Korn, Tobias; Lacroix, Jean

Working Paper

Local reallocation: Lessons from bankruptcies during Britain's market integration

IOS Working Papers, No. 408

Provided in Cooperation with:
Leibniz Institute for East and Southeast European Studies (IOS), Regensburg

Leibniz Institute for
EAST AND SOUTHEAST
EUROPEAN STUDIES

Arbeitsbereich Ökonomie

IOS Working Papers

No. 408 July 2025

Local Reallocation: Lessons from Bankruptcies during Britain's Market Integration

Tobias Korn*, Jean Lacroix**

* Leibniz Universität Hannover & Universität Heidelberg, Email: korn@mak.uni-hannover.de; **RITM, Université Paris-Saclay & and CESifo Münich, Email: jean.lacroix@universite-paris-saclay.fr;

Leibniz Institute for
EAST AND SOUTHEAST
EUROPEAN STUDIES

Landshuter Straße 4
D-93047 Regensburg
Telefon: (09 41) 943 54-10
Telefax: (09 41) 943 54-27
E-Mail: info@ios-regensburg.de
Internet: www.ios-regensburg.de
ISSN: 2199-9465

Local Reallocation:

Lessons from Bankruptcies during Britain's Market Integration*

Tobias Korn†, Jean Lacroix‡

July 22, 2025

Abstract

This paper documents a new consequence of market integration: *local reallocation*, i.e., the exit of some workers from production even though employment increases in the same area and industry. Thanks to new data on over 150,000 personal bankruptcies combined with detailed microcensus data from 19th-century Britain, we estimate the causal impact of railway access on employment growth and personal bankruptcies. Market integration increased both employment and bankruptcy probability solely in the manufacturing sector. Studying the mechanisms of local reallocation, we show that market integration increased the number and size of manufacturing firms that employed cheap, task-differentiated labour. Our results extend existing research focused primarily on reallocation either across sectors or across locations.

Keywords: Bankruptcies, Market Integration, Reallocation, Structural Transformation

JEL Codes: N63, L16, O33, R40, K35

*This work has been supported as part of the France 2030 program "ANR-11-IDEX-0003". Previous drafts benefited from the insights of Dan Bogart, Leah Boustan, Peter Hull, Réka Juhász, Miren Lafourcade, Claire Lelarge, Stelios Michalopoulos, Martin Rotemberg, Peter Solar, Alex Trew, and Hans-Joachim Voth. We also thank the participants of the 2024 EPCS, the IfW Kiel Research Seminar, the GDE Conference 2024, the ZEW-ifo Young Scholar Political Economy Workshop 2024, the Beyond Basic Questions Workshop 2024, the EHA Meeting 2024, the ETSG 2024 conference, the 2024 Rome Workshop on Transport and History, the KOF research seminar at ETH Zurich, the RITM Seminar at Paris Saclay, the CEPR-Banco de España Symposium in Economic History, the Happy Days Workshop at LEM (Université de Lille), the University of Mannheim Econ History Colloquium, and the Yale Econ History Lunch for useful comments and suggestions. We want to thank Dana Behnsen, Basile Blanc and Victor Westergren for their research assistance.

† korn@mak.uni-hannover.de, Leibniz Universität Hannover, Königsworther Platz 1 30167 Hannover & Universität Heidelberg, Bergheimerstr. 58, 69115 Heidelberg.

‡jean.lacroix@universite-paris-saclay.fr, RITM, Universite Paris-Saclay, 54 boulevard Desgranges, 92330 Sceaux and CESifo Münich.

1 Introduction

In *Capitalism, Socialism, and Democracy*, Joseph Schumpeter observed "The opening up of new markets, foreign or domestic, and the organizational development [...] illustrate the same process of industrial mutation – if I may use that biological term – that incessantly revolutionizes the economic structure from within, incessantly destroying the old one, incessantly creating a new one" (Schumpeter, 1942, p.91). Existing research extensively documents the macroeconomic benefits of market integration, including increased productivity, structural shifts towards manufacturing, and higher economic growth (Donaldson and Hornbeck, 2016; Bogart et al., 2022; Donaldson, 2018; Hornbeck and Rotemberg, 2024). Yet, integration also comes at a cost. Previous research illustrates geographic reallocation in the form of moving employment from high-wage to low-wage locations (Autor et al., 2013, 2016; Redding, 2016). At the same time, market integration requires firms to reorganize (Melitz, 2003; Atack et al., 2008). These organizational changes also create and destroy economic structures (Juhász et al., 2024). According to these results, those benefiting from market integration and technological change belong to either a different sector or to a different location as those experiencing the dismantlement of the old economic structure.

Yet, going back to the original argument of Schumpeter and of the more recent contribution of Melitz (2003), the destruction of the old sector should also occur in those locations and industries that benefit from market integration. A significant limitation of the current literature lies, therefore, in its focus on reallocation either across sectors (e.g., agriculture to manufacturing) or across locations (e.g., the U.S. to China). This ignores the reallocation among economic actors belonging to the same sector and the same location: *local reallocation*. As an illustration, *local reallocation* would imply that China's WTO accession would not only displace workers in the U.S. labour markets, but also less productive manufacturing firms in China that forfeit their market share to bigger competitors.[1]

In line with the original argument of Melitz (2003), this paper provides novel evidence that reallocation also occurs locally. *Local reallocation* implies that neighboring workers in the same industry have diverging experiences from market integration. In addition, we generalize the evidence for the reallocation among workers (Autor et al., 2013) or the reallocation among firms (Juhász et al., 2024) by observing employment changes and market exits altogether in response to the same market integration shock. Our analysis for 19$^{\text{th}}$-century Britain shows that market integration induces a local economic transformation in the manufacturing sector. Manufacturing experienced both higher employment and higher personal bankruptcies, all in

[1] An interesting discussion of the reallocation induced by the WTO accession in China can be found in Feng et al. (2017).

the same location. *Local reallocation* is moreover most pronounced in those manufacturing industries that underwent the biggest transformations upon market integration.

Local reallocation differs from reallocation as observed before. In a given place and a given sector, *local reallocation* implies a positive relationship between the creation of the new economic structure and the destruction of the old structure. Previous studies have identified this reallocation either across different locations but within a sector (Autor et al., 2013, 2016; Flaaen and Pierce, 2019; Autor et al., 2020; Dippel et al., 2021) or across sectors, but at the national or global level (Pavcnik, 2002; Trefler, 2004; Bustos, 2011).

The empirical challenge in investigating *local reallocation* is that one must combine two high-resolution measures that quantify, respectively, the destruction of the old structure and the creation of a new economic structure. Previously used one-dimensional measures are therefore not suited to test the existence of *local reallocation*. To measure *local reallocation*, it is hence necessary to build a measure of the "*destruction of the old sector*" and contrast it with a measure for the "*emergence of the new sector*" in the same location.

This paper leverages a unique historical data source to introduce such a measure: personal bankruptcies in 19^{th}-century Britain. We have collected and curated these new data on personal bankruptcies from the London Gazette using Optical Character Recognition and text recognition algorithms. The London Gazette regularly issued announcements of all newly opened personal bankruptcy cases starting in the late 18^{th}-century across all of Britain. The announcements of personal bankruptcies often identify bankrupted business owners and workers. They hence offer a complete view of the "*destruction of the old sector*." Each bankruptcy case records the home address of the bankrupt, the date of bankruptcy, and the bankrupt's occupation, which we assign to an economic sector. Thanks to this granularity, we have built a high-dimensional dataset at the location-sector-time level, aggregating information on around 150,000 bankruptcy cases between 1851 and 1890. We combine these bankruptcy data with British microcensus information on employment and firms at the location-sector-time level to arrive at the two measures needed to investigate *local reallocation*: the number of active firms and workers to quantify *"the creation of the new economic structure"* and the number of personal bankruptcies to quantify *"the destruction of the old economic structure"*, all within the same place and sector.

Our study combines these variables with information on rail station locations in 1851, 1861, and 1881 to exploit variation in market integration from the expansion of the railway in 19^{th}-century Britain (Bogart, 2014). High-dimensional fixed effects control for the local correlates of the rail and changes within sectors over time. Pseudo Poisson Maximum Likelihood (PPML) estimations regress bankruptcies and employment on a binary rail connection variable, interacted with indicator variables for different sectors.

The results show that railway access prompts a local reallocation only in the manufacturing sector. In locations with railway access, bankruptcies increase by around 40 percent among employees in the manufacturing sector, holding employment constant. At the same time, manufacturing employment in those same places increases by approximately 32 percent. We observe this simultaneous increase in both bankruptcies and employment in no other sectors. As discussed below, only the manufacturing sector showed signs of economic transformation triggering *local reallocation*. The main results are robust to alternative definitions of the treatment variable, such as using a market access measure instead of a binary rail variable, or by investigating different industrial branches within the manufacturing sector. Leveraging exogenous variation in the railway expansion's spatial and temporal dimensions demonstrates the causality of our baseline results. On the space dimension, both an inconsequential places approach based on a Least Cost Path between railway nodes (see Faber, 2014; Banerjee et al., 2020; Bogart et al., 2022) and controlling for a counterfactual railway network address selection into treatment. On the time dimension, placebo estimations show that the effect corresponds precisely to the rail construction period. Moreover, even a marginal variation in the duration of a rail connection explains an increase in bankruptcies.

Section 7 illustrates the mechanisms explaining this *local reallocation* in the manufacturing sector. The sector underwent significant organizational changes, which impacted firms and transformed the nature of labour (subsection 7.1). Locations connected to the rail comprised more medium-sized manufacturing firms that hired workers in more diverse occupations. At the same time, market integration changed market structures (subsection 7.2). Bankruptcies became more likely where the railway linked people to large firms in their sector at an intermediate distance. Similarly, the first areas connected to the rail experienced a lower hike in manufacturing bankruptcies than later connected areas. Subsection 7.3 rules out higher investment incentives of creditors and workers' repeated entry and exit as other potential mechanisms. It further shows that access to local financial institutions reduced the impact of the rail on bankruptcies in the manufacturing sector. This heterogeneity also suggests that individual financial situations affect the baseline results.

Our results contribute to three strands of the literature. First, they demonstrate that market integration spurs *local reallocation*. Neighboring workers in the same sectors potentially experience reallocation differently (Melitz, 2003). On the one hand, these findings conceptually refine the recent research on global reallocation (Autor et al., 2016, 2020; Heblich et al., 2024). Integrated markets not only reallocate jobs from one location to another. Also within a location, market integration benefits some but poses a burden on others. On the other hand, within a location, reallocation does not only occur across different sectors as has been shown by previous works that link differences in exposure to tariffs to (local) sectoral growth (Pavcnik, 2002; Trefler, 2004; Bustos, 2011; Brandt et al., 2017; Flaaen and Pierce, 2019). Investigating *local reallocation* is possible thanks to a new, direct measure of market exits: personal bankruptcies. Only by observing the effect of market integration on the receiving side of reallocation (firms and employment) and the losing side of reallocation (bankruptcies) at the same time, it is possible to investigate *local reallocation*. Even though bankruptcies are repeatedly mentioned in academic and policy works, to our knowledge, they have not yet been directly linked to economic outcomes. While previous scholars mainly studied the legal environment of bankruptcies (Davydenko and Franks, 2008; Ponticelli and Alencar, 2016; Bose et al., 2021), their efficiencies (Ayotte, 2007; Gine and Love, 2010; Li and Ponticelli, 2022) and diffusion (Solar and Lyons, 2011; Bernstein et al., 2019), bankruptcies have not yet been used to assess reallocation. The literature focusing on their determinants mainly investigates access to credit (Del Angel et al., 2024). However, directly capturing the destruction of the old sector matters to understand the effects of economic transformation more clearly and can help design efficient policies. This is crucial as the (fear of) reallocation contributes to political unrest. Geographical reallocation spurs opposition to globalization in declining locations (Autor et al., 2013), and workers in declining industries protest the roll-out of new technologies (Caprettini and Voth, 2020).

Second, this paper emphasizes the consequences of firms' reorganization during the Second Industrial Revolution. The reorganization towards more capital-intensive production potentially reduced the demand for some skills and occupations in the past (Goldin and Katz, 1998) and potentially today (Kogan et al., 2023). Similarly, the gains of the Industrial Revolutions were unevenly distributed across sectors (Temin, 1997) and within sectors (Crafts, 2022). Juhász et al. (2024) present evidence of within-sector reallocation in the case of cotton-spinning in France. In the French cotton-spinning sector, productivity was highly dispersed among firms, and the less productive firms disappeared. As Juhász et al. (2024) define within-sector dynamics, our study adds a geographic dimension to this reallocation. Market integration increases reallocation within a fast-evolving sector despite local aggregate gains.

Third, our results offer a new perspective on the impact of railways and, more broadly, market integration. The literature emphasizes the positive effect of market integration (Donaldson, 2015), contradicting earlier arguments that the economic impact of the railway did not justify its construction costs (Fogel, 1964). Railways increase production (Donaldson, 2018) and productivity (Hornbeck and Rotemberg, 2024). As a consequence, the development of railways spurs economic growth (Donaldson and Hornbeck, 2016). These positive effects come partly from innovation (Tsiachtsiras, 2022) and the creation of new ideas (Chiopris, 2025). Ultimately, railways encourage industrialization (Berger, 2019; Bogart et al., 2022; Kaboski et al., 2024). This structural change hinged on organizational changes as railways prompted the transition from the workshop to the factory (Atack et al., 2008; Tang, 2014; Berger and Ostermeyer, 2024). This paper's findings characterize the biased nature of structural change during the rail expansion in Britain (Bogart et al., 2022). Due to this bias, some market participants underwent financial distress while the majority prospered.

2 The Heterogeneous Effects of Market Integration

2.1 Theory

This section links the literature on market integration, organizational changes, and reallocation to discuss the mechanics of *local reallocation*. According to Donaldson and Hornbeck (2016), the expansion of the rail network in the US has increased "market access." Land with higher market access became more valuable as the integration into larger markets allowed higher returns. The literature has documented this effect in various contexts. Donaldson (2018) estimates that railways in India increased real income by 16 percent. Similar effects were found in different contexts such as Sweden (Berger and Enflo, 2017), Germany (Hornung, 2015), Britain (Bogart et al., 2022) and Kenya (Jedwab et al., 2017). The productivity gains, often explained by higher total factor productivity (Crafts, 2004), were however strongly biased towards the manufacturing sector (Hornbeck and Rotemberg, 2024).

To reap the benefits of integrated markets, the manufacturing sector reorganized (Berger, 2019; Braun and Franke, 2022). As a consequence, manufacturing jobs began to dominate local labour markets (Bogart et al., 2022).[2] In the case of the U.S., Atack et al. (2008) show that the rail led to the transition to the factory system, which in turn led to a higher reliance on unskilled labour (Atack et al., 2024). Tang (2014) also shows that the rail expansion increased investment in firm capitalization in Japan, specifically in manufacturing. In Sweden, railways also have shaped the manufacturing sector and encouraged the transition to the factory system in which the division of labour yields higher returns on economies of scale (Berger and Ostermeyer, 2024).

The theoretical framework of Melitz (2003) clarifies how these structural and organizational changes prompt *local reallocation*. First, market integration and trade increase

[2] These effects extend to overall market integration. For example, Kaboski et al. (2024) also observe that highways in India and China prompted structural change.

competition, putting pressure on the least productive firms that lose market shares. Second, firms need to invest an entry cost to enter trading. The least productive firms cannot afford this investment and exit the market. Conversely, the remaining firms reorganize to increase productivity and pay the entry cost to access other markets. Such reorganization requires economies of scale, labour division, and technology adoption to compete in new markets. Juhász et al. (2024) illustrate how the re-organization of cotton-spinning in the First Industrial Revolution increased productivity. This slow reorganization however led to the disappearance of the least productive firms in the market. According to Chandler (1977), the same occurred during the Second Industrial Revolution. Some industries have experienced rampant innovation. To adopt new technologies and realize economies of scale, firms hired more white-collar workers to solve new organizational issues. This paper does not privilege one of these mechanisms above the others.

According to our working hypothesis, the railway expansion increased productivity in the manufacturing sector as it encouraged its reorganization to fully realize economies of scale and division of labour. Despite higher aggregate productivity, organizational changes in the manufacturing sector jeopardized labour's financial position for three different reasons. First, the emergence of larger firms with monopsony power can decrease wages (Autor et al., 2020). Second, if reorganization shifts labour demand, certain workers lose their appeal to the labour market (Chandler, 1977; Goldin and Katz, 1998; Atack et al., 2019, 2024). Third, workers employed in firms that cannot afford the organizational changes to engage in trade are laid off (Melitz, 2003). In all three cases, more workers in the manufacturing sector experience financial distress despite higher productivity and employment in their sector.

2.2 From Theory to Measure

How to observe this reallocation empirically? To identify the effect of market integration on *local reallocation*, it is necessary to build a measure varying at least across sectors and across space. Leveraging these dimensions gives the opportunity to distinguish *local reallocation* from place characteristics and sector characteristics explaining a decrease in activity.[3] This distinction is important to properly identify the many dimensions along which reallocation crystallizes. It also matters to design efficient policies. For example, if reallocation occurs within integrated places within sectors, then place-based policies would not be able to target all citizens experiencing the destruction of the old economic sector.

[3] Sector characteristics include, for example, sectoral subsidies, changing sectoral market structures, and sector-specific supply chain disruptions. Geographic characteristics include, for example, any place-based policy or access to resources.

Individual bankruptcies record information along these different dimensions. They moreover directly capture a direct consequence of reallocation: the inability to repay debt. In that sense, we depart from previous research on exiting firms (Pavcnik, 2002; Juhász et al., 2024). We do not infer firms' exit from their absence in the data, but directly observe the "destruction of the old system." Hence, our measure of reallocation is not affected by other potential changes in firms' reorganization and market structures, such as, for example, firm relocation or mergers.

This paper's empirical approach uses this new measure together with measures of sectoral dynamics within space to properly identify *local reallocation*: the positive relation between the destruction of the old economic structure and the creation of the new one. *Local reallocation* implies a specific pattern. A sector needs to grow and simultaneously witness more individuals in financial distress relative to other sectors in the same location. Our results section directly tests the two ends of this chain: it articulates results on increasing bankruptcies and increasing employment in manufacturing (Section 5). Later, we document the mechanisms triggering both increased employment and bankruptcies *within-sector-place* (Section 7).

3 Historical Background

3.1 Bankruptcy Procedures in 19th Century Britain

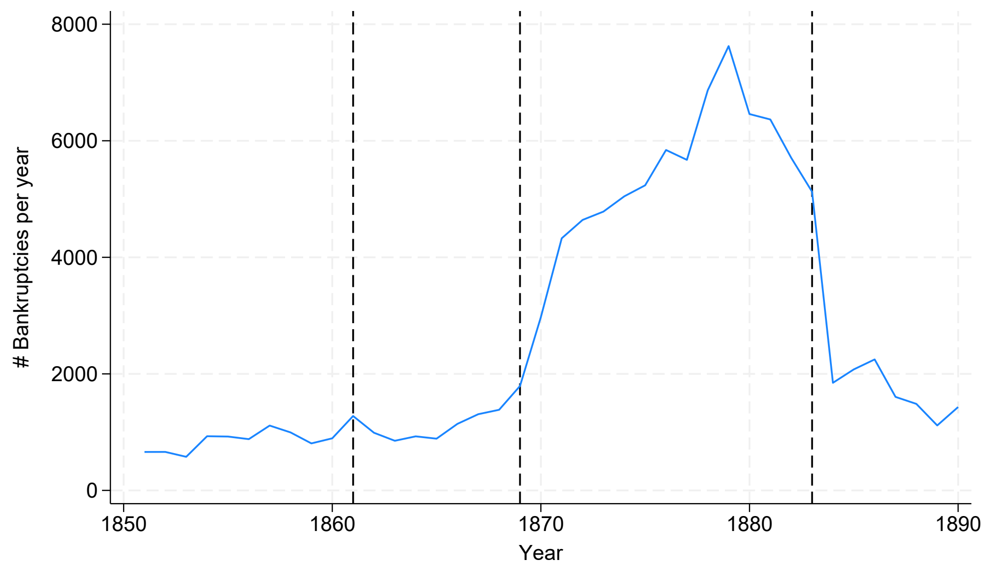
Bankruptcy procedures were at the forefront of political conversations throughout 19$^{\text{th}}$ century Britain (Lester, 1991).[4] At the beginning of the 19$^{\text{th}}$ century, it was common for insolvent debtors to be sent to prison until they could repay their debt. From 1831, the procedure appointed officials to collect and distribute the assets of bankrupts. Both debtors and creditors could then initiate bankruptcy. This doctrine of bankruptcy law called "officialism" was costly and deemed inefficient by entrepreneurs and business elites. The 1869 Bankruptcy and Debtor Acts changed this institution. After this reform, debtors' prison was limited to debtors believed to have the financial means to repay their debt but did not do so. Moreover, the reform repealed the doctrine of "officialism" and put in place a new system of bankruptcy management. If most creditors agreed, they could now proceed to manage bankruptcies themselves. Consequently, recovery rates were higher because creditors had direct incentives

[4] Debtors' prison illustrates well the conversations on bankruptcy law, how complex the system was, and how important bankruptcies were in the collective image of the time. Debtors are, for example, a common figure of Charles Dickens' work reflecting the author's father's own experience as an inmate in a debtors' prison.

to recover as much of the debt as possible. They could also avoid recovering small debts whose costs to recover were greater than the debt itself. In 1883, a new reform reintroduced "officialism."

