✓
Geographic controls			✓			✓
F-stat excluded instrument				45.76	16.90	26.79
R-squared	0.21	0.67	0.70			
Observations	575	575	575	575	575	575

Notes: Unit of observation: county. Controls include distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses and are clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01.

Columns 1 to 3 of Table 1 report the OLS results. They document a strong correlation between industrial employment and the top percentile's wealth share. While we do not attach a causal interpretation to these regressions, they are informative in that they highlight the predictive power of the variable of interest. Column 1 suggests that using the industrial employment share as the sole predictor explains 20% of the variance in our dependent variable. The correlation does not seem to be driven by unobservable district characteristics (column 2) and is robust against controlling for other geographic features (column 3).

Columns 4 to 6 of Table 1 report the 2SLS results, including the first-stage F-statistic and the second stage coefficients.[8] Relative to the OLS estimate, the IV estimates are in a reasonably close range, in particular so for columns 1 and 2 vs. 4 and 5. The F-statistic indicates that the instrument is strongest when exploiting the variation across all of Prussia. When adding restrictive district fixed effects (column 5), it falls under the threshold of 23 proposed by Olea and Pflueger (2013) for the case of 10% potential bias and a 5% significance. Column 6 adds geographic controls, increasing the precision of the first stage and hence the F-statistic.[9] Noting that this point estimate lies within the 95% confidence interval of the one in column 5, we, therefore, use $\beta = .439$ as our preferred estimate.

The identified effect is not only statistically significant but also economically meaningful: A one standard deviation increase in our independent variable (16.67) is associated with an increase in wealth inequality of 7.3 percentage points (or 0.65 SD in the dependent variable). To put the magnitude of these effects into perspective, it is useful to compare them to both historical shocks and recent changes in the wealth distribution. Albers et al. (2022), employing a counterfactual analysis in an accounting framework, report that the decline in asset prices during the Great Depression in Germany led to a decline in wealth inequality by around 7 percentage points—a similar magnitude to the standardized coefficient reported here. Saez and Zucman (2016) estimate that the top-1% in wealth inequality increased from 23% to 35% between 1980 and 2020 in the US. This increase is often considered an extreme case in the history of wealth distributions. In our analysis moving from a county at the 25th percentile in industrial employment share to the 75th percentile ($\Delta\text{Ind} = 23.7\text{pp}$) leads to a comparable increase in wealth inequality ($\Delta s_{\text{Top-1\%}} = 10.4$ pp).

As a final plausibility check, we apply the estimated coefficients to the increase in industrialization across the time dimension. Around 6.6% of the workforce was employed in proto-industries in 1800 (Pfister, 2022, p.1079), whereas 40% worked in the industrial sector in 1907. Our estimates imply that a shift in industrial employment of this magnitude increased the top-1% wealth share by 14.6pp. That this shift is smaller than the 27pp increase observed at the national level (Section 2.1) is reassuring. After all, the most influential model on the effect of industrialization on (income) inequality rests predominantly on the role of composition effects within the aggregate national distribution, i.e., that concentration at the national level changes in response to varying the share of the population that works in the low-mean income agricultural/traditional sector versus in high-mean income industrial sec-

[8]Following Andrews et al. (2019), we apply a test statistic tailored to the case of clustered standard errors.

[9]See Appendix Tables B.2 and B.3 for first stage and reduced form regressions. We also test for the effect of a potential violation of the exclusion restriction following Conley et al. (2012) in Appendix Figure B.3.

tor (Kuznets, 1955).[10] At a high spatial resolution such as the county level, the Kuznets-type composition effects ought to be smaller than they are in the joint national distribution.

Table 2: Industrialization and the fattening of the wealth distribution's top-tail

Dep. var.:	Top 1% share (1)	Top 10 % share (2)	Top 1% / top 10% (3)
Employment share industry	0.439***	0.290***	0.355***
	(0.093)	(0.065)	(0.102)
Mean dependent variable	39.48	79.27	49.17
SD dependent variable	11.06	9.08	9.89
District FE	✓	✓	✓
Geographic controls	✓	✓	✓
F-stat excluded instrument	26.79	26.79	26.79
Observations	575	575	575

Notes: Unit of observation: county. Controls include distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses and are clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01. For the corresponding OLS results see Appendix Table B.4.

While the analysis of the top-1% share reveals a positive causal relationship between industrialization and wealth concentration, it does not speak to the dynamics of wealth accumulation among the rich. In particular, it does not reveal whether concentration is driven by the top percentile's wealth accumulation or by a larger group of the wealthy, e.g., the top-10%. Both our cross-country analysis and the stylized fact on the different tail fatness of industrial vis-à-vis agricultural fortunes (Sections 2.1 & 2.2) deliver a strong prior for answering this question: The wealth distribution in areas with higher levels of industrialization and thus more entrepreneurs should exhibit fatter tails.

Table 2 provides the corresponding test. It reports the 2SLS results for the most conservative specification, including geographic controls and district fixed effects. Comparing columns 2 and 1 suggests that the increase in the top-10% share is entirely driven by the increase in the top 1% share. If anything, the share of the 90th-99th percentiles decreased with higher levels of industrialization. Column 3 documents the effect of industrialization on a classical summary measure for tail fattening: the share of the top-1% relative to the top-10%. The effect is highly significant, both in statistical and economic terms. Moving from a county in the 25th percentile in the employment share to one in the 75th percentile (23.7pp) leads to an 8.3pp. increase in the share-in-share variable (or 83% of 1 SD). The causal cross-sectional analysis confirms the descriptive patterns found across countries: Industrialization led to an increase in the top-1% share mainly because of the polarization among the relatively wealthy.

[10] Robinson (1976) shows that, in theory, no differences in the distribution of incomes within both sectors is needed to produce the Kuznets curve.

Robustness Our results are robust against a number of exercises relating to migration and the spatial nature of our data.

A first-order concern is the potential role of self-selected migration in affecting the results. During the period under consideration, a substantial share of the population moved from rural counties to cities. If the migrants were negatively selected with respect to wealth, composition effects could occur; i.e., the arrival of poor rural laborers in cities would increase wealth inequality in urban counties while reducing it in rural counties. While the existing evidence on selective migration is weak (Grant, 2005, p. 280), we confront its potential role directly. If migration was selective and relevant for local wealth distributions, adding net migration data as a variable (from Bräuer and Kersting, forthcoming) should alter the coefficient of interest. However, we do not observe such a change leading us to conclude that selective migration does not affect our results (Appendix Table B.5).

A second potential concern is that the spatial nature of the dataset leads to incorrect inference, either because regional peculiarities in German history produce outliers or because standard errors are miscalculated in the presence of spatially correlated unobserved heterogeneity. With respect to regional peculiarities, the literature has highlighted three potential factors for economic development that varied regionally: Land inequality, religion, and ethnicity (Cinnirella and Hornung, 2016; Becker and Woessmann, 2009; Kersting et al., 2020). Appendix Tables B.6 and B.7 show that the inclusion of these variables does not alter the coefficient of interest in a meaningful way as one would expect when including district fixed effects. Next, Appendix Table B.8 shows that our results are not driven by the inclusion of Germany's industrial heartland, the Rhine/Ruhr area. To confront the potential miscalculation of standard errors due to spatially correlated unobserved heterogeneity, we show that our results are robust to the inclusion of different distance cutoffs for calculating clustered standard errors (following Colella et al., 2019, see Appendix Table B.9). In sum, our findings neither appear to be driven by migration nor regional peculiarities or the spatial nature of our data.

Summary and limitations Overall, our empirical results show that, first, industrialization causally led to wealth concentration, and, second, this process was associated with the fattening of the wealth distribution's tail. However, the empirical setup has two limitations. First, an important question is whether we can generalize these results from 'county to country'. Given the unit of observation, our empirical setup ignores potential composition effects that could occur in the joint distribution of all counties. A second limitation of the empirical setup is the inability to dissect the mechanisms at work. Industrialization always coincides with both the emergence of economies of scale and the transition of wealth-holders from one sector to the other. This makes it impossible to pin down the relative importance of these two factors for wealth concentration.

3 Modeling the industrialization-wealth inequality nexus

To rationalize the observed wealth concentration, we extend the overlapping generations model by Benhabib et al. (2011). In particular, we introduce a dynamic two-sector structure with a traditional (or agricultural) sector and a modern industrial sector. This enables us to disentangle three main mechanisms through which capital returns increased wealth concentration during industrialization. First, by modeling the emergence of and transition to the new industrial sector, we induce an increase in the overall dispersion of rates of return. Second, we add dynastic type dependence–dynasties stay in the new industrial sector once they have transitioned from the traditional sector–that increases the intergenerational persistence of returns.[11] Third, we implement sector-specific scale dependence of rate of returns in the industrial sector, increasing both the dispersion and persistence of returns. We calibrate the model targeting the pre-industrial distribution and implement industrialization as a shock that hits the economy in 1800. This allows us to compare the evolution of wealth concentration in the 19th century under different scenarios with and without scale and dynastic type dependence and benchmark them against the empirically observed but not targeted level of wealth concentration in 1914. Section 3.1 and 3.2 present the model and calibration. Section 3.3 discusses the main simulation results, while Section 3.4 conducts taxation experiments.

3.1 Model and simulation approach

Baseline model The model by Benhabib et al. (2011) is an overlapping generations model of dynastic wealth accumulation with rate of return heterogeneity across dynasties and generations. In the model, each individual dies after T periods and leaves her wealth to her only child, born at that exact point in time. At each point in time, individuals of all ages are alive, with equal proportions in the population. In a given dynasty, the individual representing generation n draws her labor income and rate of return at birth, from stochastic processes $(y_n)_n$ and $(r_n)_n$, respectively. The outcomes are then fixed over her lifetime. Given both, she chooses her optimal consumption path, maximizing her discounted utility from consumption as well as from leaving wealth to her child (see Appendix C.1).

The transition of wealth within a dynasty from one generation to the next, expressed in the respective wealth values at birth, can be summed up as

$$w_{n+1} = \alpha_n w_n + \beta_n, \tag{4}$$

where $(\alpha_n, \beta_n)_n = (\alpha(r_n), \beta(r_n, y_n))_n$ are random processes representing the individual consumption and bequest choice given r_n and y_n. Under the assumed income and rate of return

[11] Note that the original model embeds two other forms of type dependence, (1) the persistent heterogeneity between individuals induced by the randomness of individual rates of return (intragenerational type dependence, meaning a fixed individual rate of return) and (2) the intergenerational persistence of some rate of return states (autocorrelation of the rate of return process).

processes, there is a negative drift ($E(\alpha_n|\alpha_{n-1}) < 1$) with a nonzero probability of multiplicative growth ($\alpha_n > 1$).

The authors show that across dynasties wealth will have a stationary distribution with a Pareto tail. This result holds if dynasties' rates of return are autocorrelated across generations.[12] The tail is independent of the distribution of income and a potential growth in income, either across generations (growth rate g', applying to the possible income states at birth), or over each individual's lifetime (rate g), or both. This is implied by the model's key insight that the shape of the tail of the wealth distribution depends exclusively on the stochastic properties of $(r_n)_n$ and thus $(\alpha_n)_n$, which impacts wealth accumulation multiplicatively, and not $(y_n)_n$, which only enters equation (4) additively through $(\beta_n)_n$ and is negligible for large wealth values. More specifically, the inverse Pareto index, a measure of concentration at the top, increases in the variance and intergenerational persistence of $(r_n)_n$.

Modeling industrialization as a rate of return shock We now turn to the adaption and extensions of the model. Conceptually, we model industrialization as a shock to the income and distribution of rates of return across households in t_0, the onset of industrialization.

Regarding income, we assume in all scenarios that an overall income growth sets in for both the wealth-holding and the hand-to-mouth portion[13] of the population in t_0. With respect to dynasties, we assume that the stochastic initial income for an individual born in $t \geq t_0$ is given by $z_n \equiv y_n e^{(t-t_0)g'}$ (instead of just y_n as before t_0), that her income then grows with age τ according to $z(\tau) = z(0)e^{\tau g}$ with $z(0) = z_n$, and that $g' = g$. While we incorporate income growth as a characteristic feature of industrialization, it has no effect on our qualitative results (Appendix C.3.5).

Regarding the rate of return process, we assume that until t_0, all dynasties draw rates of return from an old, landowner-type discrete Markov process $({r^L}_n)_n$ with state space $S^L = \{{s^L}_j\}_{j=1}^{k_L}$ and transition matrix P. This yields a steady-state wealth distribution, using the baseline model described above. At t_0, we introduce a new industrial sector with higher average and more dispersed rates of return than the traditional sector. Individuals born after t_0 now draw from a new rate of return process covering both sectors.[14]

We simulate the model four times with different versions of this new return process

[12] Modeling the rate of return process as a Markov chain, the authors induce this intergenerational autocorrelation by modifying the conditional probabilities in the transition matrix. We provide more details on this and how we adopt this formulation in our analysis in Appendix C.3.2.

[13] The original model does not feature individuals that live hand-to-mouth. To account for the historically large share of the population falling into this class, we assume that a share ξ of the population lives hand-to-mouth throughout, both before and after industrialization. These individuals draw labor income from the same process as the wealth-holding population but they do not accumulate wealth.

[14] Note that since cohorts successively draw from the new rate of return process owing to this assumption, the trajectories of wealth concentration following t_0 exhibit a wave-like pattern, reflecting cohorts' asynchronous life cycles in the aggregate, with breaks in the curvatures 45 ($= T$) and 90 ($= 2T$) years after t_0. Alternatively, one could let all living individuals draw at once at t_0, for once regardless of their age. However, this strikes us as a less realistic assumption, as it would induce a pronounced jump in the share of industrialists.

Table 3: The examined shocks to intergenerational wealth accumulation

Scale dependence	Dynastic type dependence: **no**	Dynastic type dependence: **yes**
no	Scenario 1	Scenario 2
yes	Scenario 3	Scenario 4

(Table 3). In particular, these four scenarios allow us to study the distributional impact of dynastic type dependence and scale dependence in rates of return by assuming return processes that reflect these characteristics. In the following, we describe the two features in more detail, defining the scenarios formally in the process.

The new sector and the transition of dynasties We create scenarios 1 and 2 to gauge the long-run distributional impact of dynastic type dependence, i.e., the irreversibility of the overall relatively slow transition of dynasties to the more profitable industrial sector, relative to a setting with neither scale nor dynastic type dependence. With the emergence of the industrial sector, new rate of return states become available. In scenarios 1 and 2, we account for this by defining the respective new rate of return processes $({r^1}_n)_n$ and $({r^2}_n)_n$ as Markov chains with a common post-shock economy-wide state space $S^e = \{{s^e}_j\}_{j=1}^{k_e} = S^L \cup S^I$, where $S^I = \{{s^I}_j\}_{j=1}^{k_I}$ is the new industrial subspace. Specifically, the elements of S^e are given as ${s^e}_j = {s^L}_j$ for $j = 1, 2, ..., k_L$ and ${s^e}_j = {s^I}_{j-k_L}$ for $j = k_L + 1, k_L + 2, ..., k_L + k_I$. We define a dynasty to belong to the agricultural (industrial) sector if and only if its current generation has drawn a rate of return from the subspace S^L (S^I). The difference between scenarios 1 and 2 lies in how we model dynasties' transition probabilities regarding the subspaces and, by extension, the states in S^e.

For scenario $i \in \{1, 2\}$, denote by ϕ_i the probability that a currently agricultural dynasty will move to industry in the next generation and by ω_i the probability that a currently industrial dynasty moves back to agriculture in the next generation.[15] In Scenario 2, we induce

[15] We can sum up the transition probabilities regarding the two sectors in the matrix

$$A_i = \begin{pmatrix} 1 - \phi_i & \phi_i \\ \omega_i & 1 - \omega_i \end{pmatrix}, \tag{5}$$

where the diagonal entries denote a dynasty's new generation's probability of drawing the same sector as the previous generation, and the off-diagonal entries denote the conditional probabilities of changing the sector.

dynastic type dependence by setting $\omega_2 = 0$, implying that industry is an absorbing subspace.[16] Given this value, we set ϕ_2 such that an empirically observed share π^* of industrial dynasties is reached by 1914. Dynastic type dependence thus arises only with respect to industry once dynasties have transitioned to this sector. In Scenario 1, by contrast, we set a positive probability of returning to agriculture, i.e., $\omega_1 > 0$. To pin down the probabilities (ϕ_1, ω_1), we make two assumptions regarding Scenario 1. First, we impose that it needs to result in the same average share of industrial dynasties over the simulated time period as Scenario 2, ensuring comparability of the two scenarios with respect to the actual availability of industrial returns. Given this additional assumption it holds that $\phi_1 > \phi_2$.[17] Second, we assume that $\omega_1 = 1 - \phi_1$, thereby eliminating any conditionality of a dynasty's next rate of return draw on the current sector.[18]

Based on these transition probabilities, we define the scenarios' respective overall transition matrix with respect to the economy-wide state space S^e. In doing so, we combine the possibility of transition between sectors with within-sector rate of return heterogeneity. For scenario $i \in \{1, 2\}$ and the respective Markov process $({r^i}_n)_n$, we define that matrix as

$$Q_i = \begin{pmatrix} (1-\phi_i)P & \phi_i P \\ \omega_i P & (1-\omega_i)P \end{pmatrix}, \tag{6}$$

where the entry in row l and column m represents Prob$({r^i}_{n+1} = {s^e}_m \mid {r^i}_n = {s^e}_l)$ and we assume that $k_I = k_L$, such that Q_i has dimensions $2k_L \times 2k_L$. This latter simplifying assumption allows us to employ the same $k_L \times k_L$ probability matrix P to describe within-sector transitions between rate of return states both in the preindustrial agricultural state space as well as in the industrial absorbing subspace of Scenario 2, hence also keeping the persistence on the respective lowest and highest state implied by P constant (see Appendix C.3.2 for details). Expression (6) implies that the probability of being in a given sector (ϕ_i or ω_i) scales down the probabilities in P regarding its return states. This allows us to hold the within-sector transition probabilities and within-sector persistence, both conditional on staying in that sector, constant across scenarios and thereby to isolate the effect of dynastic type dependence.

Summing up, the rate of return processes $({r^1}_n)_n$ and $({r^2}_n)_n$ of scenarios 1 and 2 are discrete Markov chains that share the new state space S^e but differ in their respective transition matrix Q_1 and Q_2, resulting from different probabilities of dynasties switching sectors.

[16] This is a deviation from the assumption of irreducibility of the Markov chain characterizing the rate of return process in Benhabib et al. (2011).

[17] Hence, in scenarios without dynastic type dependence the share of industrial dynasties will be initially higher after t_0. However, scenarios with dynastic type dependence will catch up with and eventually pass the counterfactual scenarios with respect to this share since industry is absorbing in them (Appendix Figure C.1 illustrates this for our model calibration).

[18] While all values of ω_1 larger than zero represent a counterfactual to dynastic type dependence, this assumption appears appealing as a counterfactual since it arguably reflects the opposite extreme.

Scale dependence in industry The formalization of the new sector in scenarios 1 and 2 lacks one known feature of the emerging industrial sector: scale dependence in rates of return. Following Table 3, we thus additionally create scenarios 3 and 4 to simulate the long-run implications of this feature. We do so by adding scale dependence as a conditional additive term to the rate of return processes of scenarios 1 and 2, respectively:

$$r^3{}_n = r^1{}_n + \mathbb{1}_{\{r^1{}_n \in S^I\}} \left(p(w_n) - 0.5\right) \theta \tag{7}$$

$$r^4{}_n = r^2{}_n + \mathbb{1}_{\{r^2{}_n \in S^I\}} \left(p(w_n) - 0.5\right) \theta, \tag{8}$$

where $(r^1{}_n)_n$ and $(r^2{}_n)_n$ are the rate of return processes defined as Markov chains for scenarios 1 and 2, respectively, and $\mathbb{1}_{\{\cdot\}}$ is an indicator function taking the value one if the condition in curly brackets – that the given dynasty's current state of the Markov chain component of the process is in the industrial subspace and, equivalently, the dynasty is in the industrial sector – is fulfilled and zero otherwise. $p(w_n)$ is the percentile of generation n's initial wealth value w_n in the current distribution of all initial wealth values of individuals about to earn industrial returns[19], and the coefficient $\theta > 0$ measures scale dependence. Put differently, at birth a dynasty's new generation will be subject to a boost or a subtraction of rate of return percentage points depending on its position in the distribution of all newborn industrialists.

Other than that it is sector-specific, our formulation of scale dependence is similar to that of Benhabib et al. (2019, p. 1636). However, the specific formulation in equations (7) and (8) has the advantage that it preserves the mean rate of return in the industrial sector from the scenario without scale dependence. This facilitates a fair comparison of scenarios when removing scale dependence by setting $\theta = 0$.[20]

Model summary The dispersion and intergenerational persistence of rates of return are key drivers of wealth inequality in OLG models with heterogeneous returns. Our model differentiates the underlying sources of the two in the context of industrialization by introducing a sectoral structure. The emergence of the industrial sector with higher mean and more dispersed rates of return increases the overall variance of rates of return in the economy over time relative to the preindustrial steady state. Scenario 1 isolates this effect. Moreover, the gradual, one-directional transition (dynastic type dependence) from the old to the new sector exhibiting a higher mean rate of return induces persistent differences over generations between dynasties regarding rates of return in the simulated period. Scenarios 2

[19]Thus, $p(\cdot)$ is time-varying, as the distribution of newborn industrialists' fortunes changes over time. For reasons of readability, however, we neglect a time index in the notation.

[20]Note that introducing scale dependence as a linear relationship between the position in the wealth distribution and the rate of return is a conservative assumption when it comes to quantifying the importance of scale dependence for the simulated wealth concentration, relative to the possibility that the relationship was in fact a convex one, as found for the top 20% today (Fagereng et al., 2020).

and 4 add this feature. Finally, scale dependence in industry in Scenarios 3 and 4 increases both the variance and the intergenerational persistence of rates of return, both within industry and in the aggregate. In the following simulation, we quantify the role these factors – the emergence of industrial returns per se, dynastic type dependence, and scale dependence– play for wealth concentration over the course of industrialization.

3.2 Calibration

Our aim is to study how industrialization as a rate of return shock can explain the observed wealth concentration in 1914 and to assess the relative importance of scale and dynastic type dependence. Hence, our parameterization does not target wealth concentration in 1914. Instead, we calibrate the model to the preindustrial distribution and generate a steady state with all dynasties being landowners and thus subject to the same rate of return process. We then shock this distribution with the four scenarios of changes in the distribution of rates of return presented in Table 3 and compare the resulting trajectories of wealth concentration until 1914. In the following, we describe how we set the model parameters. We provide a summary of the parametrization in Appendix Table C.1.

Population and preference parameters We simulate the model for 1,125,000 households (around 5% of the households in Germany in 1900). Following Benhabib et al. (2011), we set the working life span T to 45. We set the hand-to-mouth ratio ξ to 0.6 based on a recent study on landholdings in Hesse-Cassel around 1850 (Wegge, 2021). Based on her data, the top 40% of the wealth distribution owned 96.8% of total landholding wealth. The negligible share of wealth accumulated in the bottom 60% is likely to have persisted throughout the entire period we consider.

We shock the preindustrial wealth distribution in $t_0 = 1800$. Moreover, for our scenarios with dynastic type dependence we proxy the share π^* of industrial dynasties in 1914 by the share of wealth taxpayers living in cities, 45% (Königlich Preussisches Statistisches Landesamt, 1915b, p. 305). Based on this value and on the procedure described in the previous section, we set the sector transition probabilities to $\omega_1 = 0.690$, $\omega_2 = 0$, $\phi_1 = 0.310$, and $\phi_2 = 0.214$ (rounded values). Finally, we set the constant relative risk aversion (CRRA) parameter σ to 1.5, consistent with the macroeconomic literature, and choose the discount rate $\rho = 0.02$, slightly lower than a standard parameterization of 0.03 in the macro literature, e.g., Benhabib et al. (2019).[21]

Income process Regarding the discounted income process $(y_n)_n$, we assume throughout that individuals draw independently across generations from a fixed discretized lognormal

[21]The reason is that we require $r \geq \rho + \sigma g$ to ensure that households do not borrow, especially given the low rates of return in agriculture. Benhabib and Bisin (2009) deal with the case of $g > 0$.

distribution that we compute based on the Gini provided for the Prussian income distribution of 1852 (see Appendix C.3.4 for details). For the income states subject to growth, z_n and $z(\tau)$ given y_n, we set the growth rates g' and g to 0.0082. This is a geometric average over the simulated period, based on the growth rates provided for 1820 to 1913 by Pfister (2019) and assuming that before 1820 the lower bound of his estimate for 1820 to 1850 applies.

Rate of return process The rate of return process is a key component of the model and hence its calibration is an important aspect of the simulation. We face the challenge of finding a plausible parameterization for it as in the model it only realizes once in a dynasty's generation rather than yearly. In theory, one would thus need long-run panel data on individual rates of return over generations to obtain an estimate. This proves difficult even today. For lack of such data in our historical period, we rely on the list of Prussian millionaires introduced above. We choose the rate of return states in agriculture, S^L, such that the mean and standard deviation of the agricultural process $({r^L}_n)_n$ are 3.99% and 1.04%, respectively. These values closely follow the observed moments in the type-specific distribution observed in the millionaires dataset (Figure 2(b)). Analogously, we choose S^I such that the standard deviation of rates of return faced by industrial dynasties in the absorbing subspace of $({r^2}_n)_n$ is 1.67%.[22] Relative to the mean industrial rate of return observed in the millionaires data (6.09%), we make a slight upward correction to 6.50% to incorporate capital gains in industry, which are unobserved in the millionaires data. This correction is based on mean equity returns in Germany for the period 1880-1913 (Jordà et al., 2019). In the spirit of the model, it would be preferable to use a cross-section of lifetime rates of return for a given birth cohort. Appendix C.3.2 provides an extensive discussion of how our chosen values for the mean and standard deviation in the rates of return result in moments for the economy-wide distribution of lifetime rates of return that are similar to those found in the literature. Last, we set the scale dependence spread θ to 0.0215 in line with Benhabib et al. (2019).[23]

Calibration of free parameters Finally, we need to choose values for the bequest motive χ as well as for κ_1 and κ_2, which add intergenerational persistence to the lowest and highest rate of return states within a given sector, respectively (see Appendix C.3.2). We follow the method of simulated moments (MSM). In generating a preindustrial wealth distribution, we conduct a grid search over possible parameter values and choose the parameter vector that minimizes squared relative residuals between the generated and empirical moments.

[22] In Scenario 1, the standard deviation of rates of return observed in industry takes this value as well, but industrial dynasties face a different standard deviation in the distribution, from which they draw a new rate of return since they can return to agriculture.

[23] Rates of return in their dynastic model are also fixed at birth, such that the scale dependence spread applies to the rate of return over a generation, ensuring the compatibility of the value with our setting. For more details on the choice of θ in light of recent evidence (Bach et al., 2020; Fagereng et al., 2020), see Appendix C.3.2.

As moments, which we weight equally in our loss function, we choose the top-1% share, which is provided for Germany in 1800 by Alfani et al. (2022), as well as, for lack of these data for preindustrial Germany, the wealth-income ratio and the inheritance flow divided by national income provided for Sweden in 1810 by Ohlsson et al. (2019).

Table 4: Model fit to preindustrial distribution by MSM

Moment	Source	Target	Fit
Top-1 share	Alfani et al. (2022) for Germany 1800	0.179	0.156
Wealth-income ratio	Ohlsson et al. (2019) for Sweden 1810	2.817	3.688
Inheritance flow/national income	Ohlsson et al. (2019) for Sweden 1810	0.101	0.059

Table 4 presents the fit of the model to the preindustrial distribution. While we do not match the moments exactly, the fit appears reasonably close. Hence, it represents a realistic preindustrial state and allows us to study the model's response to the rate of return shocks to the preindustrial distribution which have been outlined above. Appendix C.3.3 provides an extended discussion of the MSM results.

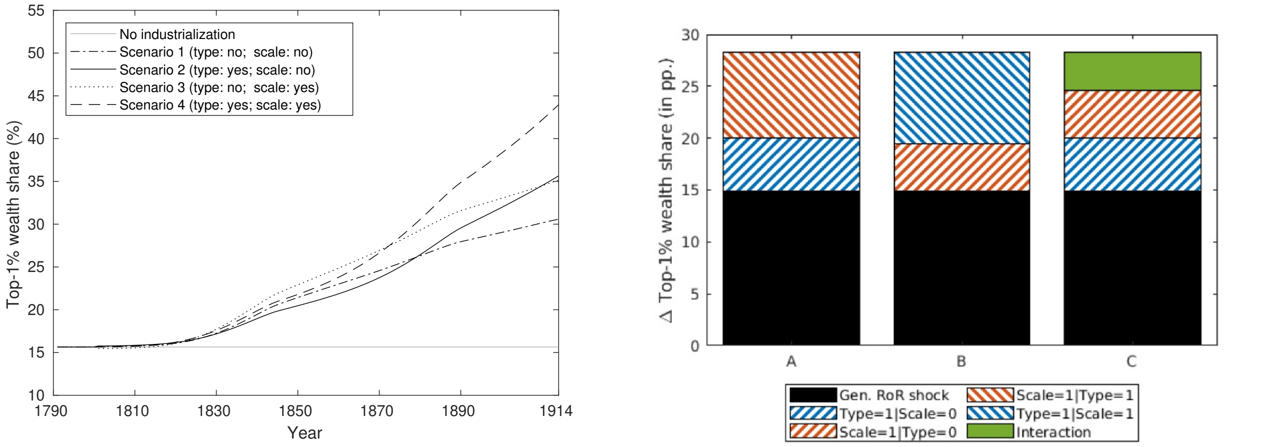
3.3 Simulation results

After calibrating the free parameters on the preindustrial state, we now simulate the model including the rate of return shock in 1800. We first explore how the resulting trajectory following each of the four scenarios of the rate of return shock matches the empirically observed levels of wealth concentration. Comparing the four scenarios also allows us to elicit the relative importance of scale and dynastic type dependence of capital returns. In the second step, we explore the role of scale and dynastic type dependence for the fattening of the tail of the wealth distribution.

Evolution of the top 1% share Figure 4(a) displays the four trajectories of the top 1 share corresponding to the four scenarios listed in Table 3 along with the steady-state level of inequality in the preindustrial society without industrialization. The scenarios with dynastic type dependence only cross their counterparts without dynastic type dependence between 1870 and 1880. This is due to the assumptions required for modeling the absence of type dependence, which necessarily overstate the accumulation among the rich in this scenario in the beginning.[24] While the curvature differs, the top-1% share increases substantially throughout the 19th century in all four scenarios and a stationary wealth distribution appears to emerge after many centuries (see Appendix C.3.6).

[24]As described in Section 3.1, in the scenarios without dynastic type dependence the initial share of industrial dynasties is higher. Hence, these scenarios likely overstate initial wealth concentration. Over time, the share of industrial dynasties in the scenarios with dynastic type dependence catches up (Appendix Figure C.1), leading to the reversal documented in the crossing of the trajectories in Figure 4(a).

Figure 4: Simulation of wealth concentration until WWI



(a) Simulated trajectories of the top-1% share

(b) Decompositions of the increase in the top-1% share

Notes: The trajectories in Figure 4(a) refer to the simulated scenarios listed in Table 3 and represent averages over 200 iterations. The abbreviations "type" and "scale" stand for the features of dynastic type dependence and scale dependence, respectively. Figure 4(b) decomposes the contribution of the two features and their interaction to the absolute change in wealth concentration. See text for details.

The dashed line corresponds to Scenario 4 featuring both scale and dynastic type dependence. It documents an increase of 28pp from 16% in 1800 to 44% in 1914. Without targeting it, this predicted level in 1914 is reassuringly close to the empirical estimate of around 45% (Albers et al., 2022).[25] The trajectories for the other three scenarios do not achieve this level of wealth concentration in 1914. While the scenarios either featuring only dynastic type dependence (Scenario 2) or only scale dependence (Scenario 3) reach a top-1% share of 35%, turning both return features off (Scenario 1) leads to an increase of only 15 percentage points vis-à-vis the preindustrial level. This implies that only half of the increase in the top-1% share can be explained with the emergence of industrial returns to capital itself. The other half of the increase originates in the scale and dynastic type dependence of capital returns. Both are required to generate realistic transition dynamics of wealth inequality in the 19th century.

Figure 4(b) adds further nuance to this finding by highlighting the importance of the interaction of scale and dynastic type dependence in the model. It decomposes the increase of the top-1% in 1914 relative to the preindustrial steady-state. The first two bars (A and B) reflect the two directions that a counterfactual analysis of the features' relative importance can take. Moving from Scenario 1 with neither of the two features (labeled "general rate of return shock") to the benchmark scenario with both features, we can either first add dynastic type dependence and then scale dependence in industry (A) or vice versa (B). The two decomposition exercises indicate that each of the two features causes a significantly larger absolute increase in the top-1 share conditional on the other. Intuitively, "locking" some dynasties into the more profitable industrial return space over generations has a larger long-run impact on aggregate wealth concentration if that space is additionally characterized by

[25]See Appendix C.3.7 for a more extensive comparison of the simulation output with observed data.

scale dependence, i.e., higher returns can potentially be earned. Conversely, the amplifying effect of scale dependence regarding the autocorrelation of rates of return across generations can fully unfold only if industrial dynasties are locked in the context determined by scale dependence.[26] The third decomposition (C) provides a quantification of the role of the interaction of scale and type dependence. It adds only the unconditional effects of scale and dynastic type dependence (4.6 and 5.1pp, respectively) and interprets the residual increase (3.8pp) as the interaction effect of the two features.

Overall, the simulation results demonstrate a sizable and approximately equal effect of both scale and dynastic type dependence of returns on the long-run top 1 share. Almost half of the total simulated increase in the top 1 share of 28pp can be explained by the co-existence of scale and dynastic type dependence. We also find that the interaction effect of these features of the return process is almost as large as each feature's unconditional effect.

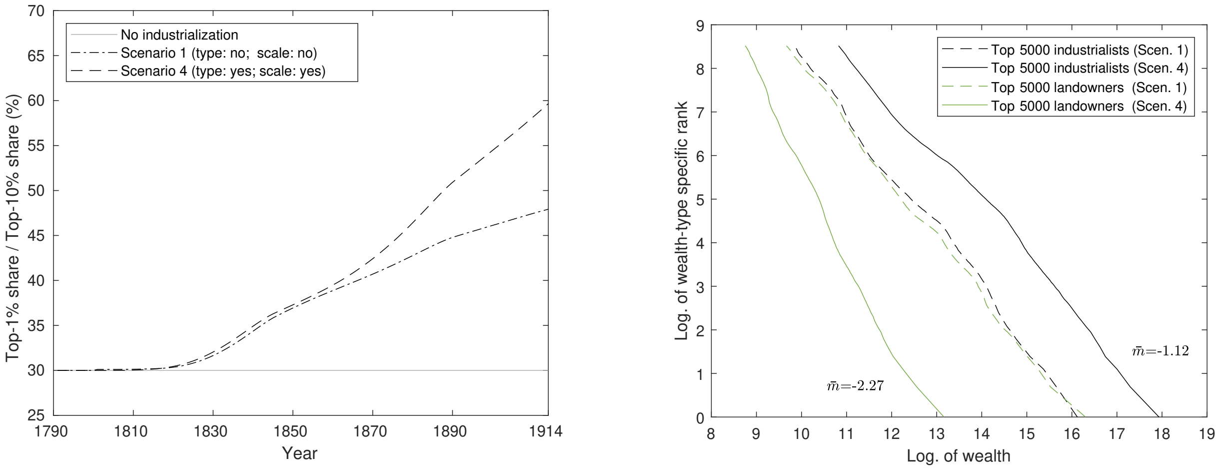
Emergence of modern tail fatness Our cross-country and within-country empirical results suggest that industrialization led to wealth concentration among the top-1% mostly through the polarization of wealth shares *within* the top decile, i.e., a fattening of the wealth distribution's tail. The model allows us to elucidate the dynamics of the emergence of heavy-tailed wealth distributions through industrialization. In particular, it generates the bifurcation of the top tail across agricultural and industrial fortunes and it enables us to pin down the importance of scale and dynastic type dependence for the tail fattening relative to that of the emergence of industrial returns in general.