Another reform during the period of our study, in 1861, broadened the scope of bankruptcy to all citizens, not only those with trading activity. The time series illustrated in Figure 1 evidences the effect of the bankruptcy regime on the number of bankruptcies. Section 7.3 leverages these reforms to inform on potential mechanisms behind the paper's main results.

Figure 1: The Evolution of Bankruptcies



Notes: This figure plots the aggregate number of bankruptcies per year based on new data collected by the authors. Dashed vertical lines indicate three significant reforms to the bankruptcy law: In 1861, the bankruptcy law was extended to all occupations. In 1869, bankruptcy management was put into the hands of most creditors. In 1883, bankruptcy management returned to "officialism" where courts presided over bankruptcy cases.

3.2 The Rail Expansion in Britain

The rail expansion was the last step in Britain's transport revolution (Bogart, 2014). In the second half of the 19th century, the rail became a cheap alternative to transport goods, resources, and people. Between 1840 and 1870, the output of the rail sector was multiplied by 44 (Bogart, 2014). The railway mania of the 1840's structured this expansion but, according to Casson (2009), resulted in an inefficient network.[5] Still, the rail became denser and more

[5]The railway mania was guided by private interests (Esteves and Mesevage, 2021).

productive. Between 1851 and 1881, the length of the railway network in England and Wales nearly doubled (Bogart et al., 2022). In 1851, the network covered mostly the central region of England. By 1881, it expanded to Wales and the South-Western part of England. At the end of the 19th century, the spatial extension and efficiency gains made the rail the main mode of transportation for passengers and materials (Bogart et al., 2022). In effect, it became a cheap alternative to other modes of transportation.

Previous research has debated the rail's overall impact on the British economy. Mitchell (1964), for example, argues that "the introduction of railways in Britain did not have a very great immediate impact on the economy." Hawke (1970) mentions that the social savings generated by railways account for approximately 7.5 percent of Britain's income, and only 4 percent if passengers' comfort is not accounted for in the savings. Overall, Crafts (2004) estimates that 0.05 percent of per capita growth from 1830 to 1860 in Britain can be attributed to the rail. The magnitude of these estimates suggests a rather low impact of the rail on the British economy, at least in the first phase of its development. More recent studies investigated the impact of the rail within Britain during the second development phase. Gregory and Henneberg (2010), for example, argue that areas connected to the railway in Britain experienced an increase in population. Bogart et al. (2022) show that this effect is causal and triggered structural change in Britain.

4 Empirical Strategy

4.1 Data

The main dataset has three dimensions: grid cell, sector, and census year. Hexagonal grid cells with an average area of 214 square kilometers constitute the spatial unit of observation.[6] Within each cell, observations record information for four big sectors (agriculture, manufacturing, trade, and services) across three different census years (1851, 1861, and 1881). We use data on bankruptcies and employment at this level of granularity. Other control variables, such as connection to the rail network and population count at the grid-cell census-year level, complement these data.

Bankruptcy Data. We collect information on personal bankruptcy cases from publications in the London Gazette. Starting in the 18th century, British bankruptcy law required publicizing insolvencies so potential creditors could make their claims official and be considered in debt-clearing. The *London Gazette* contained a separate section that announced

[6] This area is approximately the size of the city of Hannover.

new bankruptcy adjudications and informed debtors of ongoing cases.[7] The first bankruptcy notice was published in the issue of June 5$^{\text{th}}$, 1712. For this study, we accessed all digitized London Gazette issues from January 1851 until December 1890 via the official London Gazette homepage.[8]

To gather the personal bankruptcy announcements, we web-scraped scans of the 5,063 London Gazette issues published from 1850–1890 from the London Gazette homepage. We have found 4,086 regular issues with at least one bankruptcy statement. Figure 2 illustrates two examples. After converting the scans into machine-readable text using Optical Character Recognition (OCR), text recognition algorithms searched for specific keywords to detect personal bankruptcy announcements and extract the name, address, and occupation from each announcement. In a final computation step, we geolocated each address, usually at the city- or parish-level, and assigned people's occupation description to a History of Work Information (HISCO) code. Appendix A describes the data collection process in more detail. Appendix Section B further presents evidence that our dataset is consistent with other data on bankruptcies aggregated at higher levels.

Our main dependent variable is the sector-level annualized number of bankruptcies for each census year. This variable aggregates all bankruptcy cases in a sector and grid cell between two census years, and divides this aggregate by the number of years between the two census years to control for the longer time span between 1861 and 1881. We divide occupations into four main occupation groups.[9] We drop bankruptcies mentioning pensioners, rentiers, or unemployed as occupations from the analysis.

British Microcensus. We use British micro census data to observe sectoral employment together with a number of additional covariates. These data were made available as part of the Integrated Census Microdata (I-CeM) dataset (Schurer and Higgs, 2023). All census entries contain information on occupations with the associated HISCO codes. We assigned coordinates to all census observations based on the sub-district stated in the survey and intersected the subdistrict coordinates with our grid cells. Sub-districts provide a good compromise between geographic resolution and overall representativeness. Even though census sub-districts constitute bigger spatial entities in some rural areas, we could geocode every sub-district from each census wave.[10]

Additional Data. We complement our dataset with additional data sources that vary

[7]The *London Gazette* started out as the main public mouthpiece of the British government in 1665, was delivered on average two to three times per week, and is still being published today.

[8]For more information and to access the London Gazette issues, see https://www.thegazette.co.uk/.

[9]Appendix A.2 describes how we create these categories.

[10]Appendix A.3 describes how we identify locations.

Figure 2: Examples of Bankruptcy Announcements

WHEREAS a Petition for adjudication of Bankruptcy, bearing date the 7th day of August, 1858, hath been filed against John Harris Blakemore, of Wednesbury, in the county of Stafford, Brass and Iron Founder, and he being declared bankrupt, is hereby required to surrender himself to John Balguy, Esq., one of Her Majesty's Commissioners of the Birmingham District Court of Bankruptcy, at Birmingham, on the 27th day of August instant, and on the 16th day of September next, at half past eleven of the clock in the forenoon, on each of the said days, and make a full discovery and disclosure of his estate and effects; when and where the creditors are to come prepared to prove their debts, and at the first sitting to choose assignees, and at the last sitting the said bankrupt is required to finish his examination. All persons indebted to the said bankrupt, or that have any of his effects, are not to pay or deliver the same but to Mr. Frederick Whitmore, No. 19, Temple-street, Birmingham, the Official Assignee whom the Commissioner has appointed, and give notice to Mr. John Smith, Solicitor, Birmingham.

(a) Bankruptcy Announcement 1858

The Bankruptcy Act, 1869.
In the London Bankruptcy Court.

In the Matter of a Bankruptcy Petition against George Dighjohn, of No. 31, Walworth-road, in the county of Surrey, Hair Dresser, trading under the name, style, or description of George Alma Gage.

UPON the hearing of this Petition this day, and upon proof satisfactory to the Court of the debt of the Petitioners, and of the trading, and of the act of Bankruptcy alleged to have been committed by the said George Dighjohn having been given, it is ordered that the said George Dighjohn be, and he is hereby, adjudged bankrupt.—Given under the Seal of the Court this 18th day of October, 1877.

By the Court,
***James R. Brougham*, Registrar**

The First General Meeting of the creditors of the said George Dighjohn is hereby summoned to be held at the London Bankruptcy Court, Lincoln's-inn-fields, in the county of Middlesex, on the 6th day of November, 1877, at eleven o'clock in the forenoon, and that the Court has ordered the bankrupt to attend thereat for examination, and to produce thereat a statement of his affairs, as required by the statute.

(b) Bankruptcy Announcement 1870

Notes: This figure illustrates the layout of the original London Gazette files based on which the bankruptcy data were collected. Figure (a) shows one from the beginning of our sample period in 1858, and Figure (b) displays a later entry from 1877. From these texts, our algorithm would collect the information on the bankruptcy's name, address, and occupation.

at the grid cell level, over time, or both. First, we use data on railway station locations in England and Wales in 1851, 1861, and 1881 from Martí-Henneberg et al. (2017a,b,c) to assign each grid cell its number of stations in a given year. We further leverage data from Fernihough and O'Rourke (2020), which locate the British towns with access to coal to compute the distance of each grid cell's centroid to the closest town with coal access as a proxy for coal availability in a location. The distances to London and UK ports from every grid cell's centroid are also added as controls. To control for credit supply, we also control for the number of country banks in a cell (Heblich and Trew, 2018).

4.2 Method

The main econometric exercise estimates Equation 1 to leverage variation across space, time, and sectors.

$$Bankruptcies_{i,s,t} = exp\left[\beta_{1,s}\mathbb{1}Rail_{i,t} \times Sector_s + \beta_2\mathbb{1}Rail_{i,t} + \Gamma X_{i,s,t} + \nu_{t,s} + \eta_i\right] + \epsilon_{i,s,t}. \quad (1)$$

Using the Pseudo-Poisson-Maximum-Likelihood (PPML) estimator accounts for the overdispersed distribution of the dependent variable, $Bankruptcies_{i,s,t}$. This outcome variable is the annualized number of bankruptcies in grid cell i, sector s, and census year t. $\mathbb{1}Rail_{i,t}$ is a binary variable equal to one once a grid cell is connected to the rail network. $Sector_s$ is a binary variable that, in the main specifications, identifies the manufacturing sector. $\epsilon_{i,s,t}$ is the error term. Location (η_i) and sector-time ($\nu_{t,s}$) fixed effects reduce the identifying variation to variation within geographic areas and variation within sectors in a given census year. As a consequence, no geographic characteristic (such as the proximity to resources), and no time-varying, sector-specific characteristic (such as technological or organizational change) can affect the coefficient of interest, $\beta_{1,s}$.

This coefficient $\beta_{1,s}$ derives from interacting the two binary variables that identify locations with railway access and the manufacturing sector, respectively. $\beta_{1,s}$ hence captures the *differential effect* the railway expansion has on the manufacturing sector to assess how market integration interacts with the organizational changes of manufacturing firms to generate *local reallocation*. Leveraging this variation is new. It complements previous studies that either measure the impact of market integration across space (Autor et al., 2013) or assess within-sector reallocation (Juhász et al., 2024) but did not leverage the spatial dimension of within-sector reallocation. This approach refines the assessment of the theoretical arguments of Schumpeter (1942) and Melitz (2003) as it identifies whether the creation of a new economic structure coincides with the destruction of the old one at a highly disaggregated level (within small geographic units and sector-years).

The matrix $X_{i,s,t}$ adds several control variables. Employment in each sector-location-year accounts for a potential scale effect. Distance to coal, interacted with a manufacturing fixed effect, is an important control variable as it proxies for the propensity of manufacturing to develop in a specific area (Fernihough and O'Rourke, 2020). Similar interactions between a manufacturing fixed effect and the distances to London and the nearest port account for differences in the exposure to investment capital, production networks, international trade, and migration. Finally, we control for credit supply by interacting a manufacturing sector fixed effect with the number of country banks in the 1830's from Heblich and Trew (2018).

Eventually, our estimator uses two types of variation. First, spatial variation from the initial connection to the rail in 1851 at the start of our sample, and second the rail's expansion between 1851 and 1881. The fixed effects are not collinear with the interaction $Rail_{i,t} \times Sector_s$ in 1851. Hence, the results have to be interpreted as the effect of having a rail connection, and not as the effect of a station opening in the second phase of the expansion of the railway network. Extensions in Table 6 disentangle these two dimensions.

5 Results

5.1 Baseline Results

Table 1 introduces our main results. Column 1.1 presents the coefficient from regressing the number of bankruptcies on the binary rail variable. Being connected to the rail network is associated with a higher number of bankruptcies. On average, grid cells connected to the rail network experience around 13 times as many bankruptcies as non-connected cells.[11] Yet, this coefficient could result from selection into the rail or other geographic characteristics of connected cells explaining bankruptcies. Column 1.2 adds an indicator variable for the manufacturing sector and interacts it with the rail variable. The interaction term suggests that upon a connection to the rail, bankruptcies in the manufacturing sector increase by around 65 percent more than in other sectors.

Table 1: Main Results - The Effect of the Rail on Bankruptcies

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	2.62***	2.46***	0.05	0.08	0.05	0.12
	(0.21)	(0.21)	(0.14)	(0.15)	(0.14)	(0.14)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s		0.50***	0.50***	0.48***	0.55***	0.34***
		(0.07)	(0.07)	(0.06)	(0.07)	(0.08)
Observations	8341	8341	7481	7481	7473	7473
Pseu. R^2	.0765	.092	.8	.801	.801	.813
Geo FE			✓	✓	✓	✓
Sector FE			✓	✓	✓	
Year FE			✓	✓	✓	
Sector-Year FE						✓
Sector $\text{Employment}_{i,s,t}$				✓	✓	✓
Geo Controls_i $\times$ Manufacturing_s					✓	✓
Banks_i $\times$ Manufacturing_s					✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the number of people employed the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

The following columns progressively introduce fixed effects for sector, year, and location as well as different control variables. The coefficient for the rail variable turns insignificant

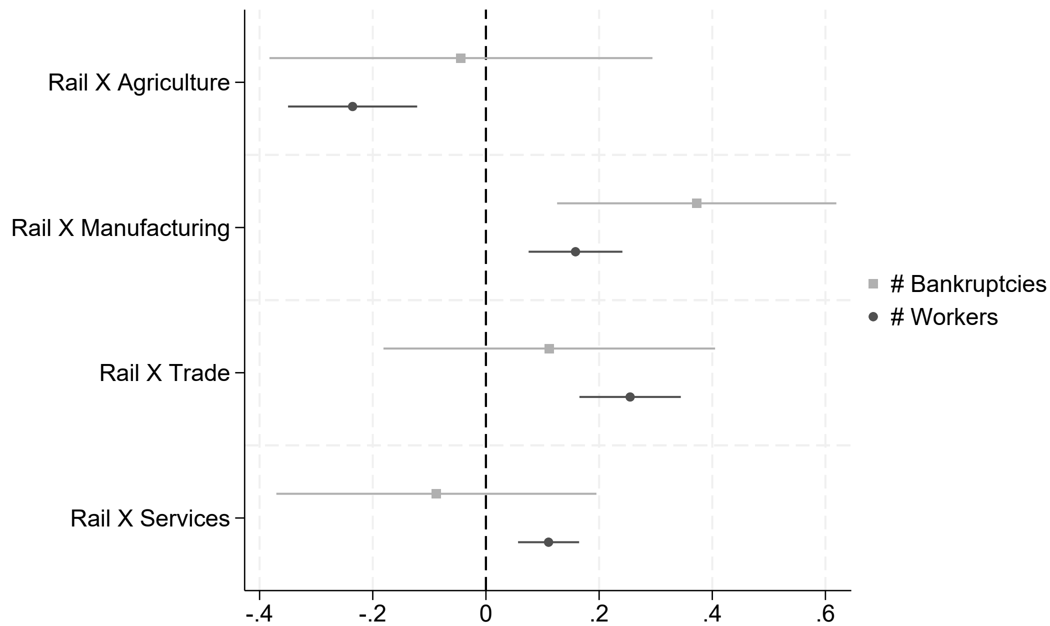
[11] The magnitude of the PPML regression coefficients is computed using the e-transformation $(e^{\beta}-1)\cdot 100\%$. Here, the coefficient 2.62 hence corresponds to an effect size of $(e^{2.62}-1)\cdot 100\% = 1273.5\%$. The map in Appendix Figure B.2 illustrates this result graphically.

after accounting for location-fixed effects. In contrast, from Column 1.3 to 1.5, the interaction term remains significant and positive. According to these estimates, a connection to the rail increases bankruptcies in the manufacturing sector by 64 to 75 percent. Neither controlling for sectoral employment nor adding high-dimensional fixed effects changes the estimate significantly. The effect of the rail on manufacturing bankruptcies, hence, goes well beyond a sector size effect and is not solely driven by between-sector reallocation. Similarly, a correlation between railway connection and special characteristics of the manufacturing sector in areas close to coal, ports, London, or with historical access to credit does not explain our effect. Column 1.6 introduces sector-time fixed effects to control for time-varying sector characteristics such as technological and organizational changes emphasized in Juhász et al. (2024). Accounting for this variation, bankruptcies are still 40 percent higher in the manufacturing sector when connected to the railway.

The railway can prompt bankruptcies in the manufacturing sector for two potential reasons. Either bankruptcies result from a decline in sector activity, or they are the product of labour reallocation because of market integration. To test which of these hypotheses holds, Appendix Table B.1 estimates the baseline regression using sectoral employment as the dependent variable. Manufacturing employment in cells connected to the rail is almost 20% higher than in unconnected cells. A connection to the rail network does not trigger a decline in the manufacturing sector but increases its dynamism. In line with Bogart et al. (2022), the expansion of the rail fosters a reallocation towards the manufacturing sector and no movement out of it. According to these results, the manufacturing sector experiences both an increase in bankruptcies and an increase in the number of employees when connected to the railway network. This pattern is characteristic of *local reallocation*. Appendix Tables C.9 and C.10 show similar results when using the share of employees and the share of bankrupts in a sector as dependent variables. The effects documented in Tables 1 and Appendix Table B.1 are relative to other sectors. To better grasp the different layers of reallocation, Figure 3 shows the coefficients from interaction terms with the binary rail variable and an indicator variable for each sector.

The estimates in dark grey indicate that a rail connection decreases the number of workers in the agricultural sector while increasing the number of workers in all other sectors. Workers' transition from the primary to the secondary sector explains the increase in manufacturing employment, whereas, at the same time, employment in services also increases to sustain larger production units (Katz and Margo, 2014). The light grey estimates show the effect of a rail connection on bankruptcies for the different sectors. As in Table 1, the Rail $\times$ Manufacturing coefficient is significant and positive. Estimates for other sectors are not different from zero. Only in the manufacturing sector does a connection to the rail increase

Figure 3: Coefficient plot – The Railway's Effect on Bankruptcies and Employment



Notes: The figure reports the coefficients for the railway expansion by sector. The coefficients result from estimating specifications following Equation 1. Dependent variables are the number of workers (in black) and the number of bankruptcies (in grey). All regressions control for the control variables and fixed effects outlined in equation 1. Confidence intervals are at the 95% level. Standard errors are clustered at the grid cell level. The results of the estimations are available in Appendix Tables B.2 and B.3.

both the number of bankruptcies and employment. The manufacturing sector is hence the only sector that exhibits this dynamic specific to *local reallocation*.

5.2 Explaining Baseline Results: Within-Manufacturing Estimates

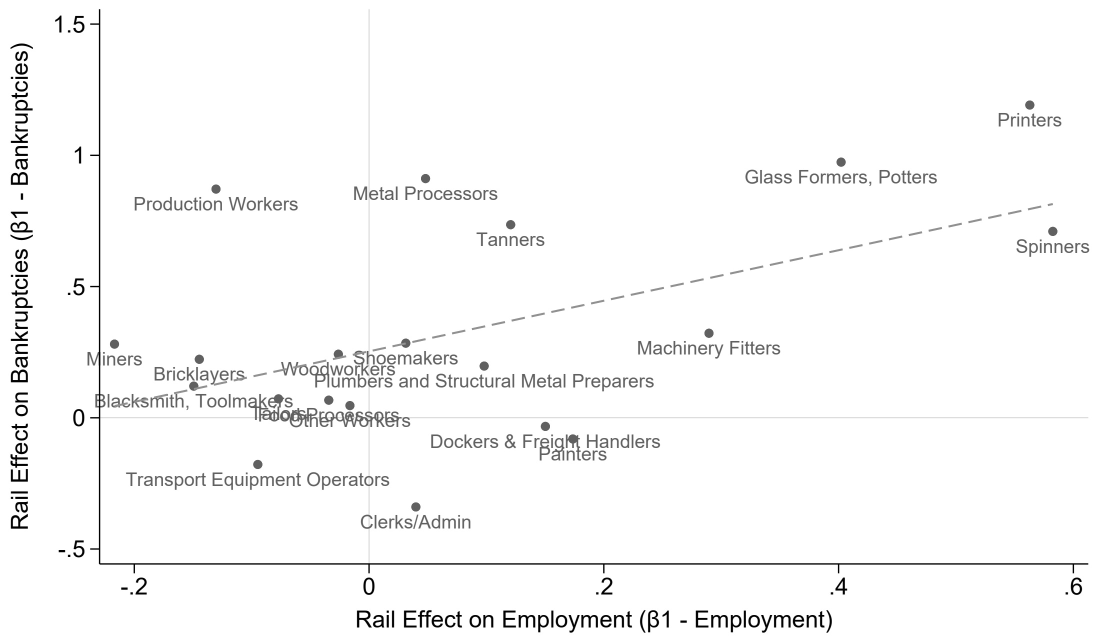
The main estimations use four broadly defined sectors. Potentially, the coefficient on the interaction $\mathbb{1}(Rail > 0) \times Manufacturing$ could hide a reallocation across subsectors. Intuitively, employment could increase in some industries that gain from market integration while bankruptcies occur in declining industries. This pattern would suggest structural change between occupations, i.e. a negative correlation between employment and bankruptcies at the occupation level. To test whether this is the case, we estimate Equation 1 using the variation across manufacturing occupations.[12] Figure 4 summarizes these coefficients in a scatter plot, where each point represents one occupation in manufacturing. The horizontal axis reports the coefficient estimates of our baseline regression (Equation 1) for occupation-level employment as the dependent variable. The y-axis reports the corresponding coefficients for bankruptcies

[12] Appendix Figures B.3 and B.4 report the occupation-specific coefficients from these regressions.

as the dependent variable (while controlling for the level of employment).