Figure 5(a) plots the trajectories of the share of the top percentile in the wealth of the richest decile for two scenarios. As described above, Scenario 1 just implements the general rate of return shock while lacking both scale and dynastic type dependence. While both scenarios generate a fattening of the tail, only Scenario 4 generates transition dynamics that lead to levels that are consistent with the observed ones. In particular, the model predicts a top-1/top-10 ratio of close to 60% in 1913, not far from the actually observed 57% for Germany (based on Albers et al., 2022, see also Appendix C.3.7). In contrast, Scenario 1 only leads to a top-1/top-10 ratio of 47% by 1913. This suggests that both dynastic type dependence and scale dependence of capital returns were of key importance for the emergence of the wealth distribution's heavy tail during industrialization.

Finally, the existence of scale and dynastic type dependence of capital returns creates the bifurcation of the top tail observed in the millionaires data (Section 2.2). To illustrate this,

[26] In Appendix C.3.2, we show how the four scenarios play out with respect to the mean and standard deviation as well as the intergenerational persistence of rates of return. The two latter features have been identified by Benhabib et al. (2011) as factors for wealth concentration. As the share of dynasties in industry increases over time, these quantities evolve dynamically in all four scenarios. In particular, dynastic type dependence increases the aggregate autocorrelation of rates of return by inducing persistent sectoral differences between dynasties over generations. Scale dependence increases the variance and the autocorrelation in rates of return, the latter particularly in combination with dynastic type dependence. Last, the emergence of the industrial sector alone increases the dispersion of rates of return, driving wealth concentration in Scenario 1.

Figure 5: Tail-fattening and type-specific concentration of top wealth



(a) Relative top shares

(b) Type-specific tails in 1914

Notes: Figure 5(a): The trajectories refer to scenarios 1 (bottom curve) and 4 (top curve) listed in Table 3 and represent averages over 200 iterations of the model. The abbreviations "type" and "scale" stand for the features of dynastic type dependence and scale dependence, respectively. Figure 5(b): For scenarios 1 and 4, we separately run 50 iterations with the household number set to 6,187,500, approximately matching a quarter of the size of Germany in 1914. For the simulated outcome in 1914, we then average the wealth values of sector-specific wealth ranks at the top over the iterations. These averages are plotted by the respective curves. Additionally, we also estimate for each scenario, wealth type, and iteration the slope in the linear fit of the logarithm of wealth-type-specific ranks on the logarithm of the respective wealth value. The reported $\bar{m}$ is the respective average slope over the iterations. The estimates for Scenario 1, not displayed in the figure, are -1.35 (landowners) and -1.36 (industrialists).

Figure 5(b) displays the largest 5,000 fortunes in agriculture and in industry in the final year of the simulated period, 1914, separately for scenarios 1 and 4 based on a simulation with a larger number of households. In particular, the curves document the relationship between wealth and wealth rank, where–unlike the typical log-log representation of such plots–the rank is 'type-specific' regarding the agricultural or industrial sector. The lines in the figures represent the average of this relationship after simulating the model 50 times. The absolute value of the slope represents the Pareto exponent.

Tails in industry and agriculture are equally heavy in Scenario 1. Due to the lack of dynastic type dependence with respect to industry, the agricultural top fortunes are those that have just switched sectors, such that the agricultural top tail closely follows its industrial counterpart. By contrast, Scenario 4 with scale and dynastic type dependence produces a bifurcation of the top tail that is consistent with our empirical evidence from the millionaires list (Section 2.2). As agricultural top fortunes have never entered the more profitable industrial sector in this scenario and the income process has no effect on the inequality between the largest fortunes, the top tail of agricultural fortunes has the same fatness as at the beginning of industrialization. Hence, the juxtaposition of the agricultural and industrial top tail for Scenario 4 visualizes the start and end point of the dynamic fattening process occurring in industry in the simulated period of industrialization. In addition to the larger standard deviation of returns, scale dependence in rates of return drives the fattening of the tail of the industrial wealth distribution relative to the original (agricultural) tail fatness.

In sum, dynastic type dependence is a necessary condition for sectoral divergence of tail fatness to emerge—a structure prevalent in the millionaires list. Scale dependence in industry, in turn, increases the sector-specific tail fatness. As the two features also increase the top-1% share, our model identifies underlying drivers linking the observed tail fattening and wealth concentration.

3.4 What level of taxation could have prevented the concentration?

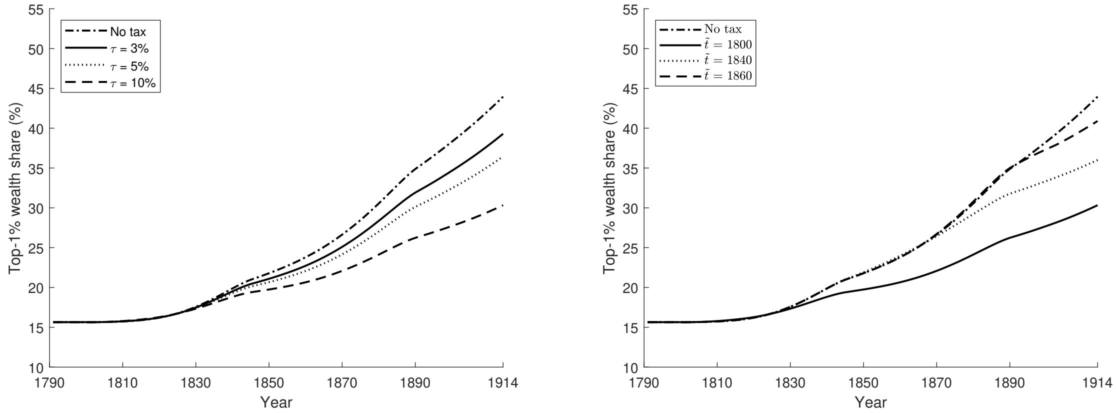
Having identified the mechanisms that led to the rapid rise of wealth concentration among the top-1% in the 19th century, it is natural to ask what policy could have done to mitigate it. To be clear, personal taxation and the taxation of capital, in particular, was limited in Germany and varied regionally throughout the 19th century (Ullmann, 2005; Spoerer, 2010). To this end, the following taxation experiment is not a plausible historical counterfactual in the sense that the existing institutions would have imposed this type of taxation in the beginning of the 19th century. Instead, it serves as a thought experiment on limiting the wealth-concentrating impact of new general purpose technologies. What levels of taxation are required and what role does the timing of implementation play?

To simplify the discussion, we focus on the taxation of capital income rather than wealth. We analyze its effect under dynastic type dependence and scale dependence in the new sector (Scenario 4). With the beginning of industrialization, individuals draw (again) from $({r^4}_n)_n$ as defined in equation (8). In contrast to before, however, in $\tilde{t} \geq t_0$ a flat capital income tax rate τ is introduced, such that individuals born in $t \geq \tilde{t}$ face a net rate of return of $(1-\tau){r^4}_n$.[27] In line with Benhabib et al. (2011), we assume that the tax proceeds are consumed by the state and not redistributed as a lump sum and that the tax has no effects on the ability to generate these returns. These assumptions are clearly not realistic, but the tax experiments are useful to demonstrate the level-timing trade-off of capital taxation as the reason for their efficacy. In particular, we explore the potential of taxation in limiting wealth concentration along two dimensions. First, we introduce the capital income tax with varying levels – $\tau \in \{0\%, 3\%, 5\%, 10\%\}$ – at the start of industrialization ($\tilde{t} = t_0 = 1800$). In a second step, we keep the tax value constant ($\tau = 10\%$) but vary the introduction date of the tax – $\tilde{t} \in \{1800, 1840, 1860\}$. Our outcome of interest is the top-1% share in 1914.

Figure 6(a) reports the results for introducing the flat tax in 1800. It reveals that modest levels of capital return taxation can have substantial effects. For example, an introduction of $\tau = 5\%$ in $t = 1800$ would have reduced the top-1% wealth share in 1914 by 8pp vis-à-vis a scenario without capital income taxation. A 10% capital income tax introduced at the same point of time (rather than in 1920 when it was actually imposed, see Deutsches Reich, 1920) would have limited the increase in the top percentile's wealth share to less than 15pp. The efficacy of this relatively modest capital flat tax originates in its ability to compress

[27] If $\tilde{t} > t_0$, individuals born in t with $t_0 \leq t < \tilde{t}$ do not have to pay a tax and hence face ${r^4}_n$ in their optimization problem and throughout their lifetime. We fade in the tax in this way for parsimony.

Figure 6: Introduction of a flat tax on capital income



(a) Varying taxation levels (introduction 1800) (b) Varying introduction dates (10% tax)

the variance of capital returns. The variance is instrumental for both the steady-state level of wealth inequality and the speed of the concentration process because the multiplicative component in the intergenerational accumulation of dynastic wealth, $\alpha(r_n)$ in equation (4), is a convex function of the capital return. Therefore, even though the capital income tax is flat, it lowers the effective intergenerational multiplicative growth factor disproportionately for dynasties that have drawn a high rate of return. In our two-sector model, against the backdrop of this convex relationship as well as dynastic type dependence, the compression of the absolute rate of return difference between sectors and of the scale dependence spread in industry are thus important channels for the tax's effectiveness.[28]

Figure 6(b) shows that the timing of introducing capital taxation is indeed crucial for its efficacy. In this experiment, we introduce the same 10% tax at different points in time. All scenarios have the same steady-state wealth shares, but the top-1% shares attained at the eve of World War I differ. Introducing the tax in 1860 has only very small effects relative to the no-tax scenario. The underlying reason is that, if its introduction is delayed ($\tilde{t} > t_0$), the top-1% share will follow the "unfettered" trajectory–that is the one without a tax, exhibiting more rapid as well as largely accelerating wealth concentration–until $\tilde{t}$.[29] Hence, the delayed introduction results in level differences in 1914. To put these into perspective, it is useful to compare the outcomes to Figure 6(a). For example, introducing a 10% tax in 1840 is less effective with respect to the reduction of the top percentile's wealth share in 1914 than introducing $\tau = 5\%$ in 1800. Put differently, a strong trade-off between the tax level and the timing of its introduction exists for time horizons relevant to society.

In sum, our results suggest that already modest levels of timely capital income taxation can slow potentially rapid increases in wealth inequality following the introduction of new

[28]See Benhabib et al. (2011, p. 134) for a discussion of the convexity-tax nexus. The convexity also implies that the tax's effect on the top-10% share is lower: While the 10% tax introduced in 1800 reduces the top-1% share increase relative to the baseline scenario of no tax by 14pp, the respective effect on the top-10% is 8pp.

[29]Both the aggregate variance in rates of return and their autocorrelation across generations, drivers of wealth concentration, increase strictly monotonically after t_0 in Scenario 4 (see Appendix C.3.2).

technologies that cause fundamental economic change such as in the case of the Industrial Revolution. However, timing appears to be crucial: A strong trade-off between the level of capital income taxation and its timing of introduction exists within our model. Whether this result generalizes to models that, unlike ours, implement the innovation-deterring effects of taxation and model growth explicitly is subject to future research.

4 Conclusion

This paper sheds new light on the distributional consequences of industrialization, the mother of all technological shocks. Cross country-data suggests that industrialization coincided with an increase in the top percentile's wealth share and that this increase originated in the fattening of the wealth distribution's tail rather than a proportional wealth concentration at the top. Leveraging new regional wealth inequality data, we document that industrialization indeed caused both the increase in the top wealth share as well as tail fattening. This tail fattening reflected the emergence of a new class of capital owners: industrial entrepreneurs. These earned substantially higher rates of return than the elite in the traditional, mostly agricultural, sector. We introduce a two-sector structure into an overlapping generations model with heterogeneous returns to reflect this bifurcation of the return process. It allows us to show that the presence and interaction of scale and dynastic type dependence were responsible for almost half of the shift in the top percentile's wealth share observed in the 19th century and that the timely introduction of modest capital taxes could have reduced it substantially.

Our results lead to a larger set of questions beyond the industrialization-inequality nexus. First, what is the role of dynastic type dependence in the 21st century? In many developing economies, the dual structure laid out in our model is still intact. Dynastic type dependence may provide a non-institutional rationale for the high levels of wealth inequality attained in them. Second, what is the relationship between economies of scale and scale dependence of capital returns throughout history? For the US, Kwon et al. (2023) document persistently increasing corporate concentration for the past 100 years. Not in all periods has this process translated to higher wealth concentration. This may be due to higher taxation of capital returns in certain periods of the 20th century vis-à-vis the period of industrialization. Moreover, the changing ownership structures of companies—from family-run businesses to public companies—may have played a role in diluting the relationship between economies of scale and the scale dependence of returns. Finally, what is the role of economies of scale for wealth inequality in the future? Platform and AI-based businesses benefit from economies of scale and it is conceivable that this effect will increase over time for the latter. Unless policies are designed to reign in the corporate concentration resulting from it, these may lead to excessive levels of wealth concentration as–in the wording of our model–the economy might break once again into a traditional and modern sector.

References

Acemoglu, Daron, "Technical change, inequality, and the labor market," *Journal of Economic Literature*, 2002, *40* (1), 7–72.

_ **and Simon Johnson**, *Power and progress: Our thousand-year struggle over technology and prosperity*, Public Affairs, 2023.

Albers, Thilo N. H., Charlotte Bartels, and Moritz Schularick, "The distribution of wealth in Germany, 1895-2018," *CESifo Working Paper No. 9739*, 2022.

Alfani, Guido, "Economic inequality in preindustrial times: Europe and beyond," *Journal of Economic Literature*, 2021, *59* (1), 3–44.

_ **, Victoria Gierok, and Felix Schaff**, "Economic inequality in preindustrial Germany, ca. 1300–1850," *The Journal of Economic History*, 2022, *82* (1), 87–125.

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

Alvaredo, Facundo, Anthony B. Atkinson, and Salvatore Morelli, "Top wealth shares in the UK over more than a century," *Journal of Public Economics*, 2018, *162*, 26–47.

Andrews, Isaiah, James H. Stock, and Liyang Sun, "Weak instruments in instrumental variables regression: Theory and practice," *Annual Review of Economics*, 2019, *11*, 727–753.

Asch, Kristine, *The 1:5 million international geological map of Europe and adjacent areas*, Hannover: German Federal Institute for Geosciences and Natural Resources (BGR), 2005.

Atack, Jeremy, Fred Bateman, and Robert A. Margo, "Skill intensity and rising wage dispersion in nineteenth-century American manufacturing," *The Journal of Economic History*, 2004, *64* (1), 172–192.

Augustine, Dolores L., *Patricians and parvenus: Wealth and high society in Wilhemine Germany*, Bloomsbury Publishing, 1994.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini, "Rich pickings? Risk, return, and skill in household wealth," *American Economic Review*, 2020, *110* (9), 2703–47.

Bartels, Charlotte, Felix Kersting, and Nikolaus Wolf, "Testing Marx: Capital accumulation, income inequality, and socialism in late nineteenth-century Germany," *The Review of Economics and Statistics*, forthcoming.

Becker, Sascha O. and Ludger Woessmann, "Was Weber wrong? A human capital theory of Protestant economic history," *The Quarterly Journal of Economics*, 2009, *124* (2), 531–596.

Bengtsson, Erik, Anna Missiaia, Ilkka Nummela, and Mats Olsson, "Unequal poverty and equal industrialisation: Finnish wealth, 1750-1900," *Scandinavian Economic History Review*, 2019, *67* (3), 229–248.

_ **,** _ **, Mats Olsson, and Patrick Svensson**, "Wealth inequality in Sweden, 1750–1900," *The Economic History Review*, 2018, *71* (3), 772–794.

Benhabib, Jess, Alberto Bisin, and Mi Luo, "Wealth distribution and social mobility in the US: A quantitative approach," *American Economic Review*, 2019, *109* (5), 1623–47.

_ **,** _ **, and Ricardo T Fernholz**, "Heterogeneous Dynasties and Long-Run Mobility*," *The Economic Journal*, 12 2021, *132* (643), 906–925.

_ **,** _ **, and Shenghao Zhu**, "The distribution of wealth and fiscal policy in economies with finitely lived agents," *Econometrica*, 2011, *79* (1), 123–157.

— and —, "The distribution of wealth and fiscal policy in economies with finitely lived agents," *NBER Working Paper No. 14730*, 2009.

Blanchet, Thomas, Juliette Fournier, and Thomas Piketty, "Generalized Pareto curves: Theory and applications," *Review of Income and Wealth*, 2022, *68* (1), 263–288.

Broadberry, Stephen, "How did the United States and Germany overtake Britain? A sectoral analysis of comparative productivity levels, 1870–1990," *The Journal of Economic History*, 1998, *58* (2), 375–407.

— and Carsten Burhop, "Comparative productivity in British and German manufacturing before World War II: Reconciling direct benchmark estimates and time series projections," *The Journal of Economic History*, 2007, *67* (2), 315–349.

—, Giovanni Federico, and Alexander Klein, "Sectoral developments, 1870-1914," in Stephen Broadberry and Kevin H. O'Rourke, eds., *The Cambridge Economic History of Modern Europe*, Cambridge University Press, 2010, pp. 59–83.

Bräuer, Richard and Felix Kersting, "Trade shocks, labor markets, and migration in the first globalization," *The Economic Journal*, forthcoming.

Chandler, Alfred D., *Scale and scope. The dynamics of industrial capitalism*, Belknap Press, 1994.

Cinnirella, Francesco and Erik Hornung, "Landownership concentration and the expansion of education," *Journal of Development Economics*, 2016, *121*, 135 – 152.

Colella, Fabrizio, Rafael Lalive, Seyhun Orcan Sakalli, and Mathias Thoenig, "Inference with arbitrary clustering," *IZA Discussion Paper No. 12584*, 2019.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi, "Plausibly exogenous," *Review of Economics and Statistics*, 2012, *94* (1), 260–272.

Cummins, Neil, "The hidden wealth of English dynasties, 1892–2016," *The Economic History Review*, 2022, *75* (3), 667–702.

De Vries, Jan, "The industrial revolution and the industrious revolution," *The Journal of Economic History*, 1994, *54* (2), 249–270.

Deutsches Reich, "Deutsches Kapitalertragsteuergesetz. Vom 29. März 1920. (R.G.Bl. 1920 Nr. 54 S. 345)," *FinanzArchiv / Public Finance Analysis*, 1920, *37* (2), 283–288.

Donges, Alexander and Felix Selgert, "The social background of Prussian inventors and entrepreneurs during the first industrial revolution," *Zeitschrift für Unternehmensgeschichte*, 2021, *66* (1), 1–41.

Dray, Sacha, Camille Landais, and Stefanie Stantcheva, "Wealth and property taxation in the United States," *NBER WP No. 31080*, 2023.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity and persistence in returns to wealth," *Econometrica*, 2020, *88* (1), 115–170.

Fernihough, Alan and Kevin Hjortshøj O'Rourke, "Coal and the European industrial revolution," *The Economic Journal*, 2021, *131* (635), 1135–1149.

Franck, Raphaël and Oded Galor, "Technology-skill complementarity in early phases of industrialisation," *The Economic Journal*, 2022, *132* (642), 618–643.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The dynamics of inequality," *Econometrica*, 2016, *84* (6), 2071–2111.

Gaillard, Alexandre and Philipp Wangner, "Wealth, Returns, and Taxation: A Tale of Two Dependencies," *SSRN No. 3966130*, 2022.

Gajek, Eva Maria, "Sichtbarmachung von Reichtum. Das Jahrbuch des Vermögens und Einkommens der Millionäre in Preussen," *Archiv für Sozialgeschichte*, 2014, *54.*

Galloway, Patrick R., *Galloway Prussia database 1861 to 1914*, www.patrickgalloway.com, 2007.

Galor, Oded and Omer Moav, "From physical to human capital accumulation: Inequality and the process of development," *Review of Economic Studies*, 2004, *71* (4), 1001–1026.

_ **and** _ , "Das Human-Kapital: A theory of the demise of the class structure," *Review of Economic Studies*, 2006, *73* (1), 85–117.

Garbinti, Bertrand, Jonathan Goupille-Lebret, and Thomas Piketty, "Accounting for wealth-inequality dynamics: Methods, estimates, and simulations for France," *Journal of the European Economic Association*, 2021, *19* (1), 620–663.

Grant, O., *Migration and inequality in Germany 1870-1913*, Oxford: Oxford University Press, 2005.

Hubmer, Joachim, Per Krusell, and Anthony A. Smith Jr, "Sources of US wealth inequality: Past, present, and future," *NBER Macroeconomics Annual*, 2021, *35* (1), 391–455.

Jones, Charles I. and Jihee Kim, "A Schumpeterian model of top income inequality," *Journal of Political Economy*, 2018, *126* (5), 1785–1826.

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor, "The rate of return on everything, 1870-2015," *The Quarterly Journal of Economics*, 2019, *134* (3), 1225–1298.

Kersting, Felix, Iris Wohnsiedler, and Nikolaus Wolf, "Weber revisited: The Protestant ethic and the spirit of nationalism," *The Journal of Economic History*, 2020, *80* (3), 710–745.

Königlich Preussisches Statistisches Landesamt, *Statistisches Jahrbuch für den Preussischen Staat 1908*, Verlag des Königlichen Statistischen Landesamts, 1909.

Kopczuk, W. and E. Saez, "Top Wealth Shares in the United States, 1916-2000: Evidence from Estate Tax Returns," *National Tax Journal*, 2004, *57* (2), 445–487.

Krugman, Paul, "Increasing returns and economic geography," *Journal of Political Economy*, 1991, *99* (3), 483–499.

Kuvshinov, Dmitry and Kaspar Zimmermann, "The big bang: Stock market capitalization in the long run," *Journal of Financial Economics*, 2022, *145* (2), 527–552.

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

Kwon, Spencer Yongwook, Yueran Ma, and Kaspar Zimmermann, "100 years of rising corporate concentration," *SAFE Working Paper No. 359*, 2023.

Königlich Preussisches Statistisches Landesamt, *Statistisches Jahrbuch für den Preussischen Staat 1906*, Verlag des Königlichen Statistischen Landesamts, 1907.

_ , *Statistik der preussischen Einkommenssteuer-Veranlagung für das Steuerjahr 1908 und die Ergänzungssteuer-Veranlagung für die Steuerjahre 1908/1910*, Verlag des Königlichen Statistischen Landesamts, 1908.

_ , *Statistisches Jahrbuch für den Preussischen Staat 1911*, Verlag des Königlichen Statistischen Landesamts, 1912.

_ , *Hauptübersicht über den Geschäftsbetrieb und die Ergebnisse der preussischen Sparkassen im Rechnungsjahre 1913. Sonderabdruck aus der 'Zeitschrift des Königlich Preussisches Statistisches Landesamt'*, Buchdruckerei W. Koebke, 1915.

— , *Statistisches Jahrbuch für den Preussischen Staat 1914*, Verlag des Königlichen Statistischen Landesamts, 1915.

Lewis, W. Arthur, "Economic development with unlimited supplies of labour," *The Manchester School*, 1954, *22* (2), 139–191.

Lindert, Peter H., "Unequal English wealth since 1670," *Journal of Political Economy*, 1986, *94* (6), 1127–1162.

— , "Three centuries of inequality in Britain and America," in Anthony B. Atkinson and François Bourguignon, eds., *Handbook of Income Distribution Volume 1*, Elsevier, 2000, pp. 167–216.

Martin, Rudolf, *Jahrbuch des Vermögens und Einkommens der Millionaire in Preussen, 1912*, Verlag W. Herlet GmbH, 1912.

Marx, Karl, "Das Kapital. Kritik der Politischen Ökonomie. Erster Band," in "Karl Marx Friedrich Engels Werke (MEW). Band 23," Berlin: Dietz Verlag, 1867.

Moll, Benjamin, Lukasz Rachel, and Pascual Restrepo, "Uneven growth: Automation's impact on income and wealth inequality," *Econometrica*, 2022, *90* (6), 2645–2683.

Ohlsson, Henry, Jesper Roine, and Daniel Waldenström, "Inherited wealth over the path of development: Sweden, 1810–2016," *Journal of the European Economic Association*, 2019, *18* (3), 1123–1157.

Olea, José Luis Montiel and Carolin Pflueger, "A robust test for weak instruments," *Journal of Business & Economic Statistics*, 2013, *31* (3), 358–369.

Pfister, Ulrich, "The Crafts–Harley view of German industrialization: An independent estimate of the income side of net national product, 1851–1913," *European Review of Economic History*, 2019, *24* (3), 502–521.

— , "Economic growth in Germany, 1500–1850," *The Journal of Economic History*, 2022, *82* (4), 1071–1107.

Piketty, Thomas, *Capital in the 21st century*, Cambridge [MA]: Harvard University Press, 2014.

— , "About capital in the twenty-first century," *American Economic Review*, May 2015, *105* (5), 48–53.

— **and Gabriel Zucman**, "Capital is back: Wealth-income ratios in rich countries 1700-2010," *Quarterly Journal of Economics*, 2014, *129* (3), 1255–1310.

— **and** — , "Wealth and inheritance in the long run," in Anthony B. Atkinson and François Bourguignon, eds., *Handbook of Income Distribution*, Elsevier, 2015, pp. 1303–1368.

— **, Gilles Postel-Vinay, and Jean-Laurent Rosenthal**, "Wealth concentration in a developing economy: Paris and France, 1807-1994," *American Economic Review*, 2006, *96* (1), 236–256.

Pleijt, Alexandra De, Alessandro Nuvolari, and Jacob Weisdorf, "Human capital formation during the first industrial revolution: Evidence from the use of steam engines," *Journal of the European Economic Association*, 2020, *18* (2), 829–889.

Ridolfi, Leonardo, Carla Salvo, and Jacob Weisdorf, "The impact of mechanisation on wages and employment: evidence from the diffusion of steam power," *CEPR Discussion Paper No. 17389*, 2022.

Robinson, Sherman, "A note on the u hypothesis relating income inequality and economic development," *The American Economic Review*, 1976, *66* (3), 437–440.

Roine, Jesper and Daniel Waldenström, "Wealth concentration over the path of development: Sweden, 1873–2006," *The Scandinavian Journal of Economics*, 2009, *111* (1), 151–187.

Saez, Emmanuel and Gabriel Zucman, "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data," *The Quarterly Journal of Economics*, 2016, *131* (2), 519–578.

_ **and** _ , "The rise of income and wealth inequality in America: Evidence from distributional macroeconomic accounts," *Journal of Economic Perspectives*, 2020, *34* (4), 3–26.

Schubert, Ulrich, *Gemeindeverzeichnis Deutschland*, www.gemeindeverzeichnis.de, 2022.

Spoerer, Mark, "The evolution of public finances in nineteenth-century Germany," in José Luís Cardoso and PedroEditors Lains, eds., *Paying for the liberal State: The rise of public finance in nineteenth-century Europe*, Cambridge University Press, 2010, pp. 103–131.

Statistik des Deutschen Reichs, *Berufs- und Betriebszählung vom 12. Juni 1907. Gewerbliche Betriebsstatistik. Band 218*, Verlag von Puttkammer & Mühlbrecht, 1909.

Tilly, Richard H., "German industrialization," in M. Teich and R. Porter, eds., *The Industrial Revolution in National Context: Europe and the USA*, Cambridge University Press, 1996, pp. 95–125.

_ **and Michael Kopsidis**, *From old regime to industrial state: A history of German industrialization from the eighteenth century to World War I*, University of Chicago Press, 2020.

Ullmann, Hans-Peter, *Der deutsche Steuerstaat: Geschichte der öffentlichen Finanzen vom 18. Jahrhundert bis heute*, C.H. Beck, 2005.

Vermeulen, P., "How fat is the top tail of the wealth distribution," *Review of Income and Wealth*, 2018, *64* (2), 357–387.

Wegge, Simone A., "Inheritance institutions and landholding inequality in nineteenth-century Germany: Evidence from Hesse-Cassel villages and towns," *The Journal of Economic History*, 2021, *81* (3), 909–942.

Wehler, Hans-Ulrich, *Deutsche Gesellschaftsgeschichte. Dritter Band: Von der "Deutschen Doppelrevolution" bis zum Beginn des Ersten Weltkrieges, 1849-1914*, C.H. Beck, 2008.

_ , *Deutsche Gesellschaftsgeschichte. Erster Band: Vom Feudalismus des Alten Reiches bis zur defensiven Modernisierung der Reformära 1700-1815*, C.H. Beck, 2008.

Zanden, Jan Luiten Van, "Tracing the beginning of the Kuznets curve: Western Europe during the early modern period," *Economic History Review*, 1995, *48* (4), 643–664.

Online Appendix "Industrialization, returns, inequality"

A Data Appendix

A.1 Cross-country estimates and decomposition

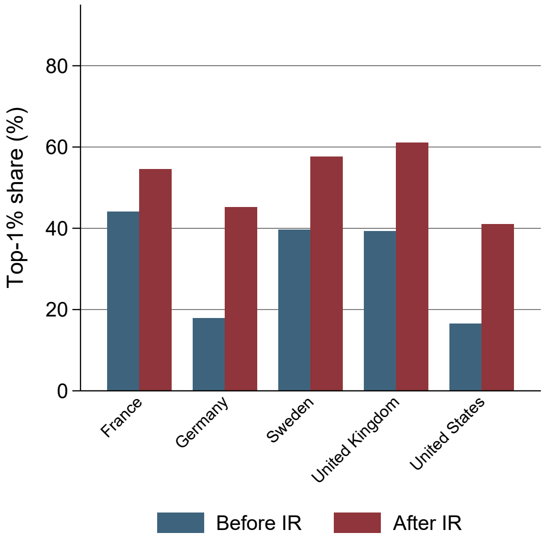
Data and data quality Table A 1 reports the years and sources included when calculating the increase of wealth shares during industrialization. The quality of existing material, especially for the pre-industrial period varies. As a basic rule, we aim to include data from the same source, as this increases the consistency and compatibility of the pre-industrial estimates and those close to the eve of World War I. For example, we prefer to use the estimate by Lindert (1986) for 1875 even though Alvaredo et al. (2018) provide estimates for 1913. The reason lies in the consistency of the tax unit concept (households in Lindert 1986 versus individuals in Alvaredo et al. 2018). Hence, the estimates for the United Kingdom, France, and Sweden have a high degree of comparability since they are from the same source and use the exact same wealth definitions and methods throughout.

Table A 1: Sources of cross-country wealth data

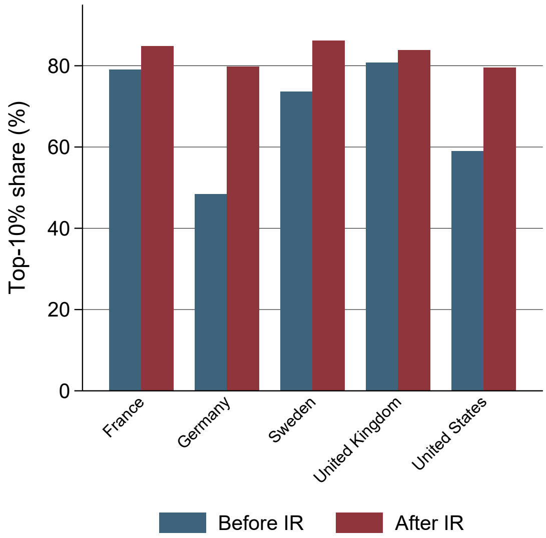
Country	Year		Sources		Tax unit
	t_0	t_1	t_0	t_1	
France	1807	1913	Garbinti et al. (2021)		Individuals
Germany	1800	1913	Alfani et al. (2022)	Albers et al. (2022)	Households
Sweden	1800	1900	Bengtsson et al. (2018)		Households
United Kingdom	1700	1875	Lindert (1986)		Households
United States	1774	1917	Lindert (2000)	Saez and Zucman (2016)	Households

For Germany and the United States, no series from a single source exists for both the pre-industrial period and the endpoint around 1913. For the United States, Lindert (2000, p. 186) discusses Alice Hanson Jones' estimate for 1774: Despite the initially small sample, follow-up studies largely confirmed her results. The big advantage of her estimate is that it spans all 13 colonies (rather than a subset of them as in other studies). We link these estimates to those from Saez and Zucman (2016) for 1917 which have the same tax unit definition (households/family) and level-wise largely agree with Kopczuk and Saez (2004) for this point in time. The least reliable estimate in our 5-country sample is that for pre-industrial Germany. Alfani et al. (2022) make a composite estimate of scattered regional urban and non-urban distributions. While this limits the comparability with the series by Albers et al. (2022), it is doubtful that better sources for Germany for the period around 1800 exist and the authors do their best to account for potential problems. As an alternative, we could have excluded the Germany series. However, as our cross-sectional data pertains to Prussia (Germany's largest state), we found it appropriate to include it despite the tentative nature of the pre-industrial estimate.

Figure A1: The Industrial Revolution and wealth concentration



(a) Top-1% wealth share levels



(b) Top-10% wealth share levels

Notes: For sources see Table A 1.

For completeness, Figure A1 provides the levels for the top-1% and top-10% wealth shares before and after the Industrial Revolution. It shows that the pre-industrial German levels reported by Alfani et al. (2022) are in line with those in the United States. In the three other countries, the levels are higher. However, note that these cross-country differences may also originate in differences in the valuation of assets or precise definitions of which assets are included. For example, under the head 'movables' the Swedish series includes all assets that are neither real estate nor financial wealth, e.g. jewelry, paintings, livestock, and consumer durables (Bengtsson et al., 2018). While these account for a substantial share of Swedish wealth, some of these assets may not be included or included at a lower value in other countries.

Decomposition Equation (1) holds as an identity. It can be derived by adding three zeros and rearranging as follows:

$$
\begin{aligned}
s_1^{\text{Top 1}} - s_0^{\text{Top 1}} &= \gamma_1 s_1^{\text{Top 10}} - \gamma_0 s_0^{\text{Top 10}} \\
&\quad + \gamma_0 s_1^{\text{Top 10}} - \gamma_0 s_1^{\text{Top 10}} + \gamma_1 s_0^{\text{Top 10}} - \gamma_1 s_0^{\text{Top 10}} + \gamma_0 s_0^{\text{Top 10}} - \gamma_0 s_0^{\text{Top 10}} \\
&= \gamma_0 s_1^{\text{Top 10}} - \gamma_0 s_0^{\text{Top 10}} + \gamma_1 s_0^{\text{Top 10}} - \gamma_0 s_0^{\text{Top 10}} \\
&\quad + \gamma_1 s_1^{\text{Top 10}} - \gamma_1 s_0^{\text{Top 10}} - \gamma_0 s_1^{\text{Top 10}} + \gamma_0 s_0^{\text{Top 10}} \\
&= \gamma_0 \Delta s^{\text{Top 10}} + (\gamma_1 - \gamma_0) s_0^{\text{Top 10}} + (\gamma_1 - \gamma_0) \Delta s^{\text{Top 10}}
\end{aligned}
$$

A.2 Millionaires list

The source The millionaires data are from the *Jahrbuch des Vermögens und Einkommens der Millionäre in Preussen* ("Yearbook of wealth and income of the millionaires of Prussia"), a rich list published by the Prussian civil servant Rudolf Martin (1912). The list ranks all approximately 8000 millionaires in 1908 on the territory of Prussia by their fortune. The households represent the top 0.06% of Prussian households and own around 20 percent of the Prussian total net personal wealth estimated by Albers et al. (2022).

The author of the rich list, Rudolf Martin (1867-1939), was a former civil servant in the Prussian ministry of interior and a controversial figure in Prussia and Imperial Germany. His motivation for publishing this list was social-reformatory as he states in the introduction. The publication of the *Jahrbuch* was a big scandal, including a house search and a temporary confiscation of the manuscript. Gajek (2014) provides an excellent account of Martin's life in connection with the publication.