Figure 4 presents the same pattern as the aggregate level: bankruptcies occur more often in occupations with higher employment, a pattern consistent with *local reallocation* at the occupation-level.

Figure 4: Within-Manufacturing estimates



This figure shows a scatterplot of coefficient estimates at the occupation level in the manufacturing sector following Equation 1. Each point represents the coefficient for one manufacturing occupation. We plot the estimated effects of the interaction term $\mathbb{1}(Rail > 0) \times Manufacturing$ on employment along the x-axis, and estimated effects on bankruptcies along the y-axis. Figures B.3 and B.4 in the Appendix report all coefficients with confidence intervals. The correlation between coefficients is $\rho = 0.31$.

The occupations where this pattern is most prominent are: spinning, printing, glass formers and potters. Historians have emphasized the importance of organizational changes in these sectors during the second half of the 19$^{\text{th}}$ century. A striking example is the paper/printing industry. It experienced both higher employment and more bankruptcies when connected to the rail. During our sample period, this sector was "dominated by a capital-intensive flow production and a process of innovation characterised by gradual technological developments" (Magee, 1997, p.4). "Economies of scale were vital to survive" in this industry (Magee, 1997, p.5). Similarly, the history of the textile industry emphasizes technologies that revolutionized spinning already before our study period (Maw et al., 2022) and caused bankruptcies (Solar and Lyons, 2011). During our period, the textile industry underwent a series of organizational changes under increasing international competition (Toms, 1997).

Similarly, Pilbin (1937) emphasizes the importance of access to coal and easy transportation to explain the organization and location of the glass industry in Britain. These findings reflect our interpretation of the main results: the manufacturing sector reorganized following market integration. This reorganization was possible because the sector itself was adopting new technologies and modes of production (as in printing and spinning) or because, within this sector, a connection to the rail was of prime importance to access materials and to deliver goods (as was the case for the glass/pottery industry). Despite its growing importance, the manufacturing sector experienced more bankruptcies following this reorganization.

5.3 Robustness

We conduct a number of additional regressions to test the robustness of the main results. These robustness tests, presented in Appendix C, show that this paper's main findings do not hinge on empirical decisions that involve the sample composition, the definition of treatment and outcome variables, or the modelling of spatial dependencies.

The main specifications exploit variation across three dimensions: space, time, and sector. To account for potentially correlated shocks across these different dimensions, we introduce two-way and three-way clustered standard errors in Table C.1. These stricter standard errors do not impede the statistical significance of the main results. A related problem is the choice of the spatial unit of observation. The modifiable areal unit problem (MAUP) suggests empirical results can be specific to a certain type of spatial unit. Table C.1 demonstrates that the main results are robust to using different sizes of grid cells as the spatial unit of observation.

Another concern relates to the specification choice. The main regressions above used a relatively narrow set of local control variables on top of the fixed effects. However, introducing additional control variables such as the number of unemployed in a grid cell, the percentage of the population born in another county, or the percentage of the male population does not impact the estimated coefficients (Appendix Table C.2). Similarly, the results are not driven by local outliers in the sample. The coefficient estimates remain of the same magnitude when excluding the 5% most populated cells, the 5% least populated cells, railway nodes, and all of them together (Appendix Table C.3).[13]

One specific concern for the main outcome variable here are policy changes to bankruptcy law. During the sample period, two such policy changes occurred. An 1861 reform extended

[13] Following Bogart et al. (2022), we define nodes as the 99 British towns that had an urban population of at least 5,000 inhabitants in 1801.

bankruptcy law from individuals that declared a trading activity to all citizens. Another reform in 1869 abandoned the "officialism" doctrine and allowed creditors to manage bankruptcy cases outside courts. Especially the latter reform led to a stark increase in bankruptcy cases until the reform was revoked in 1883. Table C.4 provides results from subsamples dropping these reform periods and shows that these episodes do not impact the main results.

Due to the spatial nature of the treatment variable, spatial autocorrelation or spillovers could lead to biased statistical inference. In Appendix Table C.5, we therefore extend the modeling of spatial dependencies by clustering standard errors at the county level and introducing spatial lags of the treatment and dependent variables. Additionally, Tables C.6 and C.7 introduce Conley standard errors with a 300km spatial cut-off to our main estimations. These spatial regression models also return identical results. Similarly, our results are robust to weighting observations by the inverse of a cell's population to rebalance our estimations towards less populated areas (Appendix Table C.8).

The main results are based on count variables and estimated with the PPML estimator. To investigate the robustness towards different definitions of the dependent variables, Appendix Tables C.9 and C.10 run OLS regressions using the share of bankruptcies over employment and the share of employment over population as dependent variables. In another extension of the dependent variable, we narrow the window to count bankruptcies to the two years following a census year (Appendix Table C.11). This last test ensures that the estimates are not driven by the repeated bankruptcies of individuals who would enter and exit the manufacturing sector between two census waves. All these alternative outcome definitions return results in line with the main specifications.

Finally, we follow Donaldson and Hornbeck (2016) and use a market access measure instead of a binary rail variable as the treatment variable. This market access measure generates similar results as the binary rail indicator used in our main specifications (Appendix Table D.1).

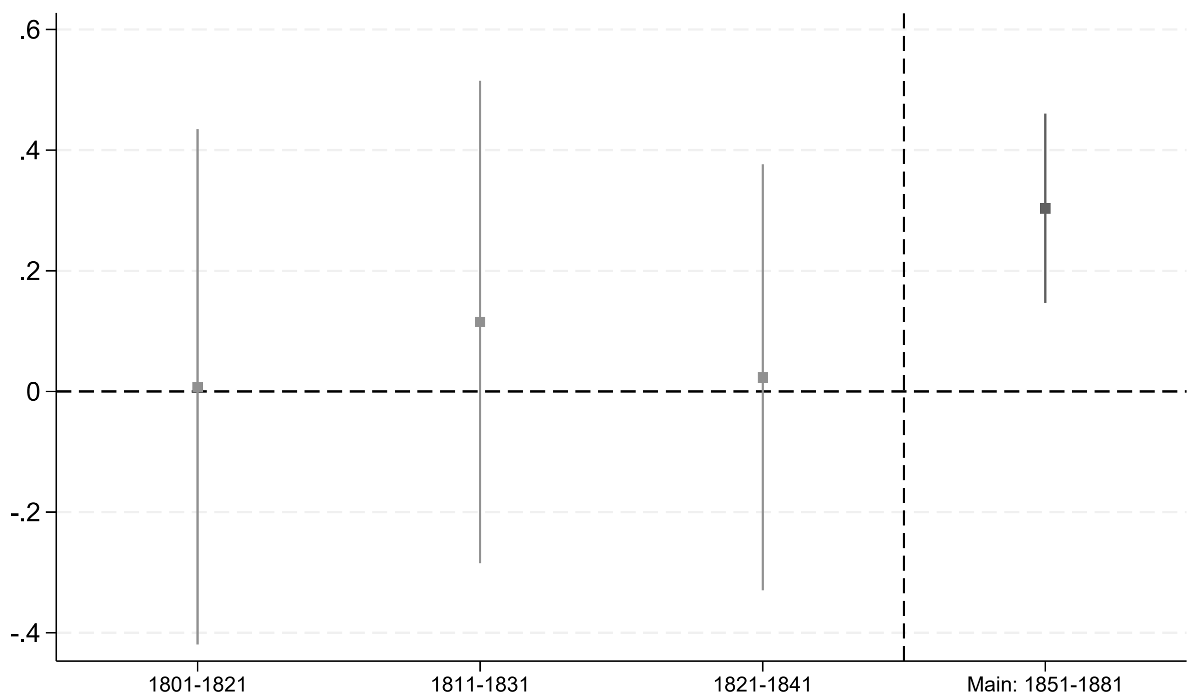
6 Identification

The fixed effects and control variables in the main specifications account for time-invariant geographic features and time-varying sector characteristics. Still, the estimated coefficient of interest can reflect other dynamics that vary over time and across space, and affect the manufacturing sector specifically. To address this concern, we show that the timing and the spatiality of the effect strongly suggest that our estimates can be considered causal.[14]

6.1 Time Dimension – Pre-treatment Placebos

The baseline estimations follow the logic of triple Difference-in-Differences estimations, exploiting variation across place, time, and sector. The main identifying assumption assumes that employment and bankruptcy trends in the manufacturing sector would have been the same across locations with and without rail access if the railway had not been built. This assumption is close to a parallel trends assumption in a difference-in-differences framework.

Figure 5: Testing Parallel Pre-Trends – Coefficients Rail$_i$ $\times$ $Manufacturing_s$ on Pre-Sample



Notes: The Figure shows the results of placebo estimations from the period before the railway network was constructed. We have created a placebo "Rail" variable that equals the rail expansion from 1851–1881, but assign it to the three earlier periods 1801-1811-1821, 1811-1821-1831, and 1821-1831-1841. We then regress bankruptcies that occurred in these earlier periods on the placebo rail expansion variables using Equation 1.

[14]Borusyak and Hull (2023) caution against the causal interpretation of the effects of market access variables because locations can (endogenously) differ in how they are affected by an extension of the transport infrastructure. This logic does not apply to our case. The identifying variation in our main specifications derives from the conditionally exogenous exposure of manufacturing firms to the railway expansion while controlling for the endogeneity of the rail network. This section demonstrates that parallel pre-trends as well as various alternative treatment definitions support the assumption of conditional exogeneity of *the manufacturing sector's* exposure to the railway expansion.

To test the plausibility of this assumption, we investigate bankruptcy trends before the railway was actually built. Leveraging our bankruptcy data that go back until 1788, we estimate placebo regressions that follow our main specifications. These placebo tests use the same specification, keeping the railway expansion from 1851-1881 interacted with the manufacturing sector variable as the explanatory variable. In contrast to the main specifications, the annualized number of bankruptcies from different pre-railway periods enters as dependent variable. Figure 5 presents the coefficient estimates by period for the $Rail_{i,t} \times Manufacturing_s$ interaction in a coefficient plot. None of the placebo coefficients is significant. For all periods, the standard errors are large and the point estimates near zero. The coefficients also do not exhibit any specific upward or downward trend. Accordingly, the main specifications do not pick up any long-term, manufacturing-specific geographic patterns.

6.2 Space Dimension – Exogenous Rail Access

A second test leverages an exogenous variation in the connection to the rail. The main specification directly controls for local correlates of railway construction. However, the rail might have developed faster in areas where the manufacturing sector was prone to bankruptcy. To ensure that this potential selection into the rail does not explain our results, we use a Least Cost Path (LCP) approach similar to Bogart et al. (2022) to model an exogenous variation in access to the rail network.[15] The approach assumes that locations along the LCP between two nodes more likely receive a railway station. We construct a 30 kilometers buffer around the LCP. Appendix Figure E.1 shows that this threshold marks a discontinuity in the probability that grid cells receive access to the rail. Proximity to the LCP only predicts the existence of at least one station, but not the number of stations in a cell (Appendix Table E.1). This suggests that the LCP does not capture factors that correlate with a denser railway network, such as economic activity or natural resources.

Table 2 presents the results from the LCP approach, starting with reduced form estimates in the upper panel. Columns 2.1 to 2.3 use different buffers to define the instrument. Being close to the Least Cost Path increases bankruptcies in the manufacturing sector by around 20 percent. Columns 2.4 to 2.6 add spatial spillovers (i.e. bankruptcies or employment in the same sector in neighboring cells) to control for potential spatial correlations that would affect the instrument. According to these estimates, proximity to the Least Cost

[15] Following Bogart et al. (2022), we select the 99 towns with a population over 5,000 inhabitants in 1801 as natural railway nodes, i.e., as towns that almost certainly would have been among the first to receive a railway station. We then construct LCPs between each of these nodes. These LCPs measure the easiest way to build railway lines between two locations, considering the bilateral distance and the variation in construction costs due to elevation and rivers.

Path increases bankruptcies in the manufacturing sector by 26 percent. In the meantime, the bottom panel shows the validity of our approach. Proximity to the LCP increases the probability of connection to the rail by 12 percentage points.

Table 2: Exogenous Variation in Rail – LCP Proximity

Instrument	LCP<30km	LCP<25km	LCP<35km	LCP<30km	LCP<30km	LCP<30km
	(1)	(2)	(3)	(4)	(5)	(6)
	Reduced Form / Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$					
$\mathbb{1}$ $\text{Instrument}_i \times \text{Manufacturing}_s$	0.18^{**}	0.20^{**}	0.18^{**}	0.15^{*}	0.26^{**}	0.23^{**}
	(0.08)	(0.08)	(0.09)	(0.09)	(0.10)	(0.12)
Spillovers Bankruptcies $_{s,i,t}$				0.11^{*}		0.06
				(0.07)		(0.06)
Spillovers Employment $_{s,i,t}$					0.26^{**}	0.24^{*}
					(0.12)	(0.12)
Observations	7473	7473	7473	7473	7473	7473
Pseu. R^2	.813	.813	.813	.813	.814	.814
	First stage / Dependent Variable: $\mathbb{1}$ Rail $_{i,t}$					
$\mathbb{1}$ Instrument_i	0.12^{***}	0.11^{***}	0.07^{**}	0.12^{***}	0.11^{***}	0.11^{***}
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Spillovers Bankruptcies $_{s,i,t}$				0.01		0.00
				(0.01)		(0.01)
Spillovers Employment $_{s,i,t}$					0.02^{**}	0.01^{**}
					(0.01)	(0.01)
Observations	7473	7473	7473	7473	7473	7473
Adj. R^2	.214	.214	.207	.214	.216	.216
Geo FE	✓	✓	✓	✓	✓	✓
Year × Sector FE	✓	✓	✓	✓	✓	✓
$\text{Sector Employment}_{i,s,t}$	✓	✓	✓	✓	✓	✓
$\text{Geo Controls}_i \times \text{Manufacturing}_s$	✓	✓	✓	✓	✓	✓
$\text{Banks}_i \times \text{Manufacturing}_s$	✓	✓	✓	✓	✓	✓

Notes: Table reports results from PPML reduced form regressions (upper panel) and OLS first stage regressions (lower panel) at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years (upper panel) and an indicator for railway access (lower panel). The main explanatory variable is an indicator for a grid cell i being located within a buffer around the Least Cost Path instrument, interacted with an indicator variable for the manufacturing sector. Spillovers indicate the number of bankruptcies and employment, respectively, in the same sector and year of neighboring grid cells. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

Potentially, the Least Cost Path may correlate with previous transport networks such as turnpike roads or waterways. To ensure that the geography of the rail does not pick up a potential effect of these other networks on bankruptcies, Table 3 directly controls for the length of turnpike roads and of waterways interacted with the manufacturing sector both in our main specification and in the specification using proximity to the LCP. The coefficients attached to the binary rail variable interacted with the manufacturing sector indicator are always significant at the one percent-level. Even after controlling for the geography of other transport networks, a connection to the rail increases bankruptcies in the manufacturing sector by a magnitude similar to the baseline results. A connection to the rail increases bankruptcies by 30 percent in the manufacturing sector (Column 3.3). In locations along the least cost path, manufacturing bankruptcies increase by 18 percent (Column 3.6).

Table 3: Previous Transport Networks – Main Estimates and LCP-Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:			#Bankruptcies$_{i,t,s}$			
$\mathbb{1}$ (Rail$_{i,t}$ > 0) × Manufacturing$_s$	0.32***	0.27***	0.26***			
	(0.09)	(0.07)	(0.08)			
$\mathbb{1}$ Instrument$_i$ × Manufacturing$_s$				0.17**	0.17**	0.16**
				(0.08)	(0.07)	(0.07)
Turnpike$_i$ × Manufacturing$_s$	0.08		0.05	0.05		0.01
	(0.06)		(0.06)	(0.06)		(0.05)
Waterways$_i$ × Manufacturing$_s$		0.05**	0.04**		0.04***	0.04**
		(0.02)	(0.02)		(0.02)	(0.02)
Observations	7473	7473	7473	7481	7481	7481
Pseu. R^2	.813	.813	.813	.813	.813	.813
Geo FE	✓	✓	✓	✓	✓	✓
Year × Sector FE	✓	✓	✓	✓	✓	✓
Sector Employment$_{i,s,t}$	✓	✓	✓	✓	✓	✓
Geo Controls$_i$ × Manufacturing$_s$	✓	✓	✓	✓	✓	✓
Banks$_i$ × Manufacturing$_s$	✓	✓	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Similarly, we add indicators for grid cells containing a turnpike as well as a navigable river or canal. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

As an additional test to ensure that the baseline results can be interpreted as causal, Appendix Table E.2 controls for a counterfactual rail network developed by Casson (2009). This counterfactual rail network is based on the (cost)-efficiency of the rail. It captures the part of the railway network built for economic (and hence endogenous) reasons. Even after controlling for this counterfactual network, the estimates remain significant at the usual level.

6.3 Control Group – Leveraging the timing of connection

While most of the previous estimations leverage the spatial variation in exposure to the rail, this section develops a design-based approach. In this approach, each comparison group consists of grid cells that received a railway station shortly *after* a census year. In the spirit of Borusyak and Hull (2023), areas connected to the rail shortly before or after a census year should be comparable in their potential treatment intensities. This section builds three new control groups based on this intuition: (1) cells getting a connection before the next census, (2) cells getting a connection within five years after the census, and (3) cells getting a connection within two years after the census. On top of the creation of these new control groups, a last approach compares grid cells that got connected to the rail in the two years before the census to grid cells connected to the rail in the two years after the census. All these approaches compare treated cells, but the duration of exposure to the treatment varies. For example, the last approach compares the number of bankruptcies between two censuses for places whose time of exposure to the rail varies by 1 to 3 years (i.e. the control group is treated but later than the treated group). Table 4 presents the results.[16]

Table 4: Timing of connection and treatment – Defining other control groups

	(1)	(2)	(3)	(4)
	Dependent Variable: #Bankruptcies$_{i,t,s}$			
Treatment	Rail	Rail	Rail	Rail 2 years before
Control group	Before next census	5 years after	2 years after	2 years after
$\mathbb{1}$ (Rail$_{i,t}$ > 0)	0.10	0.09	0.50***	
	(0.22)	(0.19)	(0.18)	
$\mathbb{1}$ (Rail$_{i,t}$ > 0) × Manufacturing$_s$	0.42***	0.39***	0.25**	0.22*
	(0.09)	(0.10)	(0.12)	(0.12)
Observations	4302	6149	5893	600
Pseu. R^2	.828	.811	.813	.841
Geo FE	✓	✓	✓	✓
Year × Sector FE	✓	✓	✓	✓
Sector Employment$_{i,s,t}$	✓	✓	✓	✓
Geo Controls$_i$ × Manufacturing$_s$	✓	✓	✓	✓
Banks$_i$ × Manufacturing$_s$	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. The four columns vary the control group that enters the estimations. Column (1) compares connected locations only to locations that receive their first railway station after the current, but before the next census year. This Column hence only considers grid cells treated in 1851 and 1861. Columns (2) and (3) further subset the control group to locations that receive a connection within 5 and 2 years, respectively, after a census year. Column (4) finally also trims the treatment group to locations that received their first station within two years before a census year. Geo controls are the number of people employed the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. *Banks$_i$* refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

Compared to cells connected right after the census, a connection to the rail still increases

[16]Appendix Table E.3 presents a similar intuition along the lines of Costas-Fernández et al. (2020). It compares treated to not-yet-treated units and controls whether a grid cell was ever connected to the rail.

bankruptcies by 20 to 40 percent (Columns 4.1 to 4.3). Already a longer duration of exposure to the rail leads to an increase in bankruptcies. Furthermore, places that got connected to the rail within 2 years before a census still have 20 percent more bankruptcies in the manufacturing sector than cells connected within 2 years after the census. In this case, only 1 to 3 additional years of connection to the rail explain an increase in manufacturing bankruptcies. Appendix Table E.4 shows that this longer exposure does neither explain the stock of bankruptcies before a census year nor the number of bankruptcies between the preceding census years. Meanwhile, within the sample, a balance test shows that the timing of connection does not correlate with demographics, labour market outcomes, and past bankruptcies.[17]

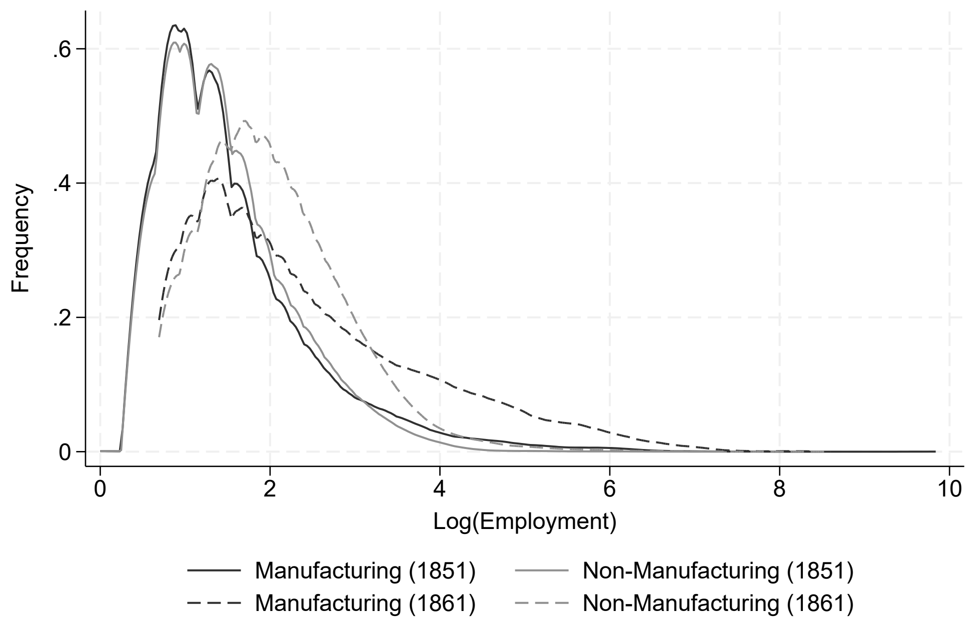
Taking stock, our baseline results appear at the actual date of rail construction and not earlier. Moreover, when we replace our main spatial variation with a variable based on an inconsequential places approach, locations that got a connection to the rail because they were along the way between two important places experienced a surge in bankruptcies in the manufacturing sector. Similarly, a marginal quasi-random variation in the duration of rail exposure also explains the number of bankruptcies in the manufacturing sector.