In general, the *Jahrbuch* contains the following information:

1. Name and titles (noble and others).
2. Address: We use this information to code the county in which the millionaire lives.
3. Income and wealth.
4. Current occupation, e.g., *Rittergutsbesitzer* (owner of large agricultural estate), *Fabrikbesitzer* (owner of a factory), or *Aufsichtsrat in Bank xx* (Board member of bank xx).
5. Short biographies for the richest millionaires. These allow us to code the actual origin of the fortune (rather than the current occupation) for the richest 500 millionaires into the 7 sectors shown in the paper.

Reliability of the source An important question is how reliable the source is. Given Martin's privileged access as a Prussian bureaucrat and that Martin reports the addresses of all millionaires, it is likely that he actually used the original tax returns. Alternatively, it is also possible that he used the tabulated official statistics—published at the district (*Regierungsbezirk*) level—and then researched the rich at the district level (including their fortune and income). It is impossible to ultimately determine how Martin generated the list. Augustine (1994) and Gajek (2014) discuss the origin and quality of Martin's yearbooks–they exist for other states and Germany as a whole as well–in depth.

As a validation of the source, Table A 2 compares the tabulation of millionaires in the statistical yearbook (Königlich Preussisches Statistisches Landesamt, 1909) with the data in the rich list by sorting the latter into the bins used in the official statistics. Meaningful differences between the official tax data and the rich list only arise at wealth levels below 3 million Marks. Even those are far below 5%. For the richest 500 millionaires—starting from the bin with the lower bound of 6 million—the differences are negligible. It appears that

Table A 2: Comparison of Martin's rich list with the official 1908 wealth tax data

Wealth bin (lower bound in m)	Count tax data		Count rich list		Difference	
	in bin	cumulative	in bin	cumulative	in bin	cumulative
1	5294	8377	5014	8032	-280	-345
2	1455	3083	1357	3018	-98	-65
3	593	1628	605	1661	12	33
4	285	1035	295	1056	10	21
5	192	750	191	761	-1	11
6	121	558	120	570	-1	12
7	90	437	97	450	7	13
8	62	347	63	353	1	6
9	53	285	56	290	3	5
10	26	232	25	234	-1	2
11	31	206	32	209	1	3
12	32	175	32	177	0	2
13	13	143	13	145	0	2
14	17	130	17	132	0	2
15	13	113	16	115	3	2
16	8	100	8	99	0	-1
17	16	92	15	91	-1	-1
18	8	76	9	76	1	0
19	4	68	3	67	-1	-1
20	24	64	23	64	-1	0
25	9	40	10	41	1	1
30	5	31	5	31	0	0
35	6	26	5	26	-1	0
40	8	20	9	21	1	1
50	8	12	8	12	0	0
100	4	4	4	4	0	0
200	0	0	0	0	0	0

Notes: The wealth data by Martin (1912) are binned according to the binning scheme of the 1908 Prussian aggregate wealth data from Königlich Preussisches Statistisches Landesamt (1909).

Martin's data is consistent with the 'aggregate' tax data, a reassuring validation in addition to the assessment by Augustine (1994, p. 22f.).

Portfolio diversification among the rich We calculate the mean and dispersion of the millionaires' capital returns by sector (agriculture vs. industry). This implicitly assumes that the diversification across asset classes of wealth-holders in these two sectors was negligible. The description of the millionaires' fortunes in Martin's yearbook suggests that this was indeed the case: Traditional landowners and industrialists did not extensively diversify their fortunes. For the rural landowners, their status was linked to the land and thus, culturally, incentives to diversify were limited (Wehler, 2008b, p. 150). For the entrepreneurial elite, it often proved difficult to diversify their wealth as their businesses were held in legal forms other than publicly traded stock companies. Albers et al. (2022, Appendix p. 80) provide further evidence on the lack of portfolio diversification among the German rich for the pre-1913, interwar, and early post-war period.[1] All this is not to say that business owners or wealthy rural landowners did not own any urban real estate. However, most of their fortune would be in their company and agricultural estate, respectively, and their portfolios would be far from representing a balanced portfolio of different asset types.

A.3 Estimates of sub-national wealth distributions

We present the procedure for estimating the wealth distributions at the county level in two steps. First, we discuss the underlying sources. Second, we discuss how we employ these to generate wealth tabulations at the county level and ultimately estimate top shares.

To improve the readability, we use the following notation. We denote the varying geographic resolutions with subscripts:

- n for national.
- d for district, dra for district-level data only including municipalities within district d with more than 2,000 inhabitants, drb for district-level data only including municipalities within district d with less than 2,000 inhabitants, dru for district-level data only including urban municipalities within district d.
- c for county.

Totals are denoted with upper-case letters (e.g., W for the wealth total and N for the number of households/tax units), whereas lower-case letters indicate totals of a geographical unit normalized by the number of households of that respective geographical unit (e.g., $w_n = W_n/N_n$ denotes households' average wealth calculated at the national level).

A.3.1 Sources

For the estimation of top shares at the county level, we require wealth totals, tabulated wealth data, and the total number of tax units at various geographic resolutions. We also make use of data on population structures and savings. We detail the sources below:

National aggregates We use the net wealth estimate as well as the number of tax units for Prussia as provided by Albers et al. (2022). These include:

- Total number of tax units N_n: 16,254,480

[1] Before the start of the Industrial Revolution, there was one group that diversified their wealth: The urban patrician elite. Next to their businesses, they acquired urban real estate and, in some cases, also rural estates (Wehler, 2008b, p. 151).

– of which are assessed for taxes because their wealth is more than 6,000 Marks: $N_n^{\text{assessed}} = 2,183,774$.

– of which are exempted because their wealth is less than 6,000 Marks: $N_n^{\text{not assessed}} = 14,070,706$.

- Total net private wealth: $W_n = 143,787,913,070$ Marks

  – of which is owned by those assessed for taxes, i.e., having wealth of more than 6,000 Marks: $W_n^{\text{assessed}} = 130,868,400,589$ Marks.

  – of which is owned by those exempted because their wealth is less than 6,000 Marks: $W_n^{\text{not assessed}} = 12,919,512,481$ Marks.

District tabulations of those assessed for the wealth tax

Tax units: Königlich Preussisches Statistisches Landesamt (1916, p. 408f) provides the number of tax units for each county and we aggregate those up to the districts.

Wealth: For each district, Königlich Preussisches Statistisches Landesamt (1914) provides the number of taxpayers and total tax paid in each tax bracket, the number of exempt households, and those paying a reduced rate. It also dis-aggregates the results for each district by the size of the municipalities, in particular:

- rural municipalities with less than 2,000 inhabitants
- rural municipalities with more than 2,000 inhabitants
- urban municipalities

This is important as it allows us to estimate the average wealth by bracket in rural counties in connection with data on each county's population structure (see below), ultimately leading to improved estimates of wealth concentrations. For urban counties, we have a source reporting this (see below).

Tabulations for urban counties of those assessed for the wealth tax

Tax units: Königlich Preussisches Statistisches Landesamt (1916, p. 408f)

Wealth: Königlich Preussisches Statistisches Landesamt (1914) provides detailed data on the number of households and tax paid in each tax bracket, those exempted, and those paying a reduced tax rate.

Tabulations for rural counties of those assessed for the wealth tax

Tax units: Königlich Preussisches Statistisches Landesamt (1916, p. 408f)

Wealth: Königlich Preussisches Statistisches Landesamt (1916, p. 408f) provides tabulated wealth data at a much lower resolution for all counties, including rural ones. In particular, it provides the number of households in 5 wealth brackets. In contrast to the district and urban county data, it does not provide the tax amount or wealth held in each of the brackets.

Total tax returns: These are reported for the year 1911 for all rural (and urban) counties in Königlich Preussisches Statistisches Landesamt (1912, p. 588f).

Data on county population structure by size of municipality: Population data for German municipalities is available via Schubert (2022). The data are mainly sourced from the official census published as: Kaiserliches Statistisches Amt (1915): 'Die Volkszählung im Deutschen Reich: Am 1. Dezember 1910.' *Statistik des Deutschen Reichs*, Band 240), Puttkammer & Mühlbrecht, Berlin.

Data on savings accounts with small deposits Königlich Preussisches Statistisches Landesamt (1915a) provides tabulated data on the number of saving deposits by the size of the deposit. We define 'small deposits' as those up to 1,500 Marks, thereby covering the following brackets: 0-60 Marks, 60-150 Marks, 150-300 Marks, 300-600 Marks, 600-1500 Marks. For simplicity, we assume that an individual in a given bracket holds the mean of the lower and the upper bound of the bracket. The data are available at the district level.

A.3.2 Estimation strategy

The main challenge is to generate wealth tabulations at the county level. This requires us to estimate the wealth of those being assessed and those not being assessed for the wealth tax. We structure the discussion of our estimates below along these two categories of households.

Those not assessed for the wealth tax Those assessed for the wealth tax made up 13.4% of Prussian tax units/household heads (Albers et al., 2022, Appendix p. 7). 86.6% of the population was not assessed as their wealth did not pass the threshold of 6,000 Marks. Even though they only held a small share in the total (9% of the total wealth in Prussia), it is important to account for the geographical distribution of this wealth. How much did these households hold in a given county? Fortunately, contemporaries faced a similar problem when trying to estimate the evolution of national wealth over time. Among them, Biedermann (1918) employs the evolution of small savings accounts to estimate the wealth of those not assessed for the wealth tax. We adopt this strategy across space.

In particular, we employ the difference in the amount of deposits per household held in small savings accounts across districts. We do so in the following five steps:

1. We normalize the total amount of small savings in district d by the number of households N_d in d to arrive at $s_d = S_d/N_d$.

2. We calculate the proportion of s_d in the mean across all districts, i.e., $s_d/\overline{s_d}$.
3. We scale the national average wealth of those not assessed ($w_n^{\text{not assessed}} = \frac{W_n^{\text{not assessed}}}{N_n^{\text{not assessed}}}$) with this proportion: $\widehat{w_d^{\text{not assessed}}} = w_n^{\text{not assessed}} \times s_d/\overline{s_d}$.
4. We estimate total wealth by those not assessed in a given county c by $\widehat{W_c^{\text{not assessed}}} = N_c^{\text{not assessed}} \times \widehat{w_d^{\text{not assessed}}}$.[2]
5. We normalize the amounts held in each county proportionally such that they are consistent with the national total.

Those assessed for the wealth tax Figure A2 summarizes schematically the estimation procedure. We possess detailed tax data for urban counties, in which those paying taxes are tabulated in 48 bins ranging from 6,000-8,000 Marks to the final bin 'above 500,000 Marks'. For these individuals, we calculate their wealth by dividing the tax amount by the tax rate (discussed and documented in Albers et al., 2022). The source also covers households that are assessed, but pay a reduced rate or are even exempt. For those, we follow the estimates by Biedermann (1918), also used by Albers et al. (2022).

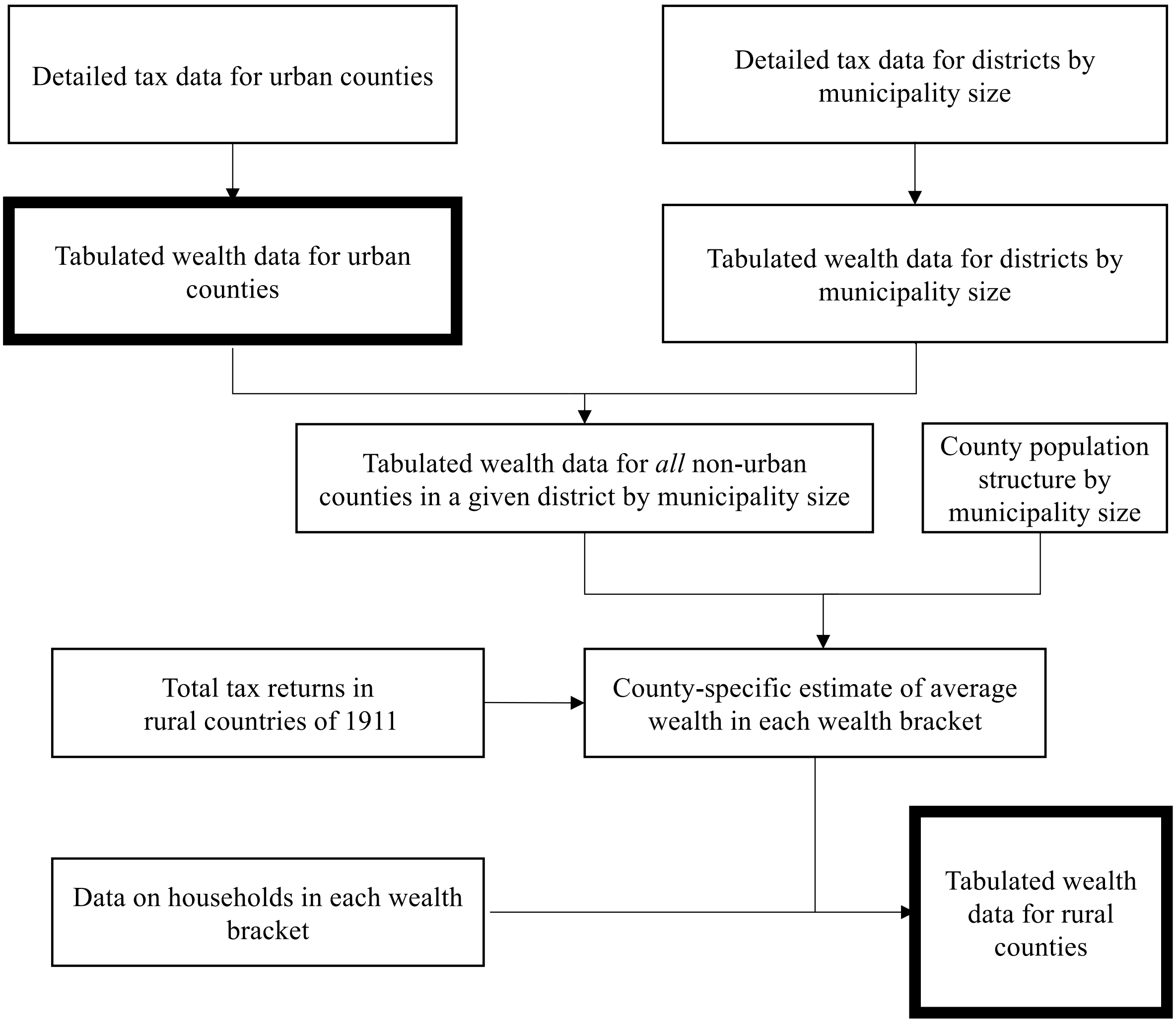
Unfortunately, the tabulated data for rural counties are less rich. They only have five brackets and, while they report the number of individuals in each of them, they do not report the tax amount or wealth in each of these brackets. This requires us to implement a more complex approach to generate tabulations for them. In particular, we combine detailed district-level data, the data for the urban counties, and data on the population structure of counties.

The district tabulations for the wealth tax are as detailed as those for the urban counties. Additionally, the district-level data are classified by the size of municipalities. For example, the data allows to describe the total amount of taxes in a district T_d as the sum of those levied in municipalities with more than 2,000 inhabitants (T_{dra}), those with less than 2,000 inhabitants (T_{drb}), and urban municipalities (T_{du}).

After estimating the wealth from the taxes for all these district types by dividing the taxes paid by the tax rates, we begin by subtracting the wealth of the urban counties from the district data summarizing the wealth for urban municipalities. This is necessary as all municipalities in urban counties are urban, but urban municipalities exist outside of urban counties (e.g., a small city in a rural county). For each district, we identify the wealth held in urban municipalities in rural counties in a district (dru) in each wealth bin b by subtracting the sum of wealth held in the urban counties in that respective bin ($\sum W_{cu}^b$), i.e., $W_{dru}^b = W_{du}^b - \sum W_{cu}^b$. Let us also note that the wealth held in all rural counties in bin b

[2] We thereby assume that the wealth of those not assessed does not vary across counties within a given district. This simplifying assumption is borne out by the lack of tabulated data on small savings accounts at the county level. However, given the limited amount of wealth owned by this group generally, this simplification is unlikely to affect the distributional estimates in a meaningful way.

Figure A2: Inferring wealth tabulations for those assessed for the wealth tax at the county level



in a given district d is given as $W^b_{dr} = W^b_{dru} + W^b_{dra} + W^b_{drb}$. This way we arrive at the box 'Tabulated wealth data for *all* non-urban counties in a given district by municipality size'.

Given the number of taxpayers living in each district in municipalities of the types dra (more than 2,000 inhabitants), drb (those with less than 2,000 inhabitants), and dru (urban municipalities, but not urban county) is known, we can also calculate the average wealth in a given district for each bracket b for the given type: w^b_{dru}, w^b_{dra}, w^b_{drb}. Since the tabulation for the rural counties provides only four relevant brackets (6,000-20,000; 20,000-52,000; 52,000-100,000, above 100,000), we calculate these averages for those rather than the 48 brackets available for the urban counties and the districts.

We then leverage the data on the population structure of the counties. For each rural county cr, we calculate the share of the population that lives in municipalities with more than 2,000 inhabitants ($\text{Share}_{cr,mra}$), municipalities with less than 2,000 inhabitants ($\text{Share}_{cr,mrb}$), and urban municipalities ($\text{Share}_{cr,mu}$). Combining these with the averages at the district level allows us to estimate the county-specific average wealth of those paying the wealth tax for the bins $b = \{[6,000 - 20,000), [20,000 - 52,000), [52,000 - 100,000)\}$ as weighted averages: $w^b_{cr} = \text{Share}_{cr,mra} \times w^b_{dra} + \text{Share}_{cr,mrb} \times w^b_{drb} + \text{Share}_{cr,mu} \times w^b_{dru}$. By multiplying w^b_{cr} with the respective number of households N^b_{cr}, we arrive at the wealth in each of these brackets in each rural county, W^b_{cr}.

We calculate the wealth held in the remaining wealth bin, 'above 100,000 Marks', by making use of additional data: the total receipt of wealth taxes in each rural county over the period 1907-1911 (as these data do not exist for 1914). In particular, we estimate the share of the total taxes that each county had within the district.[3] We then distribute the wealth in the bin held within rural counties in a given district ($W^{b(\text{'above 100,000 Mark'})}_{dr}$) using these shares. Finally, we add the number of households in each county to arrive at the tabulated wealth data for rural counties for all households.

Estimating top shares from tabulations To estimate top shares from the tabulations, we apply the Generalized Pareto Interpolation method by Blanchet et al. (2022).

External validation of top wealth share estimates We provide two validation exercises for the estimated top wealth shares. First, Table A 3 relates the estimates of the top percentile's wealth share to the concentration of land holdings (land inequality data are from Galloway, 2007). Reassuringly, there exists a strong relationship even after controlling for soil suitability, temperature, and district fixed effects.

Table A 3: Land inequality and wealth inequality

Dep. var.: Top 1% share	(1)	(2)
Concentration of large landowners (in %)	0.374**	0.390*
	(0.179)	(0.208)
Mean dependent variable	39.48	39.48
SD dependent variable	11.06	11.06
District FE	✓	✓
Geographic controls		✓
R-squared	0.58	0.64
Observations	575	575

Notes: Unit of observation: county. Concentration of large landowners in 1907 based on Galloway (2007) is defined as the share of landowners who own more than 100 hectares relative to the total number of landowners following Cinnirella and Hornung (2016). Controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

Second, we leverage the millionaires data described in Section A.2. We expect a positive correlation between the top-1% share and the number of millionaires as well as their wealth.

[3] We deduct the taxes paid in the bin '52,000-100,000 Mark' as of 1914 to make the estimate more precise. We do not subtract the other two wealth bins (6,000-20,000; 20,000-52,000) as they contain also individuals paying a reduced rate or are exempted.

Table A 4 shows that this prediction is indeed borne out by the data.

Table A 4: Millionaires, their wealth, and wealth inequality

Dep. var.: Top 1% share	(1)	(2)	(3)	(4)
Number of millionaires (log)	0.517***	0.401***		
	(0.102)	(0.080)		
Total wealth millionaires (log)			0.464***	0.369***
			(0.097)	(0.075)
Mean dependent variable	39.48	39.48	39.48	39.48
SD dependent variable	11.06	11.06	11.06	11.06
District FE	✓	✓	✓	✓
Geographic controls		✓		✓
R-squared	0.61	0.66	0.60	0.66
Observations	575	575	575	575

Notes: Unit of observation: county. For counties without millionaires, we replace the independent variable number with ln(.0001). Controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

A.4 Other data

Industry employment Occupation by sector on a county level is reported in a census conducted by the German national statistics. We divide employment in industry by total employment.

Source: Statistik des Deutschen Reichs (1909b).

Distance to carboniferous strata We calculate the distance from a county's centroid to carboniferous strata, that is strata in which coal can be found. For the calculation, we rely on a geological map provided by the German Federal Institute for Geosciences and Natural Resources (BGR).

Source: Asch (2005).

Net-migration We decompose total population growth into its components (1) migration and (2) "natural" population growth. We calculate the difference between the reported "natural" population growth based on births and deaths for each year and the actual population growth between two censuses conducted every five years provided by the German national statistics.

Source: Galloway (2007).

Land inequality Concentration of large landowners is defined as the share of landowners who own more than 100 hectares relative to the total number of landowners following Cinnirella and Hornung (2016).

Sources: For 1816 Cinnirella and Hornung (2016); for 1907 Galloway (2007).

Urban population Urban population relative to total population based on the Prussian census from 1816.

Source: Cinnirella and Hornung (2016).

Protestantism Number of Protestants on a county level is reported in a census conducted by the German national statistics in 1910. We divide Protestants by total population.
Source: Galloway (2007).

Ethnicity Number of German speakers on a county level is reported in a census conducted by the German national statistics in 1900. We divide German speakers by total population.
Source: Galloway (2007).

Distance to the nearest river The distance in hundred of kilometers of a country's centroid to the nearest navigable river is constructed using shape files on navigable rivers in the German Lands as of 1850. To improve accuracy, we add shape files on rivers from the European Environment Agency.

Sources: Kunz (2001); European Environment Agency (2012).

Average temperature and precipitation Average temperature and precipitation data (in Celsius and mm respectively), are taken from WorldClim version 2.1 climate data. The dataset provides average monthly temperatures for the period 1970-2000 at a geospatial resolution of 30 arc seconds. We take the average over this time horizon for the county polygons.

Source: Fick and Hijmans (2017).

Soil suitability Average suitability of the soil for each Prussian county is attained using the crop suitability index for current cropland in grid cells for the time period of 1981-2010 for wheat under rainfed conditions and low input level.

Source: Food and Organization (2021).

Non-housing rate of return at city level The official tax statistics provide income and wealth tax returns of those above the income threshold of 3,000 Marks total income for 95 cities. The statistics provide three types of wealth and the generated income: capital assets ("Kapitalvermögen"), housing ("Grundvermögen"), and business assets ("Vermögen aus Handel, Gewerbe, Bergbau"). We calculate the non-housing rate of return by dividing the respective income by wealth minus income.

Source: Königlich Preussisches Statistisches Landesamt (1908).

Firm size Employment and the number of firms on a county level are reported in a census conducted by the German national statistics. We divide employment by number of firms.

Source: Statistik des Deutschen Reichs (1909).

B Additional results and robustness

B.1 Tables

Table B.1: Carboniferous strata and pre-industrialization outcomes (in 1816)

Dep. var.:	Urban population (in %) (1)	Large land owner (in %) (2)
Least distance to carbon field (in log)	-1.589 (1.425)	0.084 (0.072)
Mean dependent variable	25.75	1.72
SD dependent variable	20.25	2.09
District FE	✓	✓
Geographic controls	✓	✓
R-squared	0.26	0.69
Observations	280	267

Notes: Unit of observation: county. Here, we adjust the county borders to 1800 to include the pre-industrialization outcomes based on Becker et al. (2014). Further controls include: Distance to the nearest river, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

Table B.2: Industrialization and wealth inequality, first stage

Dep. var.: Employment share industry	(1)	(2)	(3)
Least distance to carbon field (in log)	-6.094*** (0.901)	-4.449*** (1.082)	-4.397*** (0.850)
Mean dependent variable	31.35	31.35	31.35
SD dependent variable	16.66	16.66	16.66
District FE		✓	✓
Geographic controls			✓
R-squared	0.32	0.52	0.61
Observations	575	575	575

Notes: Unit of observation: county. Further controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

Table B.3: Industrialization and wealth inequality, reduced form

Dep. var.:	Top 1%			Top 1% / top 10%		
	(1)	(2)	(3)	(4)	(5)	(6)
Least distance to carbon field	-1.752**	-1.196**	-1.932***	-1.961***	-0.946*	-1.559***
(in log)	(0.745)	(0.583)	(0.463)	(0.609)	(0.509)	(0.446)
Mean dependent variable	39.48	39.48	39.48	49.17	49.17	49.17
SD dependent variable	11.06	11.06	11.06	9.89	9.89	9.89
District FE		✓	✓		✓	✓
Geographic controls			✓			✓
R-squared	0.06	0.59	0.66	0.09	0.59	0.64
Observations	575	575	575	575	575	575

Notes: Unit of observation: county. Further controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

Table B.4: Industrialization and the fattening of the wealth distribution's top-tail, OLS

Dep. var.:	Top 1% share	Top 10 % share	Top 1% / top 10%
	(1)	(2)	(3)
Employment share industry	0.251***	0.257***	0.156***
	(0.028)	(0.030)	(0.027)
Mean dependent variable	39.48	79.27	49.17
SD dependent variable	11.06	9.08	9.89
District FE	✓	✓	✓
Geographic controls	✓	✓	✓
R-squared	0.70	0.72	0.65
Observations	575	575	575

Notes: Unit of observation: county. Controls include distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses and are clustered at the district level. * p<0.1, ** p<0.05, *** p<0.01.

Table B.5: Industrialization and wealth inequality, role of migration

Dep. var.:	Top 1%		Top 1% / top 10%	
	(1)	(2)	(3)	(4)
Employment share industry	0.397***	0.384***	0.290***	0.289***
	(0.082)	(0.095)	(0.083)	(0.097)
Net-migration (in %)		0.006		0.000
		(0.014)		(0.012)
Mean dependent variable	39.04	39.04	49.02	49.02
SD dependent variable	10.19	10.19	9.04	9.04
District FE	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓
F-stat excluded instrument	38.91	31.25	38.91	31.25
Observations	441	441	441	441

Notes: Unit of observation: county. Here, we adjust the borders to 1871 to include net-migration between 1871 and 1910. That is the reason for the lower number of observations. Further controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

Table B.6: Industrialization and wealth inequality, role of land inequality

Dep. var.:	Top 1%	Top 1% / top 10%
	(1)	(2)
Employment share industry	0.411***	0.345***
	(0.098)	(0.112)
Large landowners (in %) × rural county (dummy)	0.778*	0.502
	(0.408)	(0.401)
Rural county (dummy)	-3.313*	-1.198
	(1.889)	(2.043)
Concentration of large landowners (in %)	0.111	-0.037
	(0.233)	(0.198)
Mean dependent variable	39.48	49.17
SD dependent variable	11.06	9.89
District FE	✓	✓
Geographic controls	✓	✓
F-stat excluded instrument	25.35	25.35
Observations	575	575

Notes: Unit of observation: county. Further controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

Table B.7: Industrialization and wealth inequality, role of Protestantism and ethnicity

Dep. var.:	Top 1%	Top 1% / top 10%
	(1)	(2)
Employment share industry	0.438***	0.353***
	(0.093)	(0.105)
Share Protestants in %	0.002	0.001
	(0.014)	(0.013)
Share German in %	-0.001	0.003
	(0.012)	(0.012)
Mean dependent variable	39.48	49.17
SD dependent variable	11.06	9.89
District FE	✓	✓
Geographic controls	✓	✓
F-stat excluded instrument	19.73	19.73
Observations	575	575

Notes: Unit of observation: county. Further controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

Table B.8: Industrialization and wealth inequality, sample splits

Sample:	w/o Rhine Province		w/o Arnsberg	
Dep. var.:	Top 1%	Top 1% / top 10%	Top 1%	Top 1% / top 10%
	(1)	(2)	(3)	(4)
Employment share industry	0.379***	0.309***	0.456***	0.354***
	(0.060)	(0.066)	(0.093)	(0.108)
Mean dependent variable	39.06	48.53	39.06	48.80
SD dependent variable	10.30	9.64	11.08	9.94
District FE	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓
F-stat excluded instrument	22.36	22.36	26.70	26.70
Observations	494	494	527	527

Notes: Unit of observation: county. Further controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level. * p<0.1, ** p<0.05, *** p<0.01.

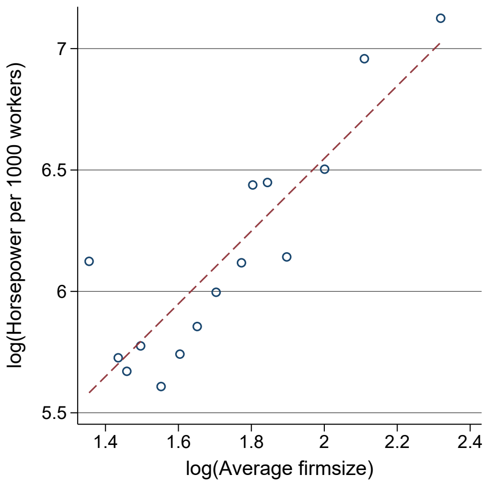
Table B.9: Industrialization and wealth inequality, spatial clustering

Dep. var.:	Top 1%			Top 1% / top 10%		
	Baseline	100km	200km	Baseline	100km	200km
	(1)	(2)	(3)	(4)	(5)	(6)
Employment share industry	0.439***	0.439***	0.439***	0.355***	0.355***	0.355***
	(0.093)	(0.080)	(0.062)	(0.102)	(0.094)	(0.080)
Mean dependent variable	39.48	39.48	39.48	49.17	49.17	49.17
SD dependent variable	11.06	11.06	11.06	9.89	9.89	9.89
District FE	✓	✓	✓	✓	✓	✓
Geographic controls	✓	✓	✓	✓	✓	✓
F-stat excluded instrument	575	575	575	575	575	575

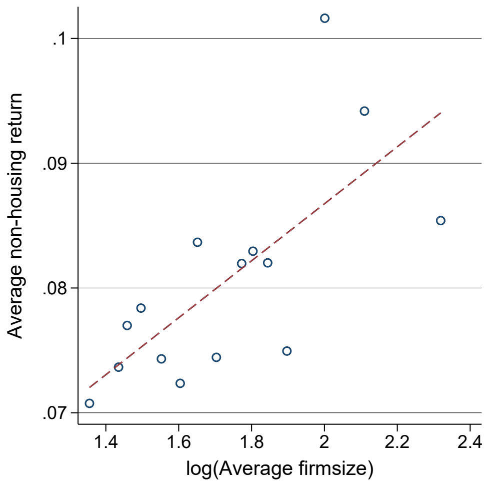
Notes: Unit of observation: county. Further controls include: Distance to the nearest river, soil suitability, latitude, longitude, average temperature, and average precipitation. Standard errors in parentheses. Standard errors are clustered on the district level (columns 1 and 4). In columns 2, 3, 5, and 6, we vary the distance cutoff, beyond which the correlation between error term of two observations is assumed to be zero. In columns 2 and 5, this cutoff is 100km and in columns 3 and 6, this cutoff is 200km. * p<0.1, ** p<0.05, *** p<0.01.

B.2 Figures

Figure B.1: Economies of scale at the city level



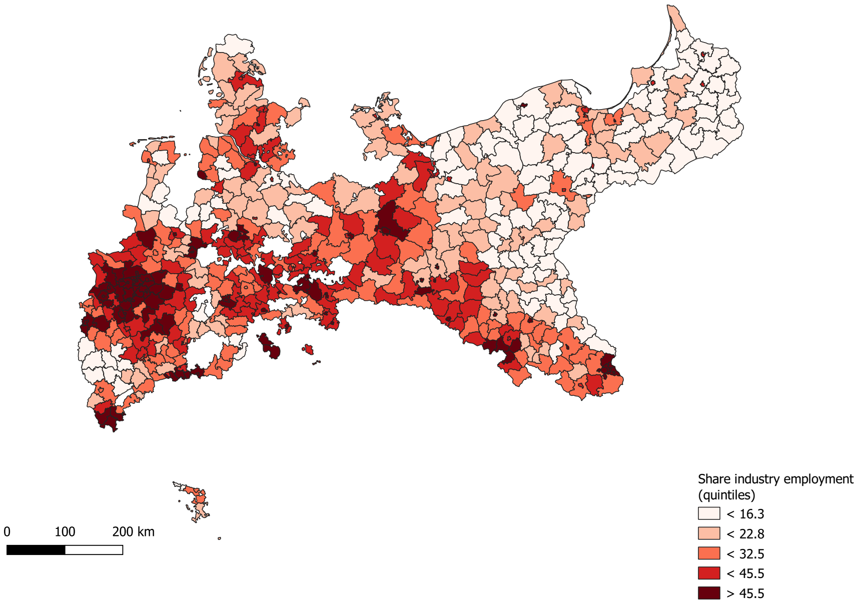
(a) Capital intensity and firm size



(b) Non-housing returns and firmsize

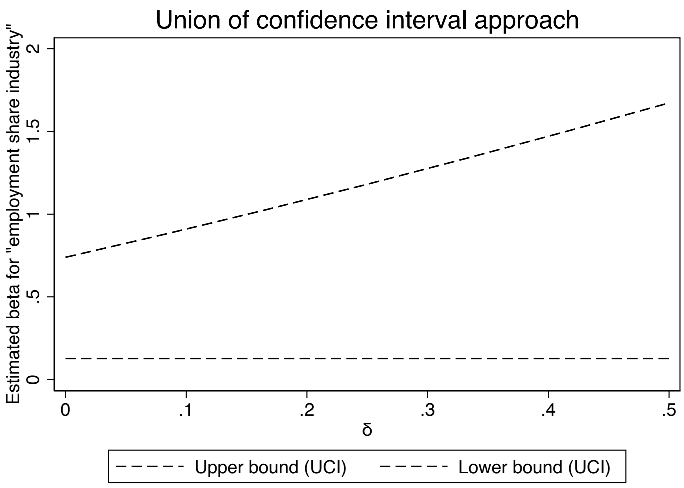
Notes: Each dot represents in the binned scatter graph represents close to 6 cities. The non-housing returns are calculated on the basis of income and wealth tax returns of those above the income threshold of 3,000 Marks total income. The data refer to 1908 and are from Königlich Preussisches Statistisches Landesamt (1908). Data on the average firm size by city is from Statistik des Deutschen Reichs (1909) and refers to year 1907, while the data on horsepower refers to the year 1904 and are from Königlich Preussisches Statistisches Landesamt (1907).

Figure B.2: Industry employment across Prussian counties, 1907



Notes: Own calculations. For detailed data sources, see Appendix A.4.

Figure B.3: Boundaries of IV estimates



Notes: This graph shows Conley et al. (2012) union of confidence interval boundaries for IV estimates of the impact of employment share in industry on the top 1% wealth share. The specification corresponds to Table 1, column 6. We plot the estimated 95 percent confidence interval bounds for σ (0, 0.5). To allow for better comparison, we standardize our variables.