7 Mechanisms: Organizational Changes

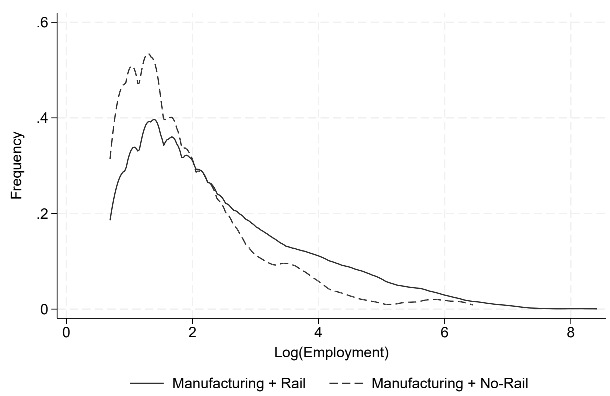
What explains this *local reallocation* in manufacturing? Tang (2014) and Atack et al. (2008) have documented that a connection to the rail prompts organizational changes in manufacturing. This section presents several pieces of evidence suggesting that these organizational changes in turn fueled *local reallocation*. First, subsection 7.1 investigates whether a connection to the rail shapes firms and their demand for labour. Second, subsection 7.2 uses variation in exposure to competition and leverages the timing of the connection to inform on the market dynamics that generate bankruptcies. Finally, subsection 7.3 shows that other mechanisms such as creditors' incentives, workers' occupational choices, and credit supply do not explain our results but are relevant to understanding bankruptcies in 19$^{\text{th}}$ century Britain.

[17] See Appendix Section E.3 for a more detailed discussion of this design-based approach.

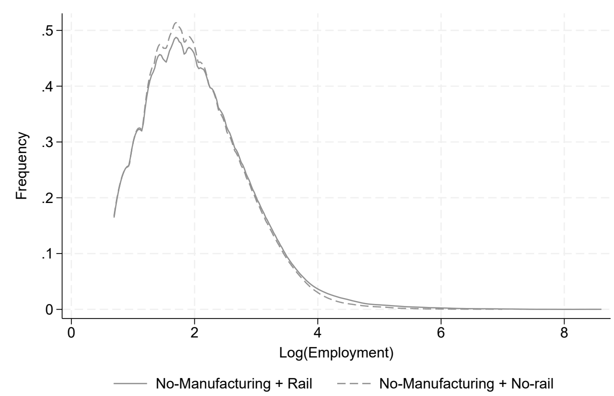
Figure 6: Firm size, Sectors and Rail Presence (1851 and 1861)



(a) Firms' Size by Sector in 1851 and 1861



(b) Firms' Size by Rail connections – Manufacturing



(c) Firms' Size by Rail connections – Non-Manufacturing

Notes: These figures display Kernel density functions for the number of employees based on data on business owners from the British 1861 census. The left panel shows the density in the manufacturing sector, the right panel shows the density for the remaining sectors. Bold lines display the density functions across locations with railway access, while dashed lines show density functions for locations without railway access in 1861.

7.1 Organizational Changes – Firms and Labour

Manufacturing was the most heterogeneous sector entering the second half of the 19th century.[18] To see how the distribution of firms evolved over time, Figure 6a plots the firm size distribution for 1851 and 1861, the two censuses that included open items where firm owners could state their occupation. In 1851, the plain lines show that the density function of the manufacturing sector is similar to the density function of other sectors. As expected, the distribution's right tail is slightly thicker for the manufacturing sector, reflecting the existence of larger firms in this sector.

The distributions diverge in 1861. Both distributions shift to the right, but they do not look alike anymore. The right tail of the manufacturing sector is now different from the right tail of the other sectors. The kernel density of manufacturing firms becomes flatter and the number of smaller firms decreases. Figures 6b and 6c further investigate the reason of this shift. They distinguish the distributions of firms by sector in 1861 and by their access to the rail network in 1851. The distribution of non-manufacturing firms (right panel) is exactly the same whether the firms are connected to the rail or not. In contrast, the distribution of firms in the manufacturing sector depends on whether they are connected to the rail in 1851 (left panel). Compared to the non-manufacturing sectors, the manufacturing firms connected to the rail (black plain line) have a density function that differs significantly from the others. The right tail is thicker, meaning that the rail promotes the growth of large firms.

Table 5 investigates this reallocation at the firm level (Columns 5.1 to 5.4) and at the labour market level (Columns 5.5 to 5.7). The estimates show that a connection to the rail increases the number of manufacturing firms more than threefold (Column 5.2). The prevalence of medium-sized firms with more than 10 employees increases by around 54% (Column 5.3). However, the rail has no impact on the average size of manufacturing firms or on the number of firms with more than 100 employees (Column 5.4).

Columns 5.5 to 5.7 test whether a connection to the rail also shapes manufacturing firms' labour demand. In particular, a connection to the rail decreases self-employment (Column 5.5) by around 7%, increases occupational diversity in the manufacturing sector (Column 5.6), and increases child labour (Column 5.7) by almost 50%.[19] These results, taken together, clarify the organizational changes brought by the rail in the manufacturing sector. Firm structures and the labour market mutated upon the arrival of the rail. Manufacturing firms became larger and more complex than firms in other sectors.

[18] Table B.4 provides measures of the heterogeneity in firm sizes captured by their level of employment as recorded in the I-CeM project (Schurer and Higgs, 2023).

[19] Child Labour is here considered as a proxy for unskilled labour (Humphries, 2013).

Table 5: Rail and Organizational changes - Firms size and labour

Dep Variable:	Firms' Avg Size (1)	Nb Firms (2)	Nb Firms $\geq$ 10 empl (3)	Nb of Firms $\geq$ 100 empl (4)	Self Employed (5)	HH Occup (6)	Child Labour (7)
$\mathbb{1}$ (Rail$_{i,t}$ > 0)	-0.04	-0.55	-0.24***	-0.29	0.02*	0.03	-0.04
	(0.19)	(0.51)	(0.09)	(0.48)	(0.01)	(0.02)	(0.03)
$\mathbb{1}$ (Rail$_{i,t}$ > 0) $\times$ Manufacturing$_s$	0.18	1.48***	0.43***	0.51	-0.07**	-0.22***	0.40***
	(0.18)	(0.38)	(0.14)	(0.43)	(0.03)	(0.03)	(0.05)
Log(Nb of Firms$_{i,s,t}$)			1.29***	0.92***			
			(0.04)	(0.06)			
Observations	5624	1876	5088	1876	8341	6008	8337
Pseu. R^2	.532	.664	.818	.722	.119	.668	.074
Geo FE	✓	✓	✓	✓	✓	✓	✓
Sector $\times$ Year FE	✓	✓	✓	✓	✓	✓	✓
Population$_{i,t}$	✓	✓	✓	✓	✓	✓	✓
Geo Controls$_i$ $\times$ Manufacturing$_s$	✓	✓	✓	✓	✓	✓	✓
Banks$_i$ $\times$ Manufacturing$_s$	✓	✓	✓	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. Columns (1) to (4) investigate firm characteristics. Here, the dependent variables are the average firm size (Column 1), the number of firms (Column 2), the number of firms with 10 or more employees (Column 3), and the number of firms with 100 or more employees (Column 4). Columns (5) to (7) investigate labour characteristics. Here, the dependent variables are the number of self-employed individuals (Column 5), the Hausman-Herfindahl index across HISCO occupation descriptions as a measure of task diversity (Column 6), and child labour as a proxy for de-skilling (Column 7). The main explanatory variable in all specifications is an indicator for a grid cell i having at least one rail station recorded in census year t, interacted with a dummy variable for the manufacturing sector. Log(Nb of Firms$_{i,s,t}$) is the natural logarithm of 1+ the number of firms in a given cell-sector year. Geo controls are the straight-line distance to the nearest city with coal deposits, the nearest port, and the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

7.2 Organizational Changes – Market Level

Appendix Table B.5 complements these findings by taking a closer look at inter-firm dynamics using data on firm owners in the I-CeM dataset. It investigates whether increased competition with large firms explains our results. Results indicate that a connection to the rail leads to a larger increase in bankruptcies when a sector's largest firms are present at a medium distance. Very close or very far large firms do not increase bankruptcies upon the arrival of the rail.[20] Hence, beyond local conditions, people's probability to experience bankruptcy increases if the railway brings them closer to large firms at a medium distance in their sector.

Table 6 investigates whether the rail's treatment effect also differs over time. It compares the effect of a rail connection in manufacturing for "first movers", i.e. locations connected already in 1851, to "late movers", i.e. locations connected by 1861 and later.

To focus on the variation over time, the specifications in Columns 6.1 and 6.5 add Sector

[20]See Appendix B for more details.

$\times$ Location fixed effects. These fixed effects account for the impact of the rail network in 1851 on the manufacturing sector. Therefore, the coefficients on $\mathbb{1}$ (Rail$_{i,t} > 0$) $\times$ Manufacturing$_s$ only measure the effect of getting connected after 1851. For both bankruptcies and employment, the coefficients are significant and positive. Getting connected to the rail increases bankruptcies by 73% and employment by 5%, which signals *local reallocation* over time. The remaining columns distinguish the timing of treatment in more detail. Columns 6.2 and 6.6 consider all locations treated in all years if they have at least one rail station in 1851. In a similar vein, Columns 6.3 and 6.7 code the binary rail variable as one for grid cells with at least one station in 1861 for the years 1861 and 1881. In Columns 6.4 and 6.8, the treatment is equal to one in the year 1881 for all grid cells that have at least one station in 1881, and zero otherwise.

Table 6: First Movers and the Effect of Railway Expansion

Dependent Variable:	#Bankruptcies$_{i,t,s}$				#Employed$_{i,t,s}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$ (Rail$_{i,t} > 0$) $\times$ Manufacturing$_s$	0.55***				0.05**			
	(0.08)				(0.02)			
$\mathbb{1}$ (Rail$_{i,1851} > 0$) $\times$ Manufacturing$_s$		0.17***				0.27***		
		(0.06)				(0.06)		
$\mathbb{1}$ (Rail$_{i,1861} > 0$) $\times$ Manufacturing$_s$			0.38***				0.27***	
			(0.08)				(0.07)	
$\mathbb{1}$ (Rail$_{i,1881} > 0$) $\times$ Manufacturing$_s$				0.56***				0.30***
				(0.19)				(0.12)
Observations	6280	7473	7473	7473	8616	8632	8632	8632
Pseu. R^2	.802	.813	.813	.813	.992	.929	.929	.928
Geo FE	✓	✓	✓	✓	✓	✓		
Sector $\times$ Year FE		✓	✓	✓		✓	✓	✓
Sector $\times$ Geo FE	✓				✓			
Sector Emp$_{i,t,s}$	✓	✓	✓	✓				
Pop$_{i,t}$					✓	✓	✓	✓
Geo Controls$_i$ $\times$ Manufacturing$_s$	✓	✓	✓	✓	✓	✓	✓	✓
Banks$_i$ $\times$ Manufacturing$_s$	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variables are the annualized number of bankruptcies that occurred between two census years (Columns (1) to (3)), and the number of employed people in a census year (Columns (4) to (6)). The main explanatory variables are indicators for a grid cell i having at least one rail station recorded (in 1851, in 1861 and in 1881) from 1851, 1861 and 1881, interacted with an indicator variable for the manufacturing sector. Geo controls are the straight-line distance to the nearest city with coal deposits, to the nearest port, and the city of London. *Banks$_i$* refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

All coefficients are significant at the one percent level, but the size of the coefficients increases over the years. Accordingly, the rail generates bankruptcies in early- and late-connected cells, but the impact becomes stronger over time. A connection to the rail from 1851 onwards increases bankruptcies by 19 percent. A connection to the rail in 1881 increases bankruptcies by 75 percent.[21]

The impact of a rail connection on manufacturing employment is constant over the years, ranging between 31 and 35 percent (Columns 6.5 to 6.8). Hence, no matter the timing of the connection, the rail exhibits a constant effect on structural change and the dynamism of the manufacturing sector.

Taken together, these results suggest that early- and late-connected locations experience market integration differently. Firms in early-connected locations are less exposed to competition but already benefit from market integration. In these locations, structural transformation reallocates labour towards the manufacturing sector with only a mild increase in bankruptcies. Later-connected locations witness a similarly growing manufacturing sector, but for them the already-stronger competition provokes a higher number of bankruptcies.

7.3 Mechanisms: Discussion and Alternative Explanations

The main interpretation of our results centers on organizational changes. Because of organizational changes, some actors face financial distress following market integration while their neighbours working in the same sector benefit from integration. Several other mechanisms may also explain the same patterns of increasing bankruptcies and employment. In this section, we discuss our results in the light of three alternative explanations: Creditors' incentives, workers' entry and exit, and credit supply.

Creditors' incentives. With the arrival of the rail, investors might put higher pressure on debtors because of better re-investment alternatives. By triggering bankruptcies faster, they would get at least part of the debtors' assets.

As a test for this alternative explanation, we leverage discontinuities in the bankruptcy procedure generated by two reforms of bankruptcy laws in 1869 and 1883. In 1869, England repealed the "officialism" doctrine for bankruptcies (Lester, 1991). Before this reform, bankruptcies were managed by local courts, often taking a long time to resolve. Moreover, the outcome was uncertain. Creditors' incentives to file for bankruptcy to "reinvest" could not impact the number of bankruptcies in this period since they were not managing bankruptcies. The 1869 reform allowed creditors to directly manage bankruptcies. This procedure advantaged creditors and increased their incentives to file for bankruptcies for quick reinvestment. In 1883, England went back to the "officialism" doctrine.

[21] According to $(e^{0.56} - 1) \cdot 100\% = 75.07\%$.

The evolution of the number of bankruptcies over time (Figure 1) illustrates the first fact about the repeal of "officialism." Creditors indeed file bankruptcies more than under "officialism," with two discontinuities in the number of bankruptcies at the repeal and re-introduction of "officialism." Bankruptcies before the 1869 reform and after the 1883 reform can be considered an imprint of the debtors' financial situation. The increase between the two reforms captures the interests of creditors. After the 1883 reform, the number of bankruptcies returns to the pre-1869 reform level lending credence to our interpretation that the surge in bankruptcies in the 1869–1883 period is mainly due to the repeal of "officialism" and variation in creditors' incentives to file for bankruptcies.

If creditor incentives would explain the baseline results, a connection to the rail would increase the number of manufacturing bankruptcies even more when "officialism" is repealed. We test this hypothesis by recoding our dataset and investigating the number of bankruptcies at the yearly level. Table 7 adds a triple-interaction $\mathbb{1}(\text{Rail}_{i,t} > 0) \times \text{Manufacturing}_s \times \text{Reform}_t$ on top of previous estimations. This triple interaction investigates whether the shift in incentives towards creditors impacts the main effects.[22] Table 7 shows that this triple-interaction does not turn significant. Moreover, the magnitude of the treatment variable is similar to the main results even though this approach uses finer fixed effects (sector-year FE instead of sector-census year FE) and a dataset with yearly observations. The main results are hence driven by the financial situations of debtors more than by the incentives of creditors.

Worker transitions across sectors Railway access increases the flow of workers into the manufacturing sector. However, censuses only provide the stock of workers. Theoretically, workers could enter the manufacturing sector after one census wave, suffer bankruptcy, and exit the sector again before one observes them in the next census. In such a case, the employment variable would not capture all the flows of workers potentially exposed to bankruptcies between two census waves. Appendix Table C.11 minimizes these concerns. The results use bankruptcies within only two years of a census wave as the dependent variable in the same specification as the baseline results. Another test further confirms this intuition. Appendix Table C.2 controls for county migrants in a cell (i.e. individuals born in a different county than the one they live in). Estimates remain significant and of a magnitude similar to the baseline estimates.

[22] We restrict our sample to 1861 to 1881 to focus on only two different bankruptcy regimes. Before 1861, only individuals with a status of "merchant" could file for bankruptcy.

Table 7: Incentives and the 1869/1883 Reforms

	Dependent Variable: #Bankruptcies$_{i,t,s}$	
	(1)	(2)
$\mathbb{1}$ (Rail$_{i,t}$ > 0) × Manufacturing$_s$	0.40***	0.38***
	(0.12)	(0.13)
$\mathbb{1}$ (Rail$_{i,t}$ > 0) × Man$_s$ × Ref$_t$	0.07	0.08
	(0.12)	(0.12)
Observations	97200	96900
Pseu. R^2	.797	.797
Geo FE	✓	✓
Sector × Year FE	✓	✓
Sector Emp$_{i,t,s}$	✓	✓
Geo Controls$_i$ × Manufacturing$_s$		✓
Banks$_i$ × Manufacturing$_s$		✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t, interacted with an indicator for the manufacturing sector and an indicator for years 1869–1881 when the "officialism" system was repealed. Geo controls are the straight-line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Credit Supply. Railway access could impact access to credit, which might correlate with the number of bankruptcies because of credit oversupply in manufacturing for example. Our baseline estimations already control for the number of country banks from Heblich and Trew (2018), interacted with an indicator variable for the manufacturing sector. Appendix Table B.6 adds a triple interaction between our railway × manufacturing indicator and the number of country banks. Checking for non-linear results depending on credit supply determines whether bankruptcies occur due to financial pressure, which credit supply may release (see Del Angel et al., 2024). If anything, the presence of country banks cushions the effect of railway access on bankruptcies in the manufacturing sector. Credit supply does not explain our effect, but rather prevents bankruptcies. The effect of this interaction moreover suggests that, if anything, the rail deteriorates the financial situation of workers in the manufacturing sector. One way to release this pressure is to facilitate access to credit.

This section illustrates that within-sector reallocation is dynamic. Railway access transforms the nature of labour in manufacturing firms, changes firm characteristics, and affects dynamics between firms. Manufacturing firms connected to the rail are larger and employ a wider variety of occupations as well as more unskilled labour. Moreover, the number of firms – especially those of medium size – increases. Market structures also mediate this effect.

Within sectors, bankruptcies are more numerous in locations in medium distance to large firms, and in locations that receive a rail connection later in the sample. The rail hence changes the dynamics within firms and between firms.

Importantly, we do not rule out that other factors beyond organizational changes between and within firms might explain why the rail increases bankruptcies in the manufacturing sector. Theoretically, creditors' incentives, workers' repeated entry and exit, and credit supply are important drivers of the effect. Focusing on within-sector, within-geography variation, we do not observe that these factors mitigate the effect of the rail. These other channels may, however, offer policy solutions to detrimental organizational changes. Bankruptcies skyrocketed when creditors were in charge of handling them. Similarly, access to credit cushions the effect of the rail in manufacturing. These results call for future research considering the factors that could solve part of the detrimental consequences of within-sector-space reallocation.

8 Conclusion

Market integration provides an opportunity for growth (Jedwab et al., 2017; Donaldson, 2018). This opportunity rests on exploiting economies of scale and a new organization of labour. This paper shows that this reorganization generates *local reallocation* – the destruction of the old economic structure while a new one is created in the same sector and place. To document this effect during market integration, we connect the literature on firms' organization during the industrial revolution (Juhász et al., 2024) to the literature on the effect of market integration (Melitz, 2003; Bogart, 2014). This approach still has value today as the merits of globalization are being questioned (Autor et al., 2020) and concerns on the future organization of firms arise (Varian, 2019).

Leveraging high-dimensional data, our estimates rigorously estimate the argument of Schumpeter (1942) and Melitz (2003). To document *local reallocation*, our study leverages a perfect setting to understand how market integration may prompt a biased firm reorganization within some sectors: the extension of the railway in England and Wales during the 19$^{\text{th}}$ century. The rapid expansion of the rail did not impact all sectors in the same way. In the manufacturing sector, firms needed to change their organization to fully exploit the division of labour and the economies of scale allowed by railways (Atack et al., 2008). This new organization was detrimental for some skills while generating aggregate welfare gains as theoretically proposed by Melitz (2003) and shown by Bogart et al. (2022).

Our results show that the expansion of the railway in Britain created financial distress among workers in the manufacturing sector. As firms reorganized to compete in larger markets, some skills lost value. At the same time, some of its workers experienced financial distress. Market integration brought *local reallocation*. The dimensionality of the effect refines previous estimates of reallocation. In particular, previous research has focused on the geographic nature of the reallocation from market integration (Redding, 2016), or the within-sector, between-places reallocation from market integration (Autor et al., 2013, 2016, 2020). Our results show reallocation within a location connected to the rail.