C Model

C.1 Individual optimization problem

We follow Benhabib et al. (2011, pp. 128-129) in formulating individual preferences as follows. Consider an individual that represents generation n of a given dynasty. At the beginning of her life, and given her drawn and thereafter fixed rate of return r_n and earnings y_n, she decides on a consumption path to maximize:

$$\int_0^T e^{-\rho\tau} u\left(c_n(\tau)\right) d\tau + e^{-\rho T} \phi\left(w_{n+1}(0)\right)$$

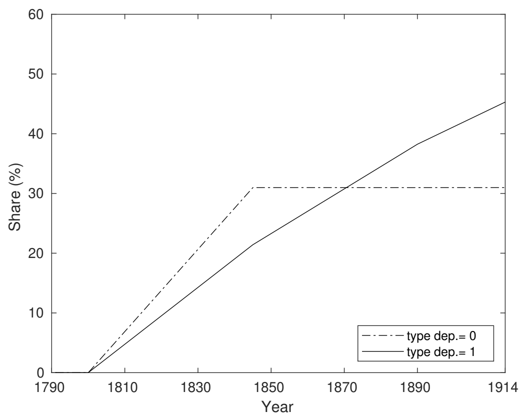
subject to

$$\begin{aligned} \dot{w}_n(\tau) &= r_n w_n(\tau) + y_n - c_n(\tau), \\ w_{n+1}(0) &= w_n(T), \end{aligned}$$

where $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ with $\sigma \geq 1$ is the utility from consumption, $\phi(w) = \chi \frac{w^{1-\sigma}}{1-\sigma}$ with $\chi > 0$ is the one-time utility from leaving a bequest w to the next generation, $\rho > 0$ is the discount rate, and $w_n(\tau)$ denotes her wealth as a function of her age. After the choice of the optimal consumption and saving plan, $w_n(\tau)$ is deterministic and can be expressed in closed form, which the authors exploit to derive the intergenerational evolution of dynastic wealth, as expressed in equation (4).

C.2 Modeling dynastic type dependence

Figure C.1: The share of industrialists among wealth-holders



Notes: The trajectory labeled "type dep = 0" reports the share of dynasties in industry for different points in time in scenarios 1 and 3 (see Table 3), i.e., those without dynastic type dependence. The trajectory labeled "type dep = 1" reports this share for scenarios 2 and 4, i.e., those with dynastic type dependence. See Section 3.1 for a description of the underlying scenario-specific sector transition probabilities ω_i and ϕ_i.

C.3 Calibration

C.3.1 Parameters

Table C.1: Parameterization overview

Parameter	Value	Description	Source / Target
Population			
T	45	Working life span	Benhabib et al. (2011). Moreover, exp. remaining lifetime at age 10 in Sweden (1810-1910), based on Roser (2023): 50.13 (avg.)
ξ	0.6	Hand-to-mouth share	Top 40% wealth share in 1850s Hesse-Cassel based on Wegge (2021): 96.8%
π^*	0.453	Share industrial dynasties 1914 (scenarios 2 and 4)	Share of wealth taxpayers living in cities based on Königlich Preussisches Statistisches Landesamt (1915b)
ω_1, ω_2, ϕ_1, ϕ_2	See text	Conditional sector transition probabilities	Calculated according to procedure described in main text
Preferences			
σ	1.5	CRRA parameter	See, e.g., Hubmer et al. (2021) and De Nardi (2004)
ρ	0.02	Discount rate	Assumed (no borrowing constraint)
χ	7.25	Bequest motive	Jointly calibrated in matching preindustrial moments (see text)
Income			
g', g	0.0082	Growth rate (init. income and over lifetime, resp.)	Geometric average based on Pfister (2019)
RoR			
θ	0.0215	Scale dependence (pp. spread 60th to top pct.)	Benhabib et al. (2019)
κ_1	0	Persistence (lowest RoR state)	Jointly calibrated in matching preindustrial moments (see text)
κ_2	0.05	Persistence (highest RoR state)	Jointly calibrated in matching preindustrial moments (see text)

C.3.2 Rate of return processes

In our model, individuals draw from a distribution of lifetime rates of return in either the agricultural or industrial sector. The drawn rate reflects the geometric mean rate of return over the lifetime. As the share of industrialists increases over time, the two-sector structure of our model gives rise to a changing aggregate distribution in lifetime rates of return over time – a difference to existing models where this distribution is fixed. Below we describe this feature and show how the model produces a standard deviation of lifetime returns that is in the ballpark of existing studies without such a sectoral structure (e.g., Bach et al., 2020).

Calibration to match the sector-specific distributions In our simulation, we leverage the observed differences between millionaires' agricultural and industrial fortunes with respect to their rates of return (Figure 2(b)). To do so, we make the assumption that the cross-sectional differences of the moments in the millionaires list reflect the *long-run* differences in the rate of return moments in the lifetime sense of the model. We show further below that the implied aggregate mean and standard deviation in the lifetime rates of return are approximately in line with existing estimates of generational rate of return moments, suggesting that this assumption is warranted. We now provide details on how we set the rate of return processes given the sector-specific rate of return moments from the millionaires list.

Following the formulation of the rate of return process by Benhabib et al. (2011), we define $({r^L}_n)_n$ as well as $({r^1}_n)_n$ and $({r^2}_n)_n$ as discrete Markov chains. We follow the authors in how we set the baseline transition matrix P, which in our case is the transition matrix regarding S^L in $({r^L}_n)_n$ and – by its inclusion in the definition of Q_i in equation (6) – the transition matrix regarding the absorbing subspace S^L in Scenario 2:

$$P = \begin{pmatrix} .8+\kappa_1 & .12-\kappa_1/3 & .07-\kappa_1/3 & .01-\kappa_1/3 \\ .8 & .12 & .07 & .01 \\ .8 & .12 & .07 & .01 \\ .8-\kappa_2/3 & .12-\kappa_2/3 & .07-\kappa_2/3 & .01+\kappa_2 \end{pmatrix}.$$

Based on the i.i.d. transition probabilities $\{.8, .12, .07, .01\}$, we choose $S^L = \{.0360, .0475, .0587, .1184\}$ to match the first and second moment of the rate of return distribution observed for landowners in the list of Prussian millionaires (mean=3.98%, SD=1.03%), and $S^I = \{.0544+\iota, .0757+\iota, .0930+\iota, .1791+\iota\}$ to match the second moment of the industrialists' distribution of rates of return (SD=1.65%). We choose $\iota = 0.0040$ to slightly increase the mean rate of return in industry to 6.494%— the average return for equity in Germany between 1880 and 1914 according to Jordà et al. (2019)—to account for capital gains that are unaccounted for in the millionaires data (see Section 2.2). We choose to base our upward correction on this relatively low estimate of the mean (annual) profitability of industrial capital compared with other estimates (see Section 2.2) because the geometric average is dampened by annual dispersion in the generational perspective. For lack of appropriate data, we refrain from making a similar correction to the standard deviation and simply stick to that observed in the millionaires list.

In choosing the rate of return states, we operate with the probabilities from the i.i.d. case to match rate of return moments because we can only calibrate the model to the preindustrial distribution, involving a grid search for κ_1 and κ_2, given set rate of return states. However, the resulting values for κ_1 and κ_2 affect the rate of return moments only marginally. As a result, $({r^L}_n)_n$ has a mean of 3.99% and a standard deviation of 1.04% and the distribution faced by industrial dynasties if the industrial sector is absorbing and lacks scale dependence (Scenario 2) has a mean of 6.50% and a standard deviation of 1.67%. While the introduction of scale dependence in scenarios 3 and 4 preserves the mean of rates of return available in industry – see equations (7) and (8) – it increases the standard deviation of rates of return in industry to 1.78%. This effect is inextricably linked to an addition of scale dependence and thus an effect that we aim to capture in our counterfactual analysis of the long-run impact of scale dependence.

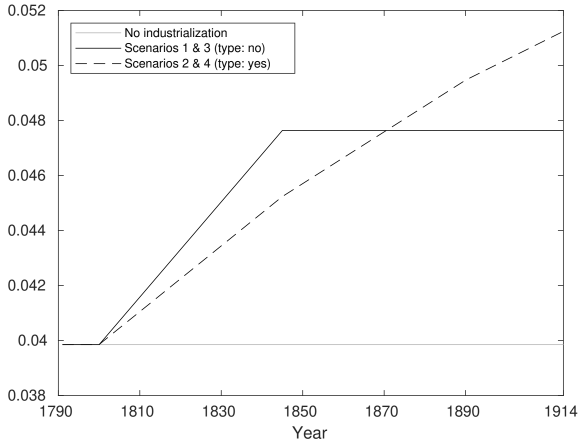
Last, note that the formulation of the transition matrices of $({r^L}_n)_n$ as well as $({r^1}_n)_n$ and $({r^2}_n)_n$ ensures that the original persistence on the sector-specific lowest and highest rate of return state induced by the parameters κ_1 and κ_2 is maintained throughout the shock in t_0. For each dynasty, we take the index of its state in t_0 and determine its transition probabilities regarding the new states from the new transition matrix's row with the same index.[4] Maintaining this persistence appears desirable since the rate of return heterogeneity also reflects idiosyncratic differences between dynasties regarding investment skill and knowledge. These features are arguably partially transmitted across generations and abstract enough to persist beyond the shock by being beneficial in the new investment circumstances as well.

Aggregate moments of the return distribution in a model with sectoral change A feature of our model is its two-sector structure and the gradual transition of dynasties from the old to the new sector. This gives rise to a changing distribution of rates of return over time, reflecting the structural change of the economy occurring at the time. In contrast to existing models which consider a stationary wealth distribution resulting from a constant rate of return process, we thus consider a process of transition, both in the distribution of rates of return and wealth.

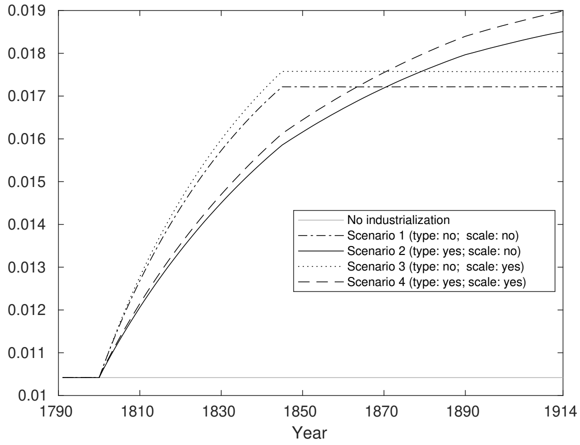
Figure C.2 illustrates this. In all simulated scenarios of the model, both the mean (left) and the standard deviation (right) of lifetime rates of return increase after the onset of indus-

[4] For example, consider a dynasty that is in the highest state of S^L in t_0 and whose rate of return process is then shocked. Focusing on the scenarios without scale dependence, i.e., scenarios 1 or 2 (or, equivalently, considering just the non-scale dependent part of processes 3 and 4), its post-shock transition probabilities are then given by the fourth row of Q_i and refer to the new state space S^e. Notably, the probability of reaching again the highest rate of return state, either of agriculture or industry, is again increased by κ_2, while the overall probability of switching to industry will be given by ϕ_i (see equation (6)).

Figure C.2: The changing aggregate distribution of lifetime rates of return



(a) Cross-sectional mean



(b) Cross-sectional standard deviation

Notes: The trajectories refer to the simulated scenarios listed in Table 3 and represent averages over 200 iterations. The abbreviations "type" and "scale" stand for the features of dynastic type dependence and scale dependence, respectively.

trialization. The scenarios without dynastic type dependence result in constant aggregate moments after one generation, when all cohorts have drawn from the new respective rate of return process. This reflects that in these scenarios the share of industrial dynasties is constant one generation after t_0 (Figure C.1).

By contrast, the scenarios with dynastic type dependence exhibit a strictly monotonic increase in both mean and standard deviation of lifetime rates of return throughout the simulated period. The reason for this is that industry is absorbing, such that once dynasties have transitioned to industry they can only draw rates of return in industry in the future. In the very long run, as the share of industrial dynasties approaches one, the aggregate moments in these scenarios will converge to the sector-specific moments of the industrial sector, which are constant throughout with a mean of 6.50% and a standard deviation of 1.67% (or 1.78% in the case of scale dependence). In the long-run, the curve of the standard deviation thus follows an inverse U.[5]

Comparison with estimates in the literature As a plausibility check, we relate the aggregate moments shown in Figure C.2 to corresponding estimates in the literature. In particular, the upper part of Table C.2 reports the moments from the simulations of our model in 1800 (equal in all four scenarios) and at the end of our simulation period (Scenario 4 with scale and dynastic type dependence). The lower part of the table reports estimates from the literature. They are comparable to our estimates in that they all refer to the moments of the

[5]Robinson (1976) provides a technical discussion of the emergence of the inverse U curve in two-sector models due to such a composition effect (specifically the Kuznets curve). In a setting (as in our case) with two sectors and an increasing relative share of the population in the sector with a higher first and second sector-specific moment, the evolution of the second moment of the aggregate distribution will follow an inverse U, largely driven by the difference in means of the two sectors.

distribution of lifetime returns (or, more specifically, geometric average rates of return over a generation). Where we report estimates of the standard deviation in brackets, these refer to a conversion of the values in models with $T = 36$ to an equivalent in our model with $T = 45$ (as also in the original model by Benhabib et al., 2011).[6] Where available, we use estimates for the top-50% of the distribution as this is closest to our model. A main general difference between our estimate and those reported in the table, however, is the geographical area and time periods these estimates refer to.

Table C.2: The distribution of lifetime rates of return

Mean	SD	Source	Notes
3.99	1.04	This model	1800
5.12	1.90	This model	1914 (Scenario 4)
8.93-9.25	2.23-3.25	Benhabib et al. (2011) (assumed)	Following Angeletos (2007)
3.06	2.69 (2.41)	Benhabib et al. (2019) (structural)	MSM estimate
4.93†	2.19 (1.96)	Bach et al. (2020) (empirical)	Top 50%; pre-tax gross wealth
4.88†	4.88 (4.36)	Bach et al. (2020) (empirical)	Top 50%; pre-tax net wealth
3.7	3.6	Fagereng et al. (2020) (empirical)	Ind. RoR fixed effects; net wealth, after tax

Notes: For the two sources that estimate the standard deviation based on the assumption of the length of a generation T of 36 years (Benhabib et al., 2019, and Bach et al., 2020), we provide in brackets a simple approximation of an estimate based on assuming $T = 45$. We calculate this estimate treating the generational rate of return as the arithmetic mean of independent and identically distributed annual rates of return. Fagereng et al. (2020) report their values based on individual rate of return fixed effects without an assumption regarding T. †: we calculate the total return by summing the return of a risk-free asset (1.5pp) and the excess return, both reported in Bach et al. (2020, p. 2716; Table 10).

The mean of lifetime rates of return in our model in 1914 is substantially lower than the value used in the original calibration by Benhabib et al. (2011). However, it is somewhat larger than recent empirical and structural estimates. The two empirical studies employ Scandinavian post-2000 data (Bach et al., 2020; Fagereng et al., 2020) whereas the structural estimate by Benhabib et al. (2019) employs US data for 2007. The estimated mean of lifetime returns in these studies varies between 3.06pp and 4.88pp for net wealth. Despite referring to a very different time period and country, the mean that emerges through our sectoral calibration in 1914 (5.12) is not far off from this upper bound and lies within the range in 1800 (3.99). In sum, our model produces a cross-sectional mean of lifetime rates of return that is well within the ballpark of estimates for more recent time periods.

With respect to the dispersion in lifetime rates of return, the standard deviation in our model is at the lower end of existing estimates. It is not reasonable to compare the 1800 value to those from industrialized OECD countries as the assets available to investors (land or government bonds) differed too much from today's portfolios across the wealth distribution.

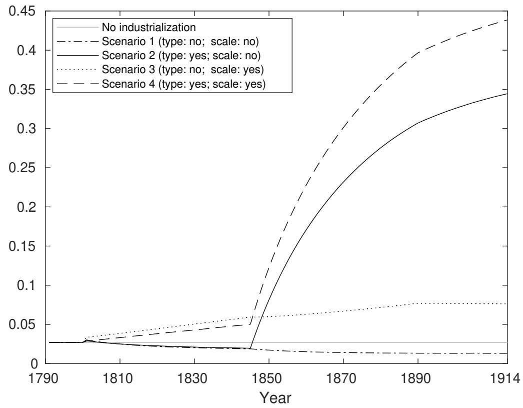
[6] In particular, we assume for this conversion that the fixed lifetime rate of return, for which the standard deviation is calculated, represents an arithmetic average of independent and identically distributed yearly rates of return.

However, our value of 1.90 for 1914 is quite close to the estimate for gross wealth by Bach et al. (2020) and not too far from the estimate by Benhabib et al. (2019) for net wealth in the US.[7] Note that the difference between Bach et al. (2020) and Fagereng et al. (2020) is quite substantial given that they refer to Sweden and Norway, respectively, for approximately the same time period. It would be even larger using the estimates by Bach et al. (2020) for the whole population and net wealth after tax (3.6 vs. 5.82). While at the lower end, our model produces a standard deviation that is comparable to those used in the literature.

In sum, our comparison with existing estimates shows that we are in the ballpark of estimates that are used to calibrate the type of model we employ. Also note that, analytically, our focus is on how sectoral *differences* in return opportunities lead to wealth concentration. As long as these differences in the moments in the sector-specific calibration of the model are maintained, our *qualitative* results are likely unaffected by varying the mean or standard deviation in the sectors as long as this is done to a similar degree.

Intergenerational persistence in rates of return For each scenario and each point in time, we calculate how rates of return of alive households vary with rates of return of their respective previous generation. Figure C.3 plots the aggregate intergenerational autocorrelation of lifetime rates of return.

Figure C.3: Autocorrelation coefficient (current and previous generation)



Notes: The trajectories refer to the simulated scenarios listed in Table 3 and represent averages over 200 iterations. The abbreviations "type" and "scale" stand for the features of dynastic type dependence and scale dependence, respectively.

Two observations are worth noting. First, scale dependence in industry increases the overall intergenerational persistence in lifetime rates of return from the start. In the transition from agriculture to industry, larger agricultural fortunes—which also tend to have accumulated with a larger rate of return—are rewarded with a scale dependence boost once they have transitioned to industry.