Market integration does not only destroy the old economic structure in declining areas and sectors. It also generates reallocation where the new economic structure develops. Beyond the theoretical argument, this result matters to better understand and maybe tackle the adversity of reallocation following market integration. Reallocation occurs even between citizens belonging to the same sector and located in the same area. Hence, policies thought to alleviate the negative consequences of market integration have to be thought out with that level of precision.

Local reallocation also matters to explain how spatial and sectoral inequality may interact today (Autor et al., 2020). This research emphasizes that despite positive aggregate effects, technology and trade are redistributive by nature. This redistribution has important (political) consequences (Frey et al., 2018; Lacroix, 2018; Autor et al., 2020; Caprettini and Voth, 2020). These results also clarify some of the dynamics driving the evolution of market structure, trade, and inequality during the Industrial Revolution and its immediate aftermath (Nye, 1987; O'Rourke and Williamson, 2005; Desmet and Parente, 2012; Desmet et al., 2020; Juhász et al., 2024).

Research must combine different indicators to better grasp the consequences of reallocation on the full distribution of workers and firms. Here, personal bankruptcies capture the destruction of the old economic structure despite aggregate gains in a given place and sector. Our research resonates with the 2030 Agenda for Sustainable Development, where UN member states pledged that "no one will be left behind".[23] Among the agenda's solutions to achieve this objective, two stand out: transport infrastructure and productivity boosts. We do see that the combination of both may actually leave some behind. Future research could build on these new results to better understand how to mitigate these distributional consequences of growth.

[23]See for more information https://sdgs.un.org/2030agenda, last visited Nov. 6th, 2024.

References

Atack, J., M. R. Haines, and R. A. Margo (2008). *Railroads and the Rise of the Factory: Evidence for the United States, 1850-1870*, Volume 14410. 2, 6, 25, 33

Atack, J., R. A. Margo, and P. W. Rhode (2019). "Automation" of manufacturing in the late nineteenth century: The hand and machine labor study. *Journal of Economic Perspectives 33*(2), 51–70. 7

Atack, J., R. A. Margo, and P. W. Rhode (2024). De-skilling: Evidence from late nineteenth century American manufacturing. *Explorations in Economic History 91*, 101554. 6, 7

Autor, D., D. Dorn, G. Hanson, and K. Majlesi (2020). Importing political polarization? The electoral consequences of rising trade exposure. *American Economic Review 110*(10), 3139–3183. 3, 4, 33, 34

Autor, D., D. Dorn, L. F. Katz, C. Patterson, and J. Van Reenen (2020). The Fall of the Labor Share and the Rise of Superstar Firms. *Quarterly Journal of Economics 135*(2), 645–709. 7, 34

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review 103*(6), 2121–2168. 2, 3, 5, 13, 34

Autor, D. H., D. Dorn, and G. H. Hanson (2016). The China shock: Learning from labor-market adjustment to large changes in trade. *Annual Review of Economics 8*, 205–240. 2, 3, 4, 34

Ayotte, K. (2007). Bankruptcy and entrepreneurship: The value of a fresh start. *Journal of Law, Economics, & Organization 23*(1), 161–185. 5

Banerjee, A., E. Duflo, and N. Qian (2020). On the road: Access to transportation infrastructure and economic growth in China. *Journal of Development Economics 145*, 102442. 4

Berger, T. (2019). Railroads and rural industrialization: Evidence from a historical policy experiment. *Explorations in Economic History 74*, 101277. 6

Berger, T. and K. Enflo (2017). Locomotives of local growth: The short-and long-term impact of railroads in Sweden. *Journal of Urban Economics 98*, 124–138. 6

Berger, T. and V. Ostermeyer (2024). Firm survival and the rise of the factory. *Economic History Review*. 6

Bernstein, S., E. Colonnelli, X. Giroud, and B. Iverson (2019). Bankruptcy spillovers. *Journal of Financial Economics 133*(3), 608–633. 5

Bogart, D. (2014). The transport revolution in industrialising Britain. *Cambridge economic history of modern Britain 1*, 368–391. 3, 9, 33

Bogart, D., X. You, E. J. Alvarez-Palau, M. Satchell, and L. Shaw-Taylor (2022). Railways, divergence, and structural change in 19th century England and Wales. *Journal of Urban Economics 128*, 103390. 2, 4, 6, 10, 15, 18, 21, 33, 70

Borusyak, K. and P. Hull (2023). Nonrandom exposure to exogenous shocks. *Econometrica 91*(6), 2155–2185. 20, 23, 67

Bose, U., S. Filomeni, and S. Mallick (2021). Does bankruptcy law improve the fate of distressed firms? The role of credit channels. *Journal of Corporate Finance 68*, 101836. 5

Brandt, L., J. Van Biesebroeck, L. Wang, and Y. Zhang (2017). Wto accession and performance of chinese manufacturing firms. *American Economic Review 107*(9), 2784–2820. 5

Braun, S. T. and R. Franke (2022). Railways, growth, and industrialization in a developing German economy, 1829–1910. *Journal of Economic History 82*(4), 1183–1221. 6

Bustos, P. (2011). Trade liberalization, exports, and technology upgrading: Evidence on the impact of mercosur on argentinian firms. *American economic review 101*(1), 304–340. 3, 5

Caprettini, B. and H.-J. Voth (2020). Rage against the machines: Labor-saving technology and unrest in industrializing England. *American Economic Review: Insights 2*(3), 305–20. 5, 34

Casson, M. (2009). *The world's first railway system: enterprise, competition, and regulation on the railway network in Victorian Britain.* OUP Oxford. 9, 23, 71

Chandler, A. D. (1977). *The visible hand.* Harvard university press. 7

Chiopris, C. (2025). Spatial Networks and the Diffusion of Ideas. *Job Market Paper (https://www.caterinachiopris.com/research-publications)*. 6

Costas-Fernández, J., J.-A. Guerra, and M. Mohnen (2020). Train to opportunity: The effect of infrastructure on intergenerational mobility. *Documento CEDE* (48). 24, 72

Crafts, N. (2004). Steam as a general purpose technology: A growth accounting perspective. *Economic Journal 114* (495), 338–351. 6, 10

Crafts, N. (2022). Slow real wage growth during the industrial revolution: Productivity paradox or pro-rich growth? *Oxford Economic Papers 74* (1), 1–13. 5

Davydenko, S. A. and J. R. Franks (2008). Do bankruptcy codes matter? A study of defaults in France, Germany, and the UK. *Journal of Finance 63* (2), 565–608. 5

Del Angel, M., G. Richardson, and M. Gou (2024). Bank failures and economic activity. Evidence from the Progressive Era. *Available at SSRN 4756705*. 5, 32

Desmet, K., A. Greif, and S. L. Parente (2020). Spatial competition, innovation and institutions: The industrial revolution and the Great Divergence. *Journal of Economic Growth 25*, 1–35. 34

Desmet, K. and S. L. Parente (2012). The evolution of markets and the revolution of industry: A unified theory of growth. *Journal of Economic Growth 17*, 205–234. 34

Dippel, C., R. Gold, S. Heblich, and R. Pinto (2021). The effect of trade on workers and voters. *The Economic Journal 132* (641), 199–217. 3

Donaldson, D. (2015). The gains from market integration. *Annual Review of Economics 7* (1), 619–647. 5

Donaldson, D. (2018). Railroads of the Raj: Estimating the impact of transportation infrastructure. *American Economic Review 108* (4-5), 899–934. 2, 5, 6, 33

Donaldson, D. and R. Hornbeck (2016). Railroads and American economic growth: A "market access" approach. *Quarterly Journal of Economics 131* (2), 799–858. 2, 5, 6, 19, 66, 67

Esteves, R. and G. G. Mesevage (2021). Private benefits, public vices: Railways and logrolling in the nineteenth-century British parliament. *Journal of Economic History 81* (4), 975–1014. 9

Faber, B. (2014). Trade integration, market size, and industrialization: Evidence from China's national trunk highway system. *Review of Economic Studies 81* (3), 1046–1070. 4

Feng, L., Z. Li, and D. L. Swenson (2017). Trade policy uncertainty and exports: Evidence from china's wto accession. *Journal of International Economics 106*, 20–36. 2

Fernihough, A. and K. H. O'Rourke (2020). Coal and the European Industrial Revolution. *Economic Journal 131*(635), 1135–1149. 12, 13

Flaaen, A. and J. R. Pierce (2019). Disentangling the effects of the 2018-2019 tariffs on a globally connected us manufacturing sector. 3, 5

Fogel, R. W. (1964). *Railroads and American economic growth.* Johns Hopkins Press Baltimore. 5

Frey, C. B., T. Berger, and C. Chen (2018). Political machinery: Did robots swing the 2016 US presidential election? *Oxford Review of Economic Policy 34*(3), 418–442. 34

Gine, X. and I. Love (2010). Do reorganization costs matter for efficiency? Evidence from a bankruptcy reform in Colombia. *Journal of Law and Economics 53*(4), 833–864. 5

Goldin, C. and L. F. Katz (1998). The origins of technology-skill complementarity. *Quarterly Journal of Economics 113*(3), 693–732. 5, 7

Gregory, I. N. and J. M. Henneberg (2010). The railways, urbanization, and local demography in England and Wales, 1825–1911. *Social Science History 34*(2), 199–228. 10

Hawke, G. R. (1970). *Railways and economic growth in England and Wales, 1840-1870.* Oxford University Press. 10

Heblich, S., S. J. Redding, and Y. Zylberberg (2024). The distributional consequences of trade: Evidence from the grain invasion. *NBER Working Paper No. 32958.* 4

Heblich, S. and A. Trew (2018). Banking and industrialization. *Journal of the European Economic Association 17*(6), 1753–1796. 12, 13, 32, 55

Hornbeck, R. and M. Rotemberg (2024). Growth off the rails: Aggregate productivity growth in distorted economies. *Journal of Political Economy 132*(11), 3547–3602. 2, 5, 6

Hornung, E. (2015). Railroads and Growth in Prussia. *Journal of the European Economic Association 13*(4), 699–736. 6

Humphries, J. (2013). Childhood and child labour in the British industrial revolution 1. *Economic History Review 66*(2), 395–418. 27

Jedwab, R., E. Kerby, and A. Moradi (2017). History, path dependence and development: Evidence from colonial railways, settlers and cities in Kenya. *Economic Journal 127*(603), 1467–1494. 6, 33

Juhász, R., M. P. Squicciarini, and N. Voigtländer (2024). Technology adoption and productivity growth: Evidence from industrialization in France. *Journal of Political Economy 132*(10), 3215–3259. 2, 5, 7, 8, 13, 15, 33, 34

Kaboski, J. P., W. Lu, W. Qian, and L. Ren (2024). Melitz meets Lewis: The impacts of roads on structural transformation and businesses. *NBER Working paper 4820042*. 6

Katz, L. F. and R. A. Margo (2014). Technical change and the relative demand for skilled labor: The United States in historical perspective. In *Human capital in history: The American record*, pp. 15–57. University of Chicago Press. 15

Kogan, L., D. Papanikolaou, L. D. Schmidt, and B. Seegmiller (2023). Technology and labor displacement: Evidence from linking patents with worker-level data. *NBER Working Paper No. 31846*. 5

Lacroix, J. (2018). Steam democracy up! Industrialization-led opposition in Napoleonic plebiscites. *European Review of Economic History 22*(2), 135–160. 34

Lester, V. M. (1991). *Insolvency and reform of English bankruptcy law, 1831-1914*. Ph. D. thesis, University of Oxford. 8, 30, 46

Li, B. and J. Ponticelli (2022). Going bankrupt in China. *Review of Finance 26*(3), 449–486. 5

Magee, G. B. (1997). *Productivity and performance in the paper industry: Labour, capital and technology in Britain and America, 1860-1914*. Number 4. Cambridge University Press. 17

Martí-Henneberg, J., M. Satchell, X. You, L. Shaw-Taylor, and E. Wrigley (2017a). 1851 England, Wales and Scotland railway stations shapefile. 12

Martí-Henneberg, J., M. Satchell, X. You, L. Shaw-Taylor, and E. Wrigley (2017b). 1861 England, Wales and Scotland railway stations shapefile. 12

Martí-Henneberg, J., M. Satchell, X. You, L. Shaw-Taylor, and E. Wrigley (2017c). 1881 England, Wales and Scotland railway stations shapefile. 12

Maw, P., P. Solar, A. Kane, and J. S. Lyons (2022). After the great inventions: technological change in uk cotton spinning, 1780–1835. *The Economic History Review 75*(1), 22–55. 17

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica 71*(6), 1695–1725. 2, 4, 6, 7, 13, 33

Mitchell, B. R. (1964). The coming of the railway and United Kingdom economic growth. *Journal of Economic History 24*(3), 315–336. 10

Nye, J. V. (1987). Firm size and economic backwardness: A new look at the French industrialization debate. *Journal of Economic History 47*(3), 649–669. 34

O'Rourke, K. H. and J. G. Williamson (2005). From Malthus to Ohlin: Trade, industrialisation and distribution since 1500. *Journal of Economic Growth 10*, 5–34. 34

Pavcnik, N. (2002). Trade liberalization, exit, and productivity improvements: Evidence from chilean plants. *The Review of economic studies 69*(1), 245–276. 3, 5, 8

Pilbin, P. (1937). External relations of the tyneside glass industry. *Economic Geography 13*(3), 301–314. 18

Ponticelli, J. and L. S. Alencar (2016). Court enforcement, bank loans, and firm investment: Evidence from a bankruptcy reform in Brazil. *Quarterly Journal of Economics 131*(3), 1365–1413. 5

Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics 101*, 148–167. 2, 34

Rhodes, J. (2025). Geo-coding addresses in historic British census data: An open methodology. *Historical Methods: A Journal of Quantitative and Interdisciplinary History 58*(1), 31–53. 45

Schumpeter, J. A. (1942). *Capitalism, socialism and democracy (2003 Edition)*. Routledge: London. 2, 13, 33

Schurer, K. and E. Higgs (2023). Integrated Census Microdata (I-CeM) Names and Addresses, 1851-1911. *UK Data Service SN: 7856*(2nd Edition), DOI: http://doi.org/10.5255/UKDA–SN–7856–2. 11, 27

Solar, P. M. and J. S. Lyons (2011). The english cotton spinning industry, 1780–1840, as revealed in the columns of the london gazette. *Business History 53*(3), 302–323. 5, 17

Tang, J. P. (2014). Railroad expansion and industrialization: Evidence from Meiji Japan. *Journal of Economic History 74*(3), 863–886. 6, 25

Temin, P. (1997). Two views of the British industrial revolution. *Journal of Economic History 57*(1), 63–82. 5

Toms, J. (1997). Textile industry, 1870-1914. *Business And Economic History 26*, 323–329. 17

Trefler, D. (2004). The long and short of the canada-us free trade agreement. *American Economic Review 94*(4), 870–895. 3, 5

Tsiachtsiras, G. (2022). Changing the perception of time: Railroads, access to knowledge and innovation in nineteenth century France. *Access to Knowledge and Innovation in Nineteenth Century France*. 6

Varian, H. (2019). Artificial intelligence, economics, and industrial organization. *Economics of Artificial Intelligence: An Agenda*, 399. 33

A Appendix A – Data Construction

A.1 Detecting Bankruptcies

We extracted personal bankruptcy announcements using a combination of Optical Character Recognition (OCR) software and text recognition algorithms. After web-scraping the 5,063 London Gazette issues that were published between 1850–1890 from the London Gazette homepage, we converted the scanned images to readable text via OCR software. Next, we wrote various text recognition algorithms to search each issue for bankruptcy announcements. We found at least one such announcement across 4,086 issues. The other issues were mostly supplements, which provide special information outside the Gazette's regular announcements.

We had to adapt our text recognition algorithms to different periods of the Gazette. Our goal was to identify individual announcements based on specific keywords. The layout of the announcements and, hence, the relevant keywords, however, changed over time. For example, until 1861 the London Gazette listed bankruptcy announcements toward the end of an issue. Each announcement received its own paragraph, starting with the introduction "Whereas a Commission of Bankrupt(cy) was awarded (and issued forth) against." Starting in 1861, the sections of bankruptcy announcements received their own headlines and internal structure. Since then, announcements have become separated into first meetings, i.e. the assessment of bankruptcy and collection of claims, later meetings to distribute funds, and final meetings to resolve open cases. For example, first meetings would be introduced under the headline "The Bankruptcy Act, 1861. Notice of Adjudications and First Meeting of Creditors." The London Gazette maintained this structure for most of the time.

To extract individual announcements from an issue, we wrote various algorithms that, depending on the announcement pattern, identified the start of a new announcement. For example, in the early issues until 1861, the algorithm looked for different variations of the text pattern "Whereas a Commission of Bankrupt is awarded and issued forth against", to determine the start of a bankruptcy announcement.[24] From 1861 onwards, we searched the issues for the headlines introducing the "First Meetings" of bankruptcies to focus our algorithm on the text between this headline and the following one, and then collecting the individual announcements with the procedure explained above.

[24] The actual pattern switch occurred with the new bankruptcy act in the issue 22,564 from November 12th, 1861. While the overall pattern remained stable across announcements, the individual solicitors who published the announcements would vary the text pattern somewhat, e.g. using past tense ("was awarded" or "has been issued forth") or dropping the "awarded" or "issued forth" part of the introduction. We went through several issues manually to include as many variations as possible in our algorithm. We returned to issues with an unusually low number of detected announcements to look for pattern variations that we might have overlooked.

Our algorithm detected a total of 134,702 bankruptcy cases, i.e., on average, 33.8 bankruptcy announcements per issue, with a median of 24 announcements per issue. For each bankruptcy case we detected, we extracted the first 300 letters after starting a bankruptcy paragraph for further processing. Within each text sample, we let our algorithm find the information on a) the name of the person, b) the person's current address, and c) the person's current occupation. To identify this information, we used detected commas in the text to separate the information. Usually, the information would be presented in the format *name, address, occupation*, so detecting commas as breakpoints helped structure the text. Using these comma-break points as general hints for where to look for certain information, we ran the specific text subsets against lists of city-, county-, borough-, and parish names and a list of (historical) census occupations, respectively, to detect matches.

Due to the occasionally bad quality of the scans, this required important pre-processing. Among other things, we corrected common typos that the OCR introduced by misreading certain letters and used fuzzy text matching procedures where direct pattern matching did not yield a result. Finally, we used the information on locations and occupations to encode it in a usable format. We geocoded the place information as accurately as possible. We could link many locations to the coordinates for a specific parish or city, but some, we could only geocode at the county level. Our analysis only uses bankruptcy cases that we could link to the city-level or below. To make use of the occupation titles, we assigned them to 5-digit historical international classification of occupations (HISCO) codes as defined by the International Institute of Social History Amsterdam.[25] Despite the pre-and post-processing steps, we were not able to acquire full information for all bankruptcy cases that our algorithm collected. We could geocode 130,203 bankruptcy cases (121,936 cases to the city- or parish-level) and assign HISCO codes to 124,363 cases.

A.2 Assigning sectors

The original HISCO coding divides occupations into ten main categories. For our purposes, we rearranged these ten groups slightly. First, we combine the groups "0" and "1", which refer to "Professional Workers," with the "Services" category. Next, we combine the groups "7", "8" and "9", which all refer to "Production and related workers", into one "Manufacturing" group. We leave the groups for "Agriculture" ("6"), "Services" ("5"), and "Trade" ("4") as is. Finally, we distribute the groups "2" ("Administrative workers") and "3" ("Clerical workers") into our four main groups based on the occupation category the I-CeM dataset assigned

[25]See their homepage https://iisg.amsterdam/en/data/data-websites/history-of-work for further information

to the individuals within these groups. For example, we assign people with an occupation description of “working and dealing with metals” to manufacturing, and “persons engaged in commercial occupations” to trade. In total, we rearranged 10 occupation categories this way. Note that all our regressions include sector fixed effects such that these coding decisions to not impact our estimates.

From jobs to sectors. Table A.1 illustrates how we assigned job descriptions in the bankruptcy announcements to the main sector groups that we use in our analysis. We mapped each occupation description into a HISCO code, and then combined the ten HISCO categories into the four main sectors we use in the analysis. Table A.1 gives some examples of common job descriptions that we encountered in the bankruptcy announcements.

Table A.1: Sector Definitions

Sector	Description
Agriculture	All occupations that involve farm work, garden work, fishing, and animal husbandry. Common job descriptions are, for example, “farmer”, "gardener”, “dairyman”, “nurseryman”.
Manufacturing	Occupations in this sector turn raw materials into intermediate or final goods, or work in the extraction of non-agricultural natural resources. Common job descriptions are, for example, “butcher”, “baker”, “brewer”, “tailor”, “engineer”, “plumber”, “jeweller”, “smith”, “shoe maker”, “builder” or “manufacturer”.
Trade	Workers in this category work in retail sales, as traveling sales agents, shop keepers, or on the import and export of goods from/to other countries. Common job descriptions are, for example, “chapman”, “merchant”, “dealer”, “grocer”, “salesman”, “retailer”, and “tobacconist”.
Services	This category combines jobs that require specific skills, but do not produce physical goods. This includes education and military personnel, artists, medical practitioners, managers, gastronomy workers, clerical workers, and other service providers. Common job descriptions are, for example, “victualler”, “inn keeper”, “hair dresser”, “contractor”, “surgeon”, “attorney”, “schoolmaster”, “artist”, “registrar”, “house keeper”, “accountant”, “clerk”, and “secretary”.

A.3 Assigning Census Locations

To match the I-CeM census observations with our grid cell dataset, we geocoded all observations based on the centroid of their subdistricts. Subdistricts are rather small in urban areas, but can extend over bigger areas in rural locations. We hence end up with some grid cells in rural regions, especially in Wales, that do not contain a subdistrict centroid due to the large size of sub-districts in these areas. We drop these grid cells from our estimations. Despite decreasing sample size, it allows us to include information on all households surveyed by the census in our estimations.

Rhodes (2025) uses exact addresses to achieve a higher spatial resolution. However, for the 1851-1881 censuses we use in our regressions, this new method achieved only a coverage of 50-60% of surveyed households. And as Rhodes (2025) further notes, coverage is higher in urban locations and among people with a manufacturing occupation. It is further fair to assume that the quality of geocoded addresses correlates with a location's railway access. Therefore, we base our regressions on census data geocoded at the sub-district level, accepting somewhat less precise estimates to avoid a potential bias from sample selection biased towards more urban and connected locations.

A.4 Descriptive Statistics

Table A.2: Descriptive Statistics

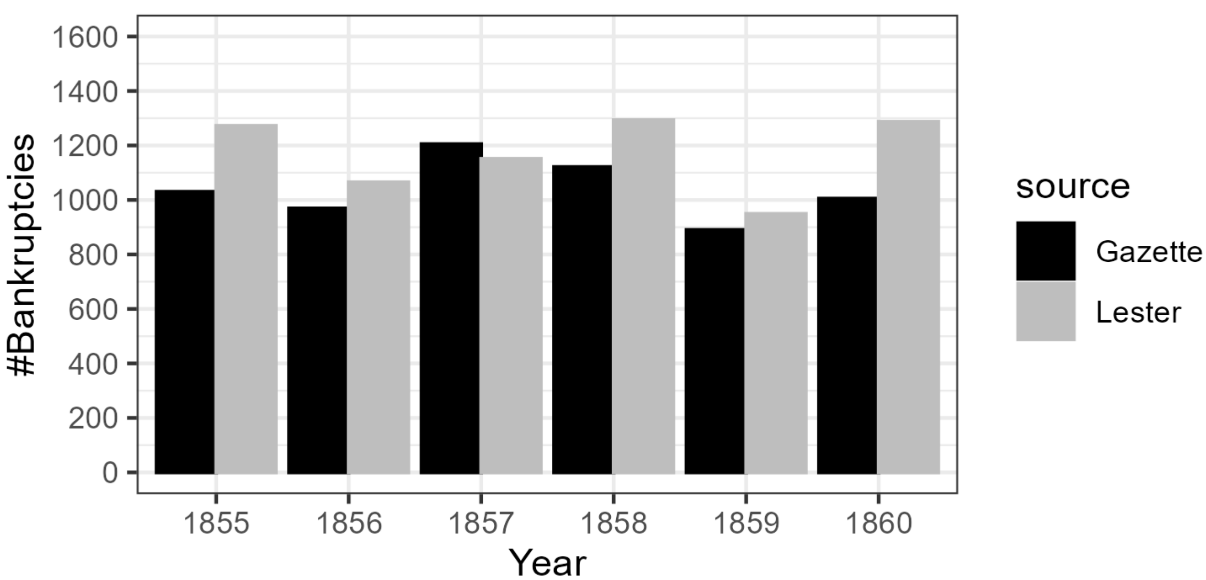
Variable	Observations	Min	Median	Mean	SD	Max
Annual Bankruptcies	10880	0.00	0.00	0.67	4.73	173.85
Annual Bankruptcies, 2-year Window	10880	0.00	0.00	0.69	6.25	253.20
Employment in Sector	10880	0.00	1566.00	3674.21	10221.98	228390.00
Rail Access	10880	0.00	1.00	0.73	0.45	1.00
Distance to London	10880	7.77	213.14	214.16	106.93	505.72
Distance to Coal	10880	0.00	52.43	67.66	58.63	252.34
Distance to Port	10880	1.05	22.42	27.97	21.38	95.54
Self-employed Individuals	10880	0.00	66.00	197.87	592.26	14852.00
Child Employment	10880	0.00	4.40	228.60	1070.77	38085.00
Herfindahl Index, Occupations	7554	1056.01	4652.41	5148.28	2824.73	10000.00
Unemployment Share	10430	0.01	0.32	0.32	0.09	0.99
Waterway Length	10880	0.00	4.30	11.91	16.19	116.05
Turnpike Length	10880	0.00	90.79	90.82	43.80	258.45
Number of Firms	10880	0.00	0.00	6.96	33.54	1435.00
Population	10880	0.00	13938.50	26912.86	61235.83	1042426.00

B Appendix B – Supporting Evidence

This Appendix section provides additional graphs and tables to complement the baseline results. It further details the dataset used in our regressions, and displays additional results from important variations of our main specifications. The last tables help the discussion of mechanisms.

Completeness of bankruptcy records. Potentially, the digitized bankruptcy announcements from the London Gazette might not represent the universe of bankruptcies. The coding of bankruptcies naturally misses some cases. Some cases might be missed by our algorithm where the OCR software misidentifies crucial words. For a few cases of bankruptcy announcements, assigning an occupation or place was impossible. We were able to assign an occupation to around 92% of the bankruptcies in our dataset. We were able to assign coordinates at the town- or parish level in 91% of cases. To investigate the comprehensiveness of our dataset, Figure B.1 compares the yearly number of bankruptcies in our dataset to officially published statistics at the national level as collected by Lester (1991). Our coding follows the general trend very closely. Moreover, the number of bankruptcies in our dataset is close to the benchmark provided in Lester (1991). Sampling bias is therefore unlikely to affect our estimations other than by increasing standard errors due to random measurement error.

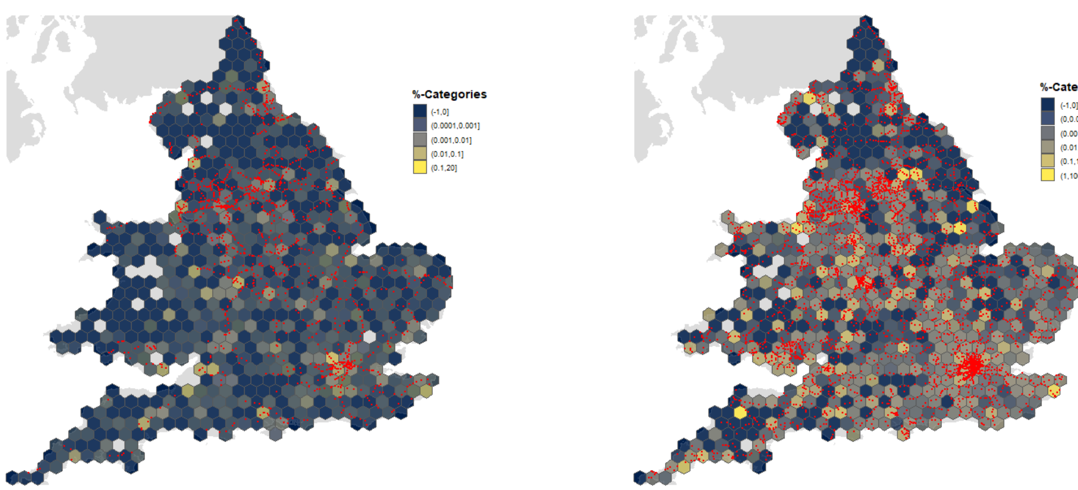
Figure B.1: Comparison to National Statistics



Notes: This figure displays yearly aggregates of the bankruptcy cases in our dataset (black), and compares them to official national statistics collected by Lester (1991) in grey.

Dataset. Figure B.2 illustrates our empirical analysis. The two maps show England and Wales covered by hexagonal grid cells with an edge-length of 16km, our spatial unit of observation. We spatially intersect the geocoded bankruptcy- and census-information with these grid cells, and then aggregate the variables of interest at the cell level. Grey cells indicate missing data from the I-CEM census data due to too large subdistricts. Colours indicate the share of bankruptcies with respect to the location's total employment, where we assign the shares to categories for ease of display. The red dots indicate the locations of railway stations. Both maps illustrate the intuition behind our analysis. In 1851, the British railway system was still in its infancy. Only 56% of cells had at least one railway station, and the overall density of railway stations was still low. Similarly, only a small number of bankruptcies occurred between 1851 and 1860. Many grid cells not even experienced one. Of those grid cells that experienced bankruptcies, almost all contain at least one railway station. In 1881, the railway network was much more advanced. More than 90% of cells had at least one railway station. And not only does the overall number of bankruptcy cases increase; we also see many grid cells lighting up now that did not exhibit any bankruptcy cases in the period before. Yet, bankruptcy cases still closely trace the spatial extent of the railway network.

Figure B.2: Bankruptcy Rates and Railway Expansion



(a) Bankruptcies 1851-1860, Rail Stations 1851.

(b) Bankruptcies 1881-1890, Rail Stations 1881.

Notes: The figures show the share of bankruptcies in total employment by location. Brighter colours indicate higher shares of bankruptcies. The red points indicate railway stations that were established at the beginning of the respective data sample. Light-grey locations are low-populated places and were omitted from the dataset because they do not contain a census sub-district, so we lack any census information for these cells.

Additional results. Several other results support our approach. Table B.1 reports the results from the main specifications using sectoral employment as the dependent variable. Instead of directly comparing manufacturing to all other sectors, Tables B.2 and B.3 report the effect of the railway expansion on bankruptcies and employment, respectively, by sector. Moving to within-manufacturing variation, Figures B.3 and B.4 provide coefficients from the main specifications with, respectively, bankruptcies and employment at the 2-digit manufacturing sector level as dependent variables. Table B.4 provides further statistics illustrating firm heterogeneity in the manufacturing sector.

Table B.1 shows that railway access significantly increased employment, especially in the manufacturing sector. According to Column B.1.1, places connected to the rail network witnessed higher overall employment. Column B.1.2 suggests that this effect is larger for the manufacturing sector. Once we add the different fixed effects and the control variables, the coefficient attached to the rail variable turns negative (Columns B.1.3 to B.1.6). The coefficient for the interaction of the manufacturing sector dummy variable with the rail dummy variable remains significantly positive across all columns. The manufacturing sector in cells connected to the rail has almost 20% more employment than in not connected cells.[26]

[26] We derive this number from the joint effect of the baseline railway effect plus the manufacturing-specific railway effect: $(e^{(-0.11+0.28)} - 1) \cdot 100\% = 18.5\%$

Table B.1: Main Results - The Effect of the Rail on Employment

		Dependent Variable: #Employed$_{i,t,s}$				
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$ (Rail$_{i,t}$ > 0)	1.18***	0.91***	-0.42***	-0.27***	-0.11***	-0.10***
	(0.11)	(0.10)	(0.06)	(0.04)	(0.03)	(0.03)
$\mathbb{1}$ (Rail$_{i,t}$ > 0) × Manufacturing$_s$		0.60***	0.60***	0.60***	0.27***	0.26***
		(0.08)	(0.08)	(0.08)	(0.07)	(0.07)
Observations	8704	8704	8704	8704	8632	8632
Pseu. R^2	.0744	.215	.873	.906	.926	.929
Geo FE			✓	✓	✓	✓
Sector FE			✓	✓	✓	
Year FE			✓	✓	✓	
Sector-Year FE						✓
Sector Employment$_{i,s,t}$				✓	✓	✓
Geo Controls$_i$ × Manufacturing$_s$					✓	✓
Banks$_i$ × Manufacturing$_s$					✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the number of people employed in sector s at census year t and in grid cell i. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the cell's total population and the straight-line distance to the nearest city with coal deposits, to the nearest port, and the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$

Our main specifications compare the effects of the railway expansion on bankruptcies and employment of manufacturing to all three other sectors. This treats all three other sectors as one homogeneous control group. In Tables B.2 and B.3, we allow for more cross-sectoral variation by providing estimates of the railway effect on each sector separately. From both Tables, the final column represents the estimates reported in Figure 3 above.

Table B.2: Estimates by Sector – Bankruptcies

	Dependent Variable: #$\text{Bankruptcies}_{i,t,s}$		
	(1)	(2)	(3)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Agr_s	-0.10	0.32*	-0.04
	(0.19)	(0.19)	(0.17)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s	0.56***	0.56***	0.37***
	(0.13)	(0.14)	(0.13)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Trade_s	0.11	-0.04	0.11
	(0.16)	(0.15)	(0.15)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Services_s	0.07	0.11	-0.09
	(0.15)	(0.16)	(0.14)
Observations	7481	7473	7473
Pseu. R^2	.801	.805	.816
Geo FE	✓	✓	✓
Sector FE	✓	✓	
Year FE	✓	✓	
Sector × Year FE			✓
Sector $\text{Employment}_{i,s,t}$	✓	✓	✓
Geo Controls_i × Manufacturing_s		✓	✓
Banks_i × Manufacturing_s		✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t, interacted with an indicator variable for the observed sector. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

According to Table B.2, the other three sectors do not experience a robust railway-effect. Only in manufacturing, the railway produces a significant increase in bankruptcies. In the other three sectors, the effects are insignificant. When looking at employment in Table B.3, the picture looks different. The agricultural sector experienced a large and significant decrease in employment upon railway connection across all specifications. The services and trading sectors experienced significant increases in employment, as does manufacturing. Hence, our main employment-regressions compare the growing manufacturing sector to a mixture of growing trade- and services sectors and a declining agricultural sector.

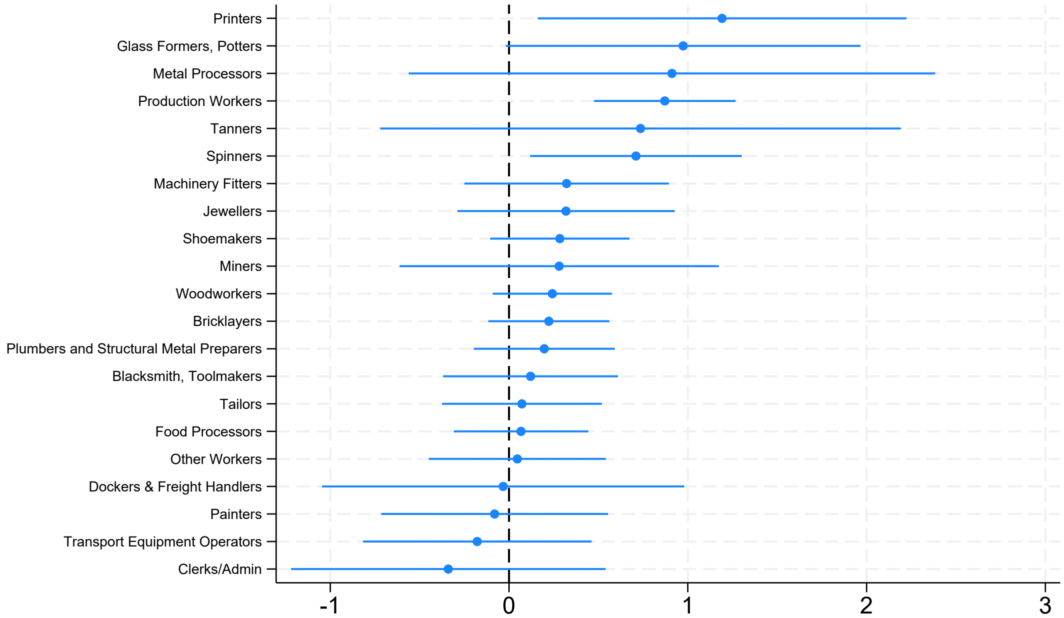
Table B.3: Estimates by Sector – Employment

	Dependent Variable: #Employed$_{i,t,s}$		
	(1)	(2)	(3)
$\mathbb{1}$ (Rail$_{i,t} > 0$) × Agr$_s$	-0.84***	-0.35***	-0.24***
	(0.09)	(0.05)	(0.06)
$\mathbb{1}$ (Rail$_{i,t} > 0$) × Manufacturing$_s$	0.33***	0.15***	0.16***
	(0.05)	(0.04)	(0.04)
$\mathbb{1}$ (Rail$_{i,t} > 0$) × Trade$_s$	0.52***	0.35***	0.25***
	(0.06)	(0.05)	(0.05)
$\mathbb{1}$ (Rail$_{i,t} > 0$) × Services$_s$	0.15***	0.13***	0.11***
	(0.05)	(0.03)	(0.03)
Observations	8704	8632	8632
Pseu. R^2	.913	.949	.951
Geo FE	✓	✓	✓
Sector FE	✓	✓	
Year FE	✓	✓	
Sector × Year FE			✓
Population$_{i,t}$	✓	✓	✓
Coal$_i$ × Manufacturing$_s$		✓	✓
Port$_i$ × Manufacturing$_s$		✓	✓
London$_i$ × Manufacturing$_s$		✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the number of people employed in sector s at census-year t in location i. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t, interacted with an indicator variable for the observed sector. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

Figures B.3 and B.4 repeat these sector-specific estimations, but consider only variation across subsectors within the manufacturing sector. All reported coefficients were derived from our main specifications, including all controls and interacted fixed effects. The coefficients from Figures B.3 and B.4 further resemble the coefficients reported in Figure 4 above.

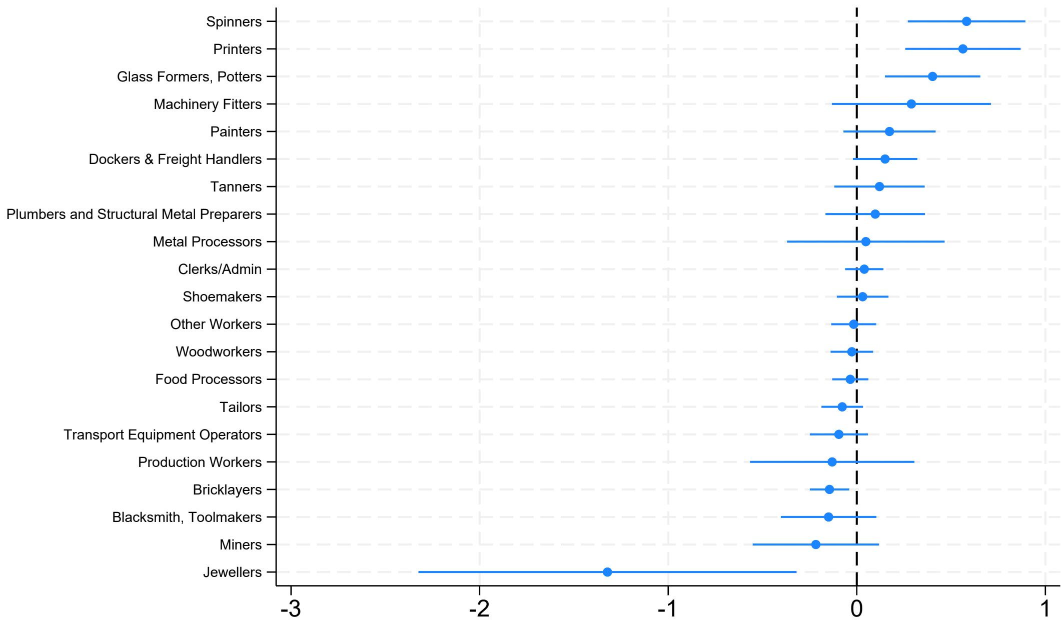
Figure B.3: Within-Manufacturing estimates: Bankruptcies



This figure plots coefficient estimates from PPML regressions following Equation 1 at the occupation level and for the subset of manufacturing occupations. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The specification includes all control variables and fixed effects as in Table 1, Column (6). Whiskers report 95% confidence intervals.

The two figures show a significant heterogenenity in the railway effect across subsectors. The professions of printers, glass formers, spinners, and general production workers experience a strong and significant increase in bankruptcies once they are connected to the railway network. Most of these professions however also experienced an increase in employment.

Figure B.4: Within-Manufacturing estimates: Employment



This figure plots coefficient estimates from PPML regressions following Equation 1 at the occupation level and for the subset of manufacturing occupations. The dependent variable is the number of people employed in sector s at census-year t in location i. The specification includes all control variables and fixed effects as in Table 1, Column (6). Whiskers report 95% confidence intervals.

Table B.4 provides more information on the manufacturing sector. It illustrates important differences of manufacturing firms compared to other sectors already at the beginning of our sample. In the manufacturing sector, the standard deviation in the number of employees was 4 to 10 times bigger than in other sectors. The fifth-largest firm was 334 times larger than the median firm. In other sectors, it was only 35 to 115 times larger. The manufacturing sector also has the highest Gini coefficient and the highest general entropy score for the number of employees. While already some large firms employing hundreds of people existed, others were still in the transition to the factory system.[27]

[27]For the years 1851 and 1861, these census tables include occupation descriptions for tens of thousands of firm owners. We combine text recognition algorithms with an updated occupation dictionary to assign one of over 1,500 occupation titles in our dictionary to each occupation description. Then, we assign the occupation titles to the manufacturing, trade, services, or agricultural sector based on the "History of Work (HISCO)" classification. See https://historyofwork.iisg.amsterdam/index.php for more information. We were able to assign over 98% of firms in the business census to one of the four sectors that way.