Second, dynastic type dependence has a large impact on the intergenerational persistence

[7]The model by Benhabib et al. (2019) excludes borrowing and is calibrated to the US net wealth distribution.

of rates of return in the aggregate, taking effect one generation after the initial shock in t_0. In contrast to scenarios without dynastic type dependence, industrial dynasties are bound to their sector in the next generation, perpetuating their rate of return lead over dynasties remaining in the old sector. Similarly to the long-run evolution of the standard deviation in rates of return, the trajectories of the autocorrelation coefficient for the scenarios with dynastic type dependence follow an inverse U in the long run, as the share of industrial dynasties approaches one and the aggregate impact of the persistent sectoral differences regarding rates of return fades away.

Scale dependence spread We set the scale dependence spread θ to 0.0215. This is the difference in the annual expected rate of return between the 60th and the top percentile in the wealth distribution as estimated by Benhabib et al. (2019) by calibrating their model to the US wealth distribution from 2007 with a moment-matching approach. Crucially, rates of return in their dynastic model are also fixed at birth, such that as in our case the scale dependence spread applies to the geometric mean of returns over a generation.

Our chosen value of 0.0215 is low when compared with recent evidence on scale dependence in *annual* returns provided by Fagereng et al. (2020) and Bach et al. (2020). In estimating scale dependence, both studies account for the effect of type dependence, which could on its own cause a positive correlation of wealth and rate of return, with high types populating higher wealth ranks over time. Fagereng et al. (2020) analyze Norwegian data and capture type dependence by individual fixed effects. Their estimated spread in pre-tax returns, applied to the top 40 percent of the wealth distribution, amounts to 0.0554.[8] Bach et al. (2020) estimate a corresponding spread of 0.0353 using Swedish data and accounting for type dependence by only exploiting variation within twin pairs.

Besides their documentation of scale dependence in annual rates of return, Bach et al. (2020) also estimate the moments of the geometric average of a generational rate of return process spanning 36 years, which is relevant for calibrations of models such as of Benhabib et al. (2011, 2019). They show that such an estimation should consider the negative effect of annual dispersion in rates of return on the long-run geometric mean. Since their panel spans only eight years, they employ a structural model to estimate the moments conditional on the initial position in the wealth distribution. Driven by a large increase in dispersion of annual rates of return at the top of the wealth distribution, they estimate a long-run negative scale dependence at the top starting around the top 5-2.5 percent bracket, with even negative mean excess returns of minus 4.43 percent per year for the top 0.01. Even though in our case we do not see an indication of a substantially higher dispersion in returns of larger compared with smaller industrial fortunes, the argument causes us to stick to the lower estimate by Benhabib et al. (2019) rather than the higher estimates of scale dependence in

[8]The corresponding values for after-tax return, which we regard as less relevant in our context, are 0.1004 and 0.0402, respectively.

annual returns provided by Fagereng et al. (2020) and Bach et al. (2020).

C.3.3 Extended discussion of results from the MSM approach

The fact that the empirical moments are not matched exactly reflects the fact that the parameter values resulting from the MSM procedure are partially corner solutions. In our grid search, we consider values for κ_1 and κ_2 by iterating in steps of 0.01 over the intervals [0,0.02] and [0,0.05], respectively. Relative to the specifications considered by Benhabib et al. (2011), this increases the upper bound of the parameter space of κ_1 from 0.01 to 0.02 and follows the scope of the parameter space regarding κ_2, but adds the intermediate values 0.03 and 0.04. The parameter estimates from our MSM approach are $\hat{\kappa_1} = 0$ and $\hat{\kappa_2} = 0.05$. Nevertheless, we choose not to extend the possible difference between the two parameters. Limiting this difference to the parameter space applied by Benhabib et al. (2011) coincides with our prior that the difference in the persistence of the highest and the lowest rate of return states should not be overstated. Regarding χ, we consider the parameter space [0.25,12], over which we iterate in steps of 0.25. While the estimated bequest parameter $\hat{\chi} = 7.25$ might seem high, the resulting generated aggregate bequest flow measured in terms of the aggregate income is moderate and in fact below the empirical preindustrial value for Sweden. In all of our simulated scenarios, the bequest flow remains close to this initial generated value over the simulated period of industrialization.

C.3.4 Calibration of the income process

Tilly (2010) estimates a Gini coefficient $G = 0.3989$ for the Prussian income distribution in 1852, and a mean income of 41.75 Taler. From these two values we can calculate the parameters μ and σ of a lognormal distribution, by which we approximate that income distribution. From Crow and Shimizu (1987, p. 11) we then know that $G = 2\Phi\left(\frac{\sigma}{\sqrt{2}}\right) - 1 \iff \sigma = \sqrt{2}\Phi^{-1}(0.5(G+1))$, where Φ is the distribution function of the standard normal distribution. Since the rth moment about the origin of a lognormally distributed variable X is given by $E(X^r) = \exp(r\mu + 0.5r^2\sigma^2)$ (p. 9) we have $m = \exp\left(\mu + 0.5\sigma^2\right) \iff \mu = \ln(m) - 0.5\sigma^2$.

Given μ and σ, we then discretize the distribution according to the procedure proposed by Miller and Rice (1983)[9] to obtain a discrete distribution with four values (26.018, 109.633, 422.655, and 1780.958) and corresponding probabilities (0.818, 0.181, 0.002, and $3.082 * 10^{-7}$).

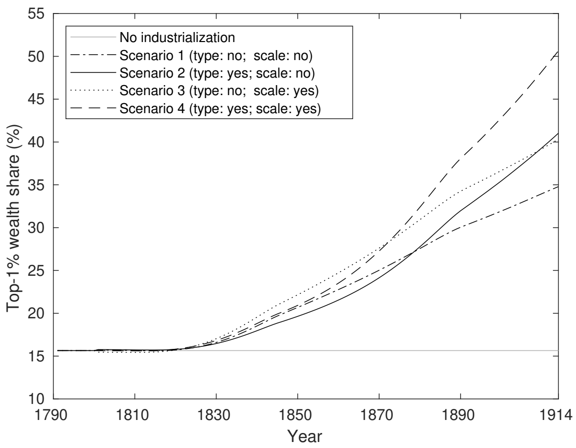
C.3.5 Robustness to eliminating income growth

In the following, we analyze whether our results are robust to eliminating income growth (setting $g = g' = 0$). In our preferred model parameterization, we incorporate income

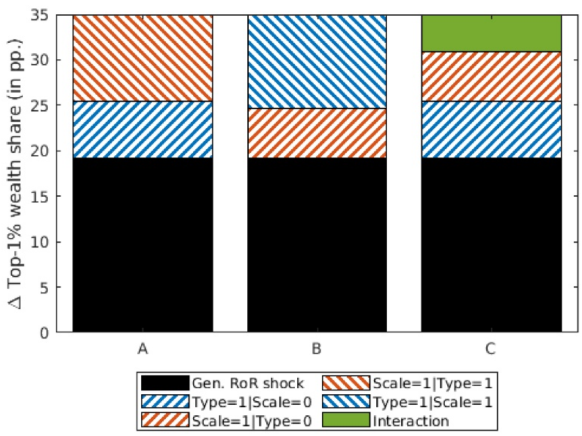
[9] According to Hammond and Bickel (2013), this is the "gold standard" in discretizing continuous distributions.

growth as a characteristic feature of industrialization. It has a dampening effect on the increase in wealth concentration as it induces higher social mobility in lower parts of the wealth distribution and it ensures a more realistic evolution of aggregate measures such as the wealth-income ratio. Indeed, Figure C.4 shows that, without income growth, the increase in the top percentile's wealth share is larger in all four scenarios compared to our parameterization with $g' > 0$ and $g > 0$ with $g = g'$ (Figure 4). However, the decomposition suggests that the relative importance of scale and dynastic type dependence as well as their interaction is unaffected by the introduction of income growth to the model.

Figure C.4: Simulation of wealth concentration until WWI (no income growth)



(a) Simulated trajectories of top 1 share



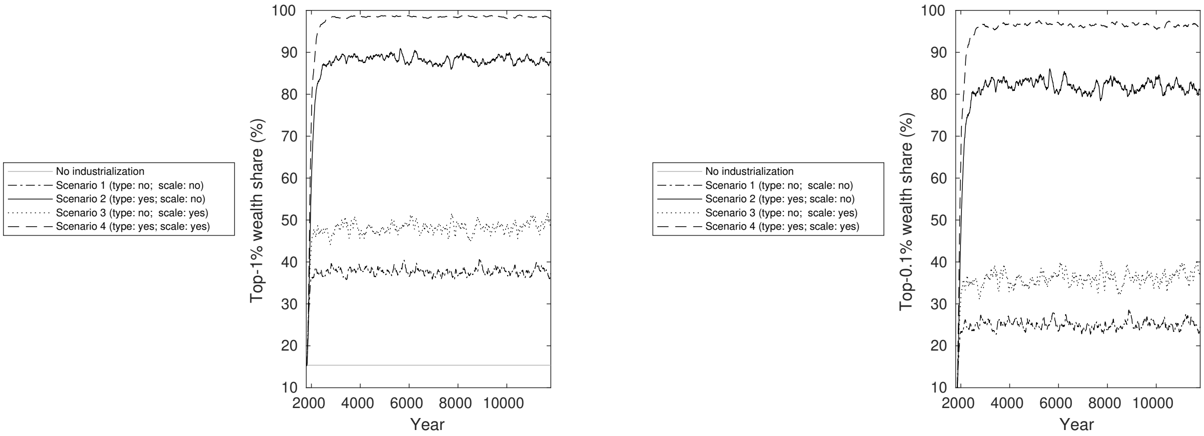
(b) Decompositions of the increase in the top-1% share

Notes: The trajectories in Figure 4(a) refer to the simulated scenarios listed in Table 3 but setting $g = g' = 0$, and represent averages over 200 iterations. The abbreviations "type" and "scale" stand for the features of dynastic type dependence and scale dependence, respectively. Figure 4(b) decomposes the contribution of these two features and their interaction to the absolute change in wealth concentration relative to the benchmark. See main text for the analogous analysis of the scenarios including income growth for details.

C.3.6 Stability of long-run wealth concentration

To determine the behavior of the model trajectories of wealth concentration following industrialization in the very long run, we simulate the model for 10,000 years (Figure C.5). This numerical analysis suggests that all four scenarios result in stable top shares below unity after a period of convergence.[10]

Figure C.5: Long-run stability of top shares



(a) Top-1% wealth share

(b) Top-0.1% wealth share

Notes: The trajectories refer to the simulated scenarios listed in Table 3 and represent averages over 50 iterations. The abbreviations "type" and "scale" stand for the features of dynastic type dependence and scale dependence, respectively.

To provide an idea of the scenario-specific convergence time, Table C.3 additionally reports the number of years each scenario takes until it reaches a steady state in the top-1 and the top-0.1 share. We define the steady state as the average of the last 3000 of the 10,000 years of the respective simulated series in Figure C.5 and we report the years it takes the moving average (centralized; taking into account the past 150 and the future 150 years) of the respective series to reach that steady state.

Table C.3: Convergence time for the simulated scenarios

Scenario	Dynastic type dep.	Scale dep.	Years until steady state... ...top1 share	...top 0.1 share
1	no	no	900	903
2	yes	no	1148	1143
3	no	yes	1303	1307
4	yes	yes	1797	1800

Note that the two scenarios with dynastic type dependence feature a long-run convergence of the share of wealth-holders in industry to unity, which prolongs the period of

[10]Note that the noise in the series is due to the limited number of simulated households (562,500) and iterations (50) – a choice made against the backdrop of increased computing time for the extended time period.

convergence to the steady state, whereas that share reaches an early plateau at around 0.3 in the counterfactuals without dynastic type dependence (see Figure C.1).

C.3.7 Comparison of simulation output and empirical data

In Table C.4, we document the model fit regarding the start and endpoint of our simulation period. The reported simulated values for 1914 refer to Scenario 4, which features scale and dynastic type dependence. The relatively good fit regarding the preindustrial top 1 share is unsurprising given that we target it in our moment-matching approach (see Table 4). By contrast, the remaining reported simulated moments, and notably all those for 1914, are also quite close to their empirical counterparts even though they are not targeted.

Table C.4: Simulation output vs. empirical data

Measure	Value in 1800		Value in 1914	
	simulation	data	simulation	data
Top-1 share	16%	18%	44%	45%
Top-10 share	52%	49%	74%	80%
Top 1 share/Top 10 share	30%	36%	60%	57%
Wealth-income ratio	369%	NA	434%	477%

Notes: Empirical data for 1800 and 1914 are from from Alfani et al. (2022) and Albers et al. (2022), respectively. The reported simulated values for 1914 refer to scenario 4 of Table 3, i.e., including scale and dynastic type dependence.

Appendix References

Albers, T. N. H., C. Bartels, and M. Schularick (2022). The distribution of wealth in Germany, 1895-2018. *CESifo Working Paper No. 9739*.

Alfani, G., V. Gierok, and F. Schaff (2022). Economic inequality in preindustrial Germany, ca. 1300–1850. *The Journal of Economic History 82*(1), 87–125.

Angeletos, G.-M. (2007). Uninsured idiosyncratic investment risk and aggregate saving. *Review of Economic Dynamics 10*(1), 1–30.

Asch, K. (2005). *The 1:5 million international geological map of Europe and adjacent areas*. Hannover: German Federal Institute for Geosciences and Natural Resources (BGR).

Bach, L., L. E. Calvet, and P. Sodini (2020). Rich pickings? Risk, return, and skill in household wealth. *American Economic Review 110*(9), 2703–47.

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

Benhabib, J., A. Bisin, and M. Luo (2019). Wealth distribution and social mobility in the US: A quantitative approach. *American Economic Review 109*(5), 1623–47.

Benhabib, J., A. Bisin, and S. Zhu (2011). The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica 79*(1), 123–157.

Biedermann, E. (1918). Die Einkommens- und Vermögensverhältnisse der preussischen Bevölkerung im Zeitraum 1895–1914 und ihre Übertragunsmöglichkeit auf das deutsche Reich. *Zeitschrift des Königlich Preussischen Statistischen Landesamts 58*, 60–86.

Blanchet, T., J. Fournier, and T. Piketty (2022). Generalized Pareto curves: Theory and applications. *Review of Income and Wealth 68*(1), 263–288.

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

Conley, T. G., C. B. Hansen, and P. E. Rossi (2012). Plausibly exogenous. *Review of Economics and Statistics 94*(1), 260–272.

Crow, E. L. and K. Shimizu (1987). *Lognormal distributions: Theory and applications*. Routledge.

De Nardi, M. (2004, 07). Wealth Inequality and Intergenerational Links. *The Review of Economic Studies 71*(3), 743–768.

European Environment Agency (2012). *WISE Large rivers and large lakes*. European Environment Agency.

Fagereng, A., L. Guiso, D. Malacrino, and L. Pistaferri (2020). Heterogeneity and persistence in returns to wealth. *Econometrica 88*(1), 115–170.

Fick, S. E. and R. J. Hijmans (2017). Worldclim 2: New 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology 37*(12), 4302–4315.

Food and A. Organization (2021). *Global agro-ecological zones (GAEZ)*. Food and Agriculture Organization.

Galloway, P. R. (2007). *Galloway Prussia database 1861 to 1914*. www.patrickgalloway.com.

Hammond, R. K. and J. E. Bickel (2013). Reexamining discrete approximations to continuous distributions. *Decision Analysis 10*(1), 6–25.

Hubmer, J., P. Krusell, and A. A. Smith Jr (2021). Sources of US wealth inequality: Past, present, and future. *NBER Macroeconomics Annual 35*(1), 391–455.

Jordà, Ò., K. Knoll, D. Kuvshinov, M. Schularick, and A. M. Taylor (2019). The rate of return on everything, 1870-2015. *The Quarterly Journal of Economics 134*(3), 1225–1298.

Kunz, A. (2001). *Schiffahrtsstraßen im Deutschen Zollverein 1850*. IEG-Maps.

Königlich Preussisches Statistisches Landesamt (1908). *Statistik der preussischen Einkommenssteuer-Veranlagung für das Steuerjahr 1908 und die Ergänzungssteuer-Veranlagung für die Steuerjahre 1908/1910*. Verlag des Königlichen Statistischen Landesamts.

Königlich Preussisches Statistisches Landesamt (1914). *Statistik der preussischen Einkommenssteuer-Veranlagung für das Steuerjahr 1914 und die Ergänzungssteuer-Veranlagung für die Steuerjahre 1914/1916*. Verlag des Königlichen Statistischen Landesamts.

Königlich Preussisches Statistisches Landesamt (1915). *Statistisches Jahrbuch für den Preussischen Staat 1914*. Verlag des Königlichen Statistischen Landesamts.

Königlich Preussisches Statistisches Landesamt (1916). *Statistisches Jahrbuch für den Preussischen Staat 1915*. Verlag des Königlichen Statistischen Landesamts.

Miller, A. C. and T. R. Rice (1983). Discrete approximations of probability distributions. *Management Science 29*(3), 352–362.

Pfister, U. (2019). The Crafts–Harley view of german industrialization: An independent estimate of the income side of net national product, 1851–1913. *European Review of Economic History 24*(3), 502–521.

Robinson, S. (1976). A note on the u hypothesis relating income inequality and economic development. *The American Economic Review 66*(3), 437–440.

Roser, M. (2023). *Our world in data*. www.ourworldindata.org.

Statistik des Deutschen Reichs (1909a). *Berufs- und Betriebszählung vom 12. Juni 1907. Gewerbliche Betriebsstatistik. Band 218*. Verlag von Puttkammer & Mühlbrecht.

Statistik des Deutschen Reichs (1909b). *Berufs- und Betriebszählung vom 12. Juni 1907. Kleinere Verwaltungsbezirke. Band 209*. Verlag von Puttkammer & Mühlbrecht.

Tilly, R. H. (2010). The distribution of personal income in Prussia, 1852 to 1875: An exploratory study. *Jahrbuch für Wirtschaftsgeschichte 51*(1), 175–194.

Wegge, S. A. (2021). Inheritance institutions and landholding inequality in nineteenth-century germany: Evidence from hesse-cassel villages and towns. *The Journal of Economic History 81*(3), 909–942.

Wehler, H.-U. (2008). *Deutsche Gesellschaftsgeschichte. Erster Band: Vom Feudalismus des Alten Reiches bis zur defensiven Modernisierung der Reformära 1700-1815*. C.H. Beck.