Table B.4: Heterogeneity in Firms' Size by Sector – 1851

Measure	Manufacturing	Agriculture	Trade	Services
S.d	149.3	10.5	38.6	27.8
5^{th} largest/Median	334.3	115.7	150.0	35
Gini	0.77	0.55	0.66	0.70
GE(1)	2.05	0.64	1.32	1.32

Notes: This table displays different measures of heterogeneity based on data on firm owners from the British 1851 census. All heterogeneity measures were calculated based on firms' number of employees, separately for the four main sectors manufacturing, agriculture, trade, and services. The heterogeneity measures are 1) the standard deviation across employee numbers, 2) the share of employees in the 5th largest to the median firm, 3) the Gini coefficient and 4) the general entropy score across employee numbers.

Additional specifications. Further results from alternative specifications illustrate the mechanisms behind our main results. Table B.5 shows the importance of competition. Bankruptcies increase significantly more in locations that the railway connects to large firms in an intermediate distance. Based on the British business census for 1851, we identify large firms as those that belong to the top decile in terms of employment. Our first variable identifies the employment in these large firms located in each cell and sector. Next, we construct different "market access" measures by counting large-firm employment in 1851 for each sector in grid cells connected to the railway within 100km, over 100km but within 250km, over 250km but within 500km, and over 500km away. Table B.5 then estimates the effect of a connection to the rail depending on the exposure to large firms in the same sector that are either located in the same cell or located farther away but connected to the railway network. We should note that the coefficients are to be interpreted as the effect of large firms within each buffer compared to large firms outside these buffers and in the same sector.

Column B.5.1 tests the argument using the employment of the largest firms in the same cell in 1851. The interaction with the rail variable bears a negative sign and implies that when cells where many workers are already employed in large firms get connected to the rail, bankruptcies were less likely to occur. In those markets, large firms might have benefited from the connection to the rail and, hence, did not suffer as much from competition. At the same time, local competition might already have driven less productive units to bankruptcy. Once the cells with large firms are connected to the rail, no small firms in their sector would potentially suffer from a connection to the rail.

Columns B.5.2 to B.5.5 test the effect of a connection to the rail interacted with the employment in large firms at different distance buffers. According to Column B.5.2, large firms within a 100km radius do not impact the number of bankruptcies in their sector when connected to the rail. Large firms within a 100km radius were already quite accessible using other modes of transportation. Yet, large-firm employment in cells between 100km and 500km from a treated cell is associated with a significant increase in bankruptcies in their sector in cells connected to the rail. The coefficients attached to employment in large firms are positive and significant at the one-percent level. In this radius, a one percent increase in the number of large firm employees increased bankruptcies by 0.07 to 0.08 percent. In Column B.5.5, we observe that the effect becomes negative. Competition from firms farther away has a weaker effect on bankruptcies than large firms' employment at short distances.

Table B.5: Rail, Existing Market Structure and Bankruptcies

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Employees in Large Firms $_{i,t,s}$	-0.37***				
	(0.02)				
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Large (Dist<100km) $_{i,t,s}$		0.01			
		(0.03)			
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Large (100<Dist< 250km) $_{i,t,s}$			0.06**		
			(0.02)		
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Large (250<Dist< 500km) $_{i,t,s}$				0.08***	
				(0.02)	
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Large (Dist>500km) $_{i,t,s}$					-0.05**
					(0.03)
Observations	7473	7473	7473	7473	7473
Pseu. R^2	.813	.813	.813	.813	.813
Geo FE	✓	✓	✓	✓	✓
Sector × Year FE	✓	✓	✓	✓	✓
Sector $\text{Employment}_{i,s,t}$	✓	✓	✓	✓	✓
Geo Controls_i × Manufacturing_s	✓	✓	✓	✓	✓
Banks_i × Manufacturing_s	✓	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t, interacted with the log-transformed number of employees in large firms in that sector located in different distance categories and connected to the rail. Geo controls are the straight-line distance to the nearest city with coal deposits, the nearest port, and the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Table B.6 investigates the role of credit supply. To do so, the estimations include a triple interaction between the railway × manufacturing interaction term and the historical availability of country banks collected by Heblich and Trew (2018). This triple interaction is significantly negative in the most restrictive specification. Country banks cushioned the effect of the rail on bankruptcies in the manufacturing sector, probably by providing loans to smooth the financial shocks from the industrial transformation. This interpretation of the results also supports the overall interpretation of our estimates that a connection to the rail deteriorated the financial situation of some in the manufacturing sector. However, when they had access to credit, they could cope better with the shock.

Table B.6: The effect of the rail on bankruptcies – The role of country Banks

	(1)	(2)	(3)	(4)	(5)
	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$				
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	2.12***	0.13	0.16	0.13	0.18
	(0.25)	(0.17)	(0.17)	(0.17)	(0.18)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s	0.47***	0.47***	0.48***	0.55***	0.40***
	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ CountyBanks_i	0.30	-0.16	-0.17	-0.18	-0.12
	(0.23)	(0.15)	(0.15)	(0.15)	(0.15)
Manufacturing_s $\times$ CountyBanks_i	-0.02	-0.02	-0.01	0.00	0.22**
	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s $\times$ CountyBanks_i	0.05	0.05	0.01	0.01	-0.21**
	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)
Observations	8277	7473	7473	7473	7473
Pseu. R^2	.127	.8	.801	.801	.813
Geo FE		✓	✓	✓	✓
Sector FE		✓	✓	✓	
Year FE		✓	✓	✓	
Sector-Year FE					✓
$\text{Sector Employment}_{i,s,t}$			✓	✓	✓
Geo controls_i $\times$ Manufacturing_s				✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t, interacted with an indicator variable for the observed sector. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $CountyBank_i$ are the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

C Appendix C – Robustness checks (Main estimates)

This Appendix Section provides a number of robustness tests to our main specifications. These robustness tests constitute different changes in control variables, standard errors, or the sample composition.

Table C.1 presents two important variations of our main specifications. In Columns C.1.1 and C.1.2, we compute stricter standard errors by accounting for two-way and three-way clustering. These alternative ways of clustering standard errors do not impact the significance of the estimates. Columns C.1.3 and C.1.4 address the Modifiable Areal Unit Problem (MAUP). The results are stable across bigger and smaller grid cell sizes, respectively. Doubling the grid cell size to over 400 square kilometers leaves the results almost unchanged. The results remain significant when halving the cell size to roughly 100 square kilometers, but the coefficients become smaller. Arguably, too small grid cells make it more difficult to find an effect as it is difficult to pin down the economic effect of new railway stations in too small spatial units.

Table C.1: Robustness Tests – Clustering and cells construction

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$			
	(1)	(2)	(3)	(4)
	2 w cluster	3w Cluster	Big Cells	Small Cells
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	0.12	0.12	0.22	0.15
	(0.13)	(0.12)	(0.15)	(0.13)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s	0.34***	0.34**	0.26**	0.15*
	(0.04)	(0.14)	(0.12)	(0.08)
Observations	7473	7473	4235	11584
Pseu. R^2	.813	.813	.821	.807
Geo FE	✓	✓	✓	✓
Sector-Year FE	✓	✓	✓	✓
$\text{Employment}_{i,t}$	✓	✓	✓	✓
Geo Controls_i × Manufacturing_s	✓	✓	✓	✓
Banks_i × Manufacturing_s	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t, interacted with in indicator variable for the manufacturing sector. Geo controls are the straight-line distance to the nearest city with coal deposits, the nearest port, and the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. In Columns (3) and (4), we double (half) the average area of the grid cells on which we base our sample. Column (1) uses two-way clustered standard errors, at the grid cell and at the sector level. Column (2) uses three-way clustered Standard Errors, in all other columns Standard Errors (in parentheses) are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

The main specifications already control for various potential confounders. In Table C.2, we introduce unemployment, migration, and male population shares as additional control variables. While these variables do not vary at the sector level, they still are important correlates of local manufacturing intensity. All specifications in Table C.2 report almost identical coefficients on the main treatment variable. Hence, these additional controls do not affect the explanatory power of the rail × manufacturing indicator, suggesting that the controls and fixed effects from the main specifications already absorb most of the variation explained by unobserved local characteristics.

Table C.2: Local Shocks

	Dependent Variable: #Bankruptcies$_{i,t,s}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}$ (Rail$_{i,t}$ > 0)	0.09	0.12	0.05	0.02
	(0.15)	(0.14)	(0.14)	(0.15)
$\mathbb{1}$ (Rail$_{i,t}$ > 0) × Manufacturing$_s$	0.33***	0.34***	0.34***	0.33***
	(0.08)	(0.08)	(0.09)	(0.09)
Unemployment$_{i,t}$	0.52			0.63*
	(0.34)			(0.38)
Migrants$_{i,t}$		-0.16		-0.22
		(0.78)		(0.61)
Male pop$_{i,t}$			-0.40***	-0.41***
			(0.08)	(0.08)
Observations	7473	7473	7473	7473
Pseu. R^2	.813	.813	.814	.814
Geo FE	✓	✓	✓	✓
Sector × Year FE	✓	✓	✓	✓
Sector Employment$_{i,s,t}$	✓	✓	✓	✓
Geo Controls$_i$ × Manufacturing$_s$	✓	✓	✓	✓
Banks$_i$ × Manufacturing$_s$	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Additional control variables are the unemployment rate, the number of people born in another county, and the share of the male population in location i in census-year t. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

Table C.3 considers different samples, dropping grid cells with the 5% highest and 5% lowest population (Columns C.3.1-C.3.3), as well as cells that include railway nodes, i.e. towns with over 5,000 population in the year 1801 (Column C.3.4). Even dropping all these cells from the sample (Column C.3.5) does not change the coefficient on the interaction term of interest.

Table C.3: Excluding Cells with low/high Levels of Population

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$				
	(1)	(2)	(3)	(4)	(5)
	w.o Top5 %	w.o Bottom5 %	w.o Both5 %	w.o Nodes	w.o Previous
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	-0.09	0.12	-0.09	-0.00	-0.14
	(0.15)	(0.14)	(0.15)	(0.18)	(0.17)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s	0.35^{***}	0.34^{***}	0.35^{***}	0.40^{***}	0.40^{***}
	(0.09)	(0.08)	(0.09)	(0.09)	(0.10)
Observations	7229	7456	7212	6973	6820
Pseu. R^2	.737	.813	.738	.746	.715
Geo FE	✓	✓	✓	✓	✓
Sector × Year FE	✓	✓	✓	✓	✓
$\text{Sector Employment}_{i,t}$	✓	✓	✓	✓	✓
Geo controls_i × Manufacturing_s	✓	✓	✓	✓	✓
Banks_i × Manufacturing_s	✓	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Columns (1)-(5) employ different sample restrictions; Column (1) drops the grid cells with the 5% highest population, Column (2) the grid cells with the 5% lowest population, and Column (3) both types of extreme cells. Column (4) excludes grid cells that contain a railway node, i.e. one of the 100 most populous towns in 1851. Column (5) applies all three sample restrictions at the same time. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

Table C.4 excludes bankruptcies that occurred in years affected by reforms to bankruptcy law. Until the reform in 1861, only people declaring a merchant occupation could file for bankruptcy. Two further reforms in 1869 and 1883 ended and reintroduced the "officialism" system, where public courts preside of bankruptcy proceedings. Table C.4 presents estimates on a sample of observations after 1861 (when everybody could file for bankruptcy) and during officialism (excluding 1869 to 1883). Still, the coefficients on the main interaction remain statistically significant and similar in size to the baseline results in Table 1.

Table C.4: Excluding Reform Years

	(1)	(2)	(3)	(4)	(5)	(6)
		Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$				
Dependent variable excludes bankruptcies in 1869 - 1883 and pre-1861						
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	2.78***	2.61***	0.15	0.21	0.22	0.23
	(0.25)	(0.27)	(0.35)	(0.35)	(0.35)	(0.35)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s		0.52***	0.52***	0.47***	0.44***	0.38***
		(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Observations	5465	5465	4425	4425	4417	4417
Pseu. R^2	.0466	.0602	.687	.689	.69	.691
Geo FE			✓	✓	✓	✓
Sector FE			✓	✓	✓	
Year FE			✓	✓	✓	
Sector-Year FE						✓
Sector $\text{Employment}_{i,s,t}$				✓	✓	✓
Geo Controls_i $\times$ Manufacturing_s					✓	✓
Banks_i $\times$ Manufacturing_s					✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. The dependent variable excludes bankruptcies that occurred between 1869 and 1883, when a temporary policy change allowed settling bankruptcies outside courts as well as bankruptcies that occurred before 1861. Before a reform in 1861, only merchants were allowed to go bankrupt. In effect, the sample becomes restricted to the census years 1861 and 1881. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

The next two tables address concerns of spatial spillovers and spatial autocorrelation. Specifications in Columns C.5.1 and C.5.3 of Table C.5 repeat the main specifications but cluster standard errors at the county level to account for common unexplained shocks at the county level. This higher-level clustering does not noticeably affect the standard errors. The specifications in Columns C.5.2 and C.5.4 introduce spatial lags of the outcome variables. This is, the regressions include the number of bankruptcies and employees in the same sector, but observed in neighboring cells. These spatial lags do not significantly affect the coefficient for the number of bankruptcies, but lead to a smaller coefficient for employment.

Table C.5: Considering Spatial Autocorrelation

Dep Variable:	#Bankruptcies$_{i,t,s}$		#Employment$_{i,t,s}$	
	(1)	(2)	(3)	(4)
	County Cluster	Spatial Lag	County Cluster	Spatial Lag
$\mathbb{1}$	0.12	0.13	-0.10***	0.01
(Rail$_{i,t}$ > 0)	(0.14)	(0.14)	(0.03)	(0.03)
$\mathbb{1}$	0.34***	0.33***	0.26***	0.18***
(Rail$_{i,t}$ > 0) × Manufacturing$_s$	(0.08)	(0.08)	(0.08)	(0.07)
Bankruptcies		0.12*		
Neigh$_{i,s,t}$		(0.07)		
Employment				0.60***
Neigh$_{i,s,t}$				(0.05)
Observations	7473	7473	8632	8277
Pseu. R^2	.813	.813	.929	.949
Geo FE	✓	✓	✓	✓
Sector × Year FE	✓	✓	✓	✓
Sector Employment$_{i,t,s}$	✓	✓		
Population$_{i,t}$			✓	✓
Geo Controls$_i$ × Manufacturing$_s$	✓	✓	✓	✓
Banks$_i$ × Manufacturing$_s$	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variables are the annualized number of bankruptcies that occurred between two census years (Columns (1) and (2)), and the number of employed people in a census year (Columns (3) and (4)). The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. We additionally include spatial lags, i.e. the number of bankruptcies in neighboring cells (Column (2)) and the number of people employed in a neighboring cell (Column (4)). Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

Tables C.6 and C.7 repeat the main specifications for bankruptcies and employment as the dependent variables, respectively. The specifications allow for spatially clustered shocks among cells within 300km distance. Accounting for spatial autocorrelation does not affect the results for bankruptcies. For employment as the dependent variable, standard errors increase markedly. This suggests that employment-reactions to the railway expansion are more correlated in space than bankruptcy-reactions.

Table C.6: The Effect of the Rail on Bankruptcies (Conley Standard Errors)

		Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$				
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	2.62***	2.46***	0.05***	0.08***	0.05	0.12***
	(0.22)	(0.21)	(0.02)	(0.02)	(0.04)	(0.05)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s		0.50***	0.50***	0.48***	0.55***	0.33***
		(0.04)	(0.04)	(0.05)	(0.06)	(0.07)
Observations	8341	8341	7481	7481	7473	7473
Pseu. R^2	0.076	0.092	0.767	0.768	0.780	0.780
Geo FE			✓	✓	✓	✓
Sector FE			✓	✓	✓	
Year FE			✓	✓	✓	
Sector-Year FE						✓
$\text{Sector Employment}_{i,s,t}$				✓	✓	✓
Geo Controls_i × Manufacturing_s					✓	✓
Banks_i × Manufacturing_s					✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Conley Standard Errors in parentheses (300km threshold), * p< 0.1, ** p< 0.05, *** p< 0.01.

Table C.7: The Effect of the Rail on Employment (Conley Standard Errors)

		Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$				
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	1.18***	0.91***	-0.42***	-0.27***	-0.11**	-0.10***
	(0.10)	(0.05)	(0.02)	(0.06)	(0.04)	(0.03)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s		0.60***	0.60***	0.60***	0.27	0.26**
		(0.15)	(0.14)	(0.14)	(0.33)	(0.10)
Observations	8704	8704	8704	8704	8632	8632
Pseu. R^2	0.074	0.215	0.873	0.906	0.923	0.929
Geo FE			✓	✓	✓	✓
Sector FE			✓	✓	✓	
Year FE			✓	✓	✓	
Sector-Year FE						✓
$\text{Population}_{i,t}$				✓	✓	✓
Geo Controls_i × Manufacturing_s					✓	✓
Banks_i × Manufacturing_s					✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the number of people employed in sector s at census year t and in grid cell i. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Conley Standard Errors in parentheses (300km threshold), * p< 0.1, ** p< 0.05, *** p< 0.01.

The next three tables provide small variations of the main specifications. Regressions displayed in Table C.8 weight observations by the inverse of a cell's population. Tables C.9 and C.10 re-define the dependent variables by using the share of bankruptcies among the employed population and the share of employed people among the local population instead of count variables as the dependent variables. The coefficient estimates remain statistically significant throughout.

Table C.8: The Effect of the Rail on Bankruptcies – Inverse Probability Weighting

	(1)	(2)	(3)	(4)	(5)	(6)
		Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$				
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	2.09***	1.93***	-0.27*	-0.27*	-0.24	-0.24
	(0.24)	(0.24)	(0.15)	(0.15)	(0.16)	(0.15)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s		0.47***	0.42***	0.44***	0.35**	0.31***
		(0.10)	(0.11)	(0.10)	(0.16)	(0.11)
Observations	8341	8341	7481	7481	7473	7473
Pseu. R^2	.0942	.103	.609	.611	.611	.622
Geo FE			✓	✓	✓	✓
Sector FE			✓	✓	✓	
Year FE			✓	✓	✓	
Sector-Year FE						✓
$\text{Sector Employment}_{i,s,t}$				✓	✓	✓
Geo Controls_i $\times$ Manufacturing_s					✓	✓
Banks_i $\times$ Manufacturing_s					✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Observations in the regressions are weighted by the inverse of cell i's population. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Table C.9: The Effect of the Rail on Employment Shares

Dependent Variable:	$\#\text{Employed}_{i,t,s}/\#\text{Population}_{i,t}$		
	(1)	(2)	(3)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	-0.14***	-0.08***	-0.08***
	(0.02)	(0.02)	(0.02)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s	0.41***	0.24***	0.26***
	(0.04)	(0.04)	(0.04)
Observations	8344	8280	8280
Pseu. R^2	.0682	.0731	.074
Geo FE	✓	✓	✓
Sector FE	✓	✓	
Year FE	✓	✓	
Sector × Year FE			✓
$\text{Population}_{i,t}$	✓	✓	✓
Geo Controls_i × Manufacturing_s		✓	✓
Banks_i × Manufacturing_s		✓	✓

Notes: Table reports results from OLS regressions at the grid cell-sector-census year level. The dependent variable is the share of people employed in sector s at census year t and in grid cell i. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight-line distance to the nearest city with coal deposits, to the nearest port, and the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Finally, Table C.11 considers shorter time windows to compute the bankruptcy variables. The main specifications in Table 1 use the annualized number of bankruptcies that occur between two census waves. For 1861, this amounts to the sum of bankruptcies between 1861 and the next census in 1881. In Table C.11, we reduce this window to only two years. This short window makes it less likely that people enter a location and sector after we observe them in one census year, declare bankruptcy, and exit the location/sector before the next census year, hence artificially increasing the number of bankruptcies vis-à-vis the observed employees. Using these short windows to compute the dependent variable leads to qualitatively similar results as using the longer windows.

Table C.10: The Effect of the Rail on Bankruptcy Shares

Dependent Variable:	$\text{Log}(\text{Bankrupt}_{i,t,s}/\text{Employed}_{i,t,s})$		
	(1)	(2)	(3)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	-0.45***	-0.45***	-0.40**
	(0.17)	(0.17)	(0.16)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s	0.43***	0.43**	0.19*
	(0.13)	(0.19)	(0.10)
Observations	7481	7473	7473
Pseu. R^2	.365	.365	.367
Geo FE	✓	✓	✓
Sector FE	✓	✓	
Year FE	✓	✓	
Sector $\times$ Year FE			✓
Sector $\text{Employment}_{i,s,t}$	✓	✓	✓
Geo Controls_i $\times$ Manufacturing_s		✓	✓
Banks_i $\times$ Manufacturing_s		✓	✓

Notes: Table reports results from OLS regressions at the grid cell-sector-census year level. The dependent variable is the log-share of bankrupts over employed in sector s at census year t and in grid cell i. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight-line distance to the nearest city with coal deposits, to the nearest port, and the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Table C.11: The Effect of the Rail on Bankruptcies – Short Windows

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent Variable: #Bankruptcies$_{i,t,s}$ for 2 years					
$\mathbb{1}$ (Rail$_{i,t}$ > 0)	3.50***	3.30***	0.29	0.35	0.34	0.43*
	(0.27)	(0.27)	(0.25)	(0.25)	(0.25)	(0.24)
$\mathbb{1}$ (Rail$_{i,t}$ > 0) × Manufacturing$_s$		0.65***	0.65***	0.60***	0.62***	0.26*
		(0.12)	(0.12)	(0.12)	(0.13)	(0.14)
Observations	8341	8341	6678	6678	6674	6674
Pseu. R^2	.082	.095	.839	.841	.841	.852
Geo FE			✓	✓	✓	✓
Sector FE			✓	✓	✓	
Year FE			✓	✓	✓	
Sector-Year FE						✓
Sector Employment$_{i,s,t}$				✓	✓	✓
Geo Controls$_i$ × Manufacturing$_s$					✓	✓
Banks$_i$ × Manufacturing$_s$					✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the number of bankruptcies that occurred within 2 years of a census-year. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

D Appendix D – Spillovers

This Appendix section presents alternative specifications considering potential spillover effects from the railway expansion. We compute a market access measure following Donaldson and Hornbeck (2016), and directly investigate spillovers from rail connection in neighboring locations.

Table D.1 repeats our main specifications with the number of bankruptcies as the dependent variable, but using market access as the treatment variable. We compute two versions of the market access variable proposed in Donaldson and Hornbeck (2016). Both versions deviate slightly from Donaldson and Hornbeck (2016) because we only have information on railway stations, but not on railway lines for each of the census years. Both measures use their elasticity of substitution of $\theta = 8.22$.

Our "simple" market access measure considers the straight-line distance between any two grid cells' centroids, and use per-kilometer transportation costs estimated in Donaldson and Hornbeck (2016) to measure bilateral transportation costs. If both grid cells have at least one railway station, we compute their transport costs as 0.49 cents per kilometer, adding a 50 cent fixed cost for transhipment. If at least one grid cell is not connected to the rail, we use 23 cents per kilometer.

The second market access measure takes the density of railway stations in a location into account. The median grid cell in our sample has four railway stations. We therefore call cells with at least four stations hubs, and assign them lower transportation costs as non-hubs. We then define per-kilometer costs the following way:

- Between two hubs: 0.49 cents/kilometer.
- Between one hub and one non-hub: 1.25 cents/kilometer.
- Between two non-hubs: 1.50 cents/kilometer.

Both market access measures follow the equation of Donaldson and Hornbeck (2016):

$$MA_{i,t} = \sum_{d} \tau_{id,t}^{-\theta} N_{d,t} \tag{2}$$

Market access for location i in census year t is hence the sum of the population in all other locations, $N_{d,t}$, weighted by the inverse of the bilateral transport costs $\tau_{id,t}$.

The results using this market access measure confirm our main results. Columns D.1.1 and D.1.2 show that locations with higher market access see a significant increase in bankruptcies in the manufacturing sector. Following the intuition in Borusyak and Hull (2023), we repeat these regressions for a sub-sample of locations that received a railway station within two years before and after a census year, as in Table 4. Even in this smaller sample, we find a significant effect of market access on manufacturing bankruptcies, despite a severe loss of power due to the much smaller sample size.

Table D.1: Market Access as a measure of connections

	Dep Variable: #Bankruptcies$_{i,t,s}$			
	(1)	(2)	(3)	(4)
Sample	All	All	2y	2y
Market Access Simple$_{i,t}$ $\times$ Manufacturing$_s$	0.02***		0.02*	
	(0.01)		(0.01)	
Market Access$_{i,t}$ $\times$ Manufacturing$_s$		0.02***		0.01*
		(0.01)		(0.01)
Observations	9283	9283	685	685
Pseu. R^2	.817	.817	.839	.839
Geo FE	✓	✓	✓	✓
Sector-Year FE	✓	✓	✓	✓
Sector Employment$_{i,s,t}$	✓	✓	✓	✓
Geo Controls$_i$ $\times$ Manufacturing$_s$	✓	✓	✓	✓
Banks$_i$ $\times$ Manufacturing$_s$	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variables are two alternatives of market access measures for cell i census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * $p< 0.1$, ** $p< 0.05$, *** $p< 0.01$

Table D.2 takes another approach by directly controlling for spillovers from the railway expansion. It adds an indicator variable that takes the value of 1 if a cell's neighboring cell's are connected to the railway network. This spatial lag of the treatment variable is insignificant and does not impact the coefficient of our actual treatment variable. The last column determines whether these neighbouring cells experienced more bankruptcies than other cells far from the rail by dropping treated cells from the sample. In this subsample, we do not find that spillovers play a role.

Table D.2: Investigating the importance of spillovers

	Dep Variable: #Bankruptcies$_{i,t,s}$	
	(1)	(2)
$\mathbb{1}$ (Rail$_{i,t} > 0$) $\times$ Manufacturing$_s$	0.44**	
	(0.20)	
$\mathbb{1}$ (NeighbourRail$_{i,t} > 0$) $\times$ Manufacturing$_s$	0.10	0.21
	(0.20)	(0.16)
Observations	7473	1544
Pseu. R^2	.813	.412
Geo FE	✓	✓
Sector-Year FE	✓	✓
Sector Employment$_{i,s,t}$	✓	✓
Geo Controls$_i$ $\times$ Manufacturing$_s$	✓	✓
Banks$_i$ $\times$ Manufacturing$_s$	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. We additionally introduce a spatial lag variable that takes the value of one if at least one of cell i's neighbors has a railway connection, and also interact this spatial lag with an indicator variable for the manufacturing sector. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

E Appendix E – Selection into treatment and causality

This Appendix Section presents results from alternative identification strategies. Section 6 explains the logic of a Least Cost Path (LCP) instrument used to address the endogenous selection of locations into railway access. We give some more background on this instrument below.

In addition, we present additional specifications that take alternative approaches to account for endogenous selection into railway access. First, we control for a counterfactual network. Then, we leverage information on the timing of station openings to control for locational selection directly.

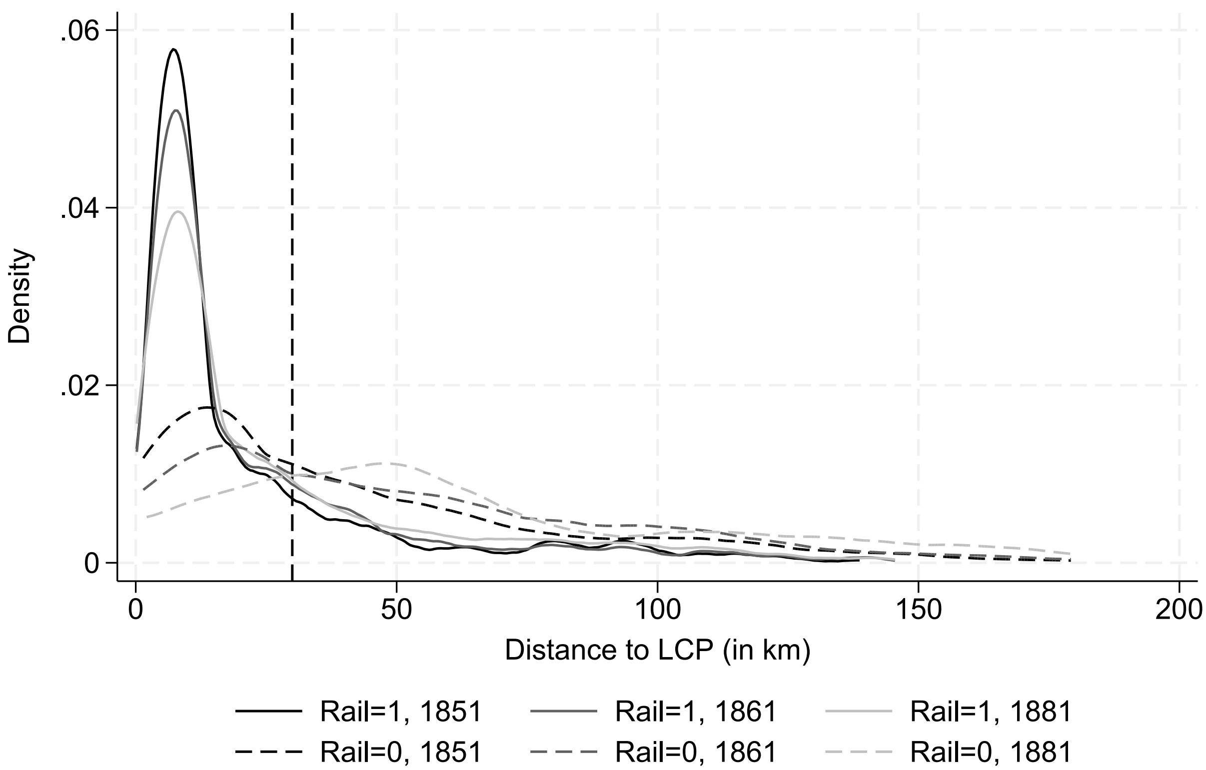
Finally, we adopt a design-based approach for our estimations. The final part of this Appendix Section investigates the plausibility of conditionally random exposure of the manufacturing sector to the railway expansion.

E.1 Least Cost Path

Our Least Cost Path (LCP) approach closely follows the design in Bogart et al. (2022). We designate all towns above 5,000 population in 1801 as natural railway nodes. We then compute the "gravitational force" between each town pair, based on towns' bilateral distance and each town's population. We then construct least cost paths based on terrain maps between the 100 pairs with the highest "gravitational force."

Other than Bogart et al. (2022), our estimations use a panel dataset that traces the railway expansion until 1891. We therefore add a buffer of 30km around the LCP and code cells as "treated" by the LCP if their centroid lies inside this buffer. As Figure E.1 shows, we observe a significant discontinuity of cells receiving railway access within 30km to the LCP. This 30km cut-off hence likely accounts for the branching-out of the railway network around the main lines constructed between the railway nodes.

Figure E.1: Kernel Density - Grid cells, Access to the Rail, and Distance to LCP



Notes: This figure displays Kernel density functions for grid cells connected to the rail in different years (plain lines) and grid cells not connected to the rail (dashed lines) depending on the distance to the LCP. Different years are represented by various shades of grey, the darker means more ancient sample.

Table E.1 shows that the LCP instrument does not predict the density of the railway network in a cell. In Columns E.1.1 to E.1.3, we regress the number of railway stations on an indicator for one station being present together with our LCP instrument. The correlation between the LCP and the number of stations is a relatively precise zero. This does not change when we control for spatial lags of employment or bankruptcies.

Table E.1: Placebo Test – LCP Proximity and Number of Stations

Instrument	LCP<30km	LCP<25km	LCP<35km	LCP<30km	LCP<30km	LCP<30km
	(1)	(2)	(3)	(4)	(5)	(6)
	First stage / Dependent Variable: Nb Rail $_{i,t}$					
$\mathbb{1}$ Rail$_{i,t}$	1.26***	1.26***	1.26***	1.26***	1.25***	1.25***
	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$\mathbb{1}$ Instrument$_i$	-0.01	-0.00	-0.06	-0.02	-0.03	-0.04
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Spillovers Bankrupt $_{s,i,t}$				0.08***		0.07***
				(0.02)		(0.02)
Spillovers Employ $_{s,i,t}$					0.06***	0.05***
					(0.01)	(0.01)
Observations	7473	7473	7473	7473	7473	7473
Adj. R^2	.728	.728	.729	.734	.733	.737
Geo FE	✓	✓	✓	✓	✓	✓
Year × Sector FE	✓	✓	✓	✓	✓	✓
Sector Employment$_{i,s,t}$	✓	✓	✓	✓	✓	✓
Geo Controls$_i$ × Manufacturing$_s$	✓	✓	✓	✓	✓	✓
Banks$_i$ × Manufacturing$_s$	✓	✓	✓	✓	✓	✓

Notes: Table reports results from OLS regressions at the grid cell-sector-census year level. The dependent variable is the number of railway stations in a location i in census year t. The main explanatory variables are an indicator variable for at least one railway station being present, and an indicator for a location i being within 30km to the Least Cost Path instrument. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

E.2 Controlling for Selection into Treatment

This section presents results directly controlling for selection into railway connection in different ways.

Estimates in Table E.2 control for a counterfactual rail network developed by Casson (2009). This counterfactual rail network is based on the (cost)-efficiency of the rail. It captures the part of the rail network built for economic (and hence endogenous) reasons. Controlling for this counterfactual network does not affect the significance of our estimates. This signals that locations having certain locational advantages that make them relevant for a railway connection does not explain our main results.

Table E.2: The Effect of the Rail on Bankruptcies – Controlling for a Counterfactual Network

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$					
	(1)	(2)	(3)	(4)	(5)	(6)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	2.05***	1.96***	0.12	0.15	0.13	0.10
	(0.22)	(0.23)	(0.15)	(0.15)	(0.15)	(0.15)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s		0.27***	0.26***	0.24***	0.29***	0.31***
		(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Observations	8341	8341	7481	7481	7473	7473
Pseu. R^2	.159	.159	.804	.806	.806	.813
Geo FE			✓	✓	✓	✓
Sector FE			✓	✓	✓	
Year FE			✓	✓	✓	
Sector-Year FE						✓
$\text{Sector Employment}_{i,s,t}$				✓	✓	✓
Geo Controls_i $\times$ Manufacturing_s					✓	✓
Banks_i $\times$ Manufacturing_s					✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. All regressions additionally include an indicator variable for location i hosting at least one station of the counterfactual, "cost-efficient" railway network. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Table E.3 directly controls for the selection into a railway connection. Our strategy is similar in spirit to Costas-Fernández et al. (2020). We consider the geographic selection of rail stations in two different ways. First, we code the indicator variable $\mathbb{1}$ ($\text{Ever Rail}_{i,t} > 0$) $\times$ Manufacturing_s to identify cells that receive a rail station at any point in time. This variable hence factors out locations with a general "potential" to receive a railway access. Second, the variable $\mathbb{1}$ ($\text{Not Yet Rail}_{i,t} > 0$) $\times$ Manufacturing_s takes the value of one for cells that do not yet have a railway station in a census year, but will receive one in the next ten years. Even after controlling for the geographic selection of stations, our estimates remain of the same magnitude as the baseline results, suggesting that our effect is driven by the opening of stations and not by the extension of the rail network towards specific geographic areas prone to bankruptcies. After controlling for geographic selection, a rail connection still increases bankruptcies in the manufacturing sector by 27 to 36 percent.[28]

[28]According to the transformation $(e^{0.31} - 1) \cdot 100\% = 36\%$ for the coefficient in Column 4.

Table E.3: Exploiting the Timing of Connection – Long-Run

	Dependent Variable: #Bankruptcies$_{i,t,s}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}$ (Ever Rail$_{i,t}$ > 0) × Manufacturing$_s$	0.18***		0.16***	
	(0.04)		(0.04)	
$\mathbb{1}$ (Not Yet Rail$_{i,t}$ > 0) × Manufacturing$_s$		0.09**		0.08**
		(0.04)		(0.04)
$\mathbb{1}$ (Rail$_{i,t}$ > 0)			0.16	0.13
			(0.15)	(0.15)
$\mathbb{1}$ (Rail$_{i,t}$ > 0) × Manufacturing$_s$			0.23***	0.31***
			(0.08)	(0.08)
Observations	7473	7473	7473	7473
Pseu. R^2	.813	.813	.813	.813
Geo FE	✓	✓	✓	✓
Sector-Year FE	✓	✓	✓	✓
Sector Employment$_{i,s,t}$	✓	✓	✓	✓
Geo Controls$_i$ × Manufacturing$_s$	✓	✓	✓	✓
Banks$_i$ × Manufacturing$_s$	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variable is the annualized number of bankruptcies that occurred between two census years. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. In addition, regressions control for an indicator whether a location receives a station at any point until 1950 (Columns 1 and 3), and an indicator whether a location receives a station until the next census year (Columns 2 and 4). Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

E.3 Design-Based Interpretation

This subsection presents results to back-up the design-based approach to causal identification, presented in Table 4. This design-based approach leverages information on the opening year of railway stations. The sample is restricted to locations that received their first railway station around the same time. Yet, some locations received their station shortly before the year a census was taken (and on which we build our dataset), while others received their station only slightly later.

Table E.4 tests whether this restriction generates a balanced sample for a design-based analysis. Columns E.4.1 and E.4.2, investigate whether this reduced sample differs from the main sample in observable pre-period characteristics. We define two variables to investigate pre-trend differences across samples. Column E.4.1 uses the number of bankruptcies that occurred between the last census and before-last census years (e.g. between 1841 and 1851 when looking at the railway network in 1861). Column E.4.2 looks at the stock of bankruptcies that occurred up to a census period. Regressing these two dependent variables on an indicator for entering the reduced sample of locations that received their first station shortly before/after a census year, we find no significant differences across samples. Hence, the narrow sample does not constitute a special version of our overall sample in terms of pre-period bankruptcy trends.

In Columns E.4.3 and E.4.4, we use the same two dependent variables to investigate whether within the narrow sample, locations that received their station before the census year (the main treated cells) show different bankruptcy trends from those locations that are treated slightly later than a census year. Also here, both coefficients are far from statistical significance. This strongly backs the assumption of a quasi-random allocation into treated and slightly-later-treated locations.

Table E.4: Pretrends for design-based sample

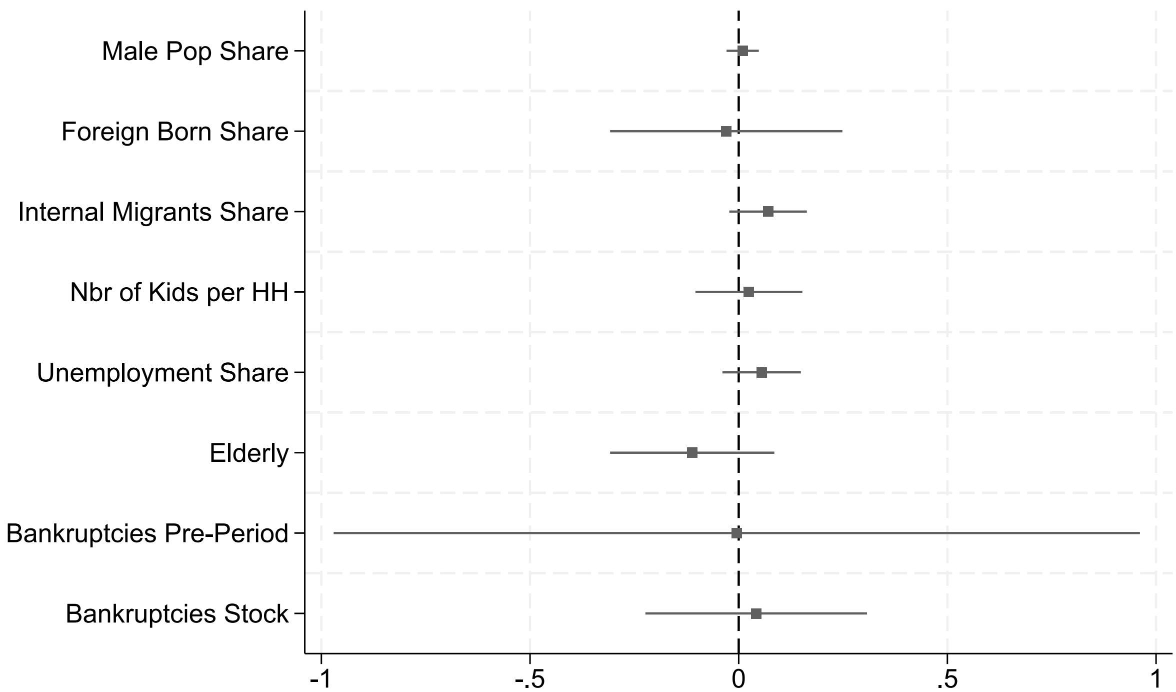
Dep. Variable is Bankruptcies as:	Flow Last Pre-Period	Stock At t	Flow Pre-Period	Stock At t
	(1)	(2)	(3)	(4)
$\mathbb{1}$ (Into Sample$_{i,t} > 0$) $\times$ Manufacturing$_s$	0.07	0.10		
	(0.11)	(0.06)		
$\mathbb{1}$ (Rail$_{i,t} > 0$) $\times$ Manufacturing$_s$			-0.21	0.09
			(0.39)	(0.18)
Observations	6108	6624	410	584
Pseu. R^2	.923	.975	.967	.983
Geo FE	✓	✓	✓	✓
Sector-Year FE	✓	✓	✓	✓
Sector Employment$_{i,s,t}$	✓	✓	✓	✓
Geo Controls$_i$ $\times$ Manufacturing$_s$	✓	✓	✓	✓
Banks$_i$ $\times$ Manufacturing$_s$	✓	✓	✓	✓

Notes: Table reports results from PPML regressions at the grid cell-sector-census year level. The dependent variables are the annualized number of bankruptcies that occurred between the last census year and the before-last census year (Columns 1 and 3), and the sum of bankruptcies that occurred up to a census year (Columns 2 and 4). The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. Geo controls are the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. $Banks_i$ refers to the number of country banks in a grid cell that were active in 1810. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Figure E.2 provides some suggestive evidence that a conditional independence assumption might also be plausible in the full sample. Our set-up differs from most other similar studies because we trace the railway expansion in a panel setting, controlling for location fixed effects. If the selection into railway connections occurred mostly due to time-invariant location characteristics, these should be controlled for in our setting.

We test this assumption by regressing a number of census-level variables that should not be related to our outcomes on our railway treatment indicator. As Figure E.2 shows, none of these variables differs across locations with and without a railway station, conditional on location and census-year fixed effects. In addition, we employ the same two pre-trend variables for bankruptcies as in Table E.4 above. Also for the full sample, we do not find a significant difference in bankruptcy-pretrends across connected and un-connected locations.

Figure E.2: Balance



Notes: Figure provides coefficients and 95% confidence intervals from balance checks. Each coefficient stems from a separate regression. The dependent variables are displayed as row names, the treatment variable is an indicator for a location having a connection to the railway. All regressions include location and year fixed effects as well as the control variables from the main specifications in Table 1.