The Bankruptcy Express
Market Integration and Intra-sector Reallocation in Industrializing Britain *

Tobias Korn†, Jean Lacroix‡

April 9, 2024

Very preliminary draft. Do not cite or circulate.

Abstract

We investigate two determinants of firm exit: technological change and market integration. As a complement to previous approaches, we put an emphasis on the interaction of these two factors. If within a sector, technology is not uniformly distributed then firms are heterogeneous. Market integration then fragilizes the least productive of these heterogeneous firms. Accordingly, the effect of technology on firm exits is conditional on market integration and vice-versa. To test this hypothesis, we introduce a new dataset on individual bankruptcies at the location-sector-year level in late 19$^{\text{th}}$ century Britain, which we combine with rich micro-level census data. In this period, we investigate the effect of the British railway expansion on firm exits and employment changes. We find that the manufacturing sector – the one with the most heterogeneous firms – experienced an increase in job creation and in firms' exits following the arrival of the rail. Accordingly, technological change and market integration work together to explain firms' failure and within-sector reallocation.

Keywords: Bankruptcies, Economic Growth, Structural transformation
JEL Codes: N63, L16, O33, R40, K35

*This work has been supported as part of France 2030 programme "ANR-11-IDEX-0003".

† 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.

1 Introduction

With their 2030 Agenda for Sustainable Development, UN member states pledged that "no one will be left behind".[1] As policymakers want to find the path to inclusive growth, we need to understand who loses when the majority prospers. Recently, the literature has proposed two main mechanisms explaining why economic growth may generate losers. First, reductions in trade costs increase profits for exporters and those depending on upstream imports but can hurt firms that compete with foreign suppliers for domestic markets (Autor et al., 2016; Melitz, 2003). Second, innovation promises gains for capital owners and big firms but can push less productive firms as well as workers with ancient skill sets out of the market (Acemoglu and Restrepo, 2022; Juhász et al., 2020).

Previous studies analyzed the effects of either of these factors separately. In this paper, we elicit an implicit assumption driving these results. Hypothetically, if markets are not integrated, the adoption of new technologies by some firms should not endanger the non-adopters. Similarly, if trade is open but firms are homogeneous, no intra-sector reallocation should occur. Hence, the effect of innovation on firm exits is conditional on market integration. Likewise, the effect of market integration on firm exits hinges on some existing productivity differentials across firms. This paper goes back to the original argument of Melitz (2003). Trade fosters intra-industry reallocation if firms are heterogeneous. This model is consistent with aggregate welfare gains at the industry level despite individual losses for some firms. If firms have different levels of productivity and entering the export market is costly, the less productive firms exit the market whereas the more productive firms gain from trade. This argument has been tested empirically in the trade literature (Autor et al., 2016).

Our study considers these theoretical arguments in a new setting: Britain during the Second Industrial Revolution. As emphasized in Juhász et al. (2020) for the case of the First Industrial Revolution in France, we make the assumption that during the Second Industrial Revolution in Britain some sectors reorganized more than others. The manufacturing sector's heterogeneity sticks out vis-à-vis the agricultural, trade, and service sectors. Indeed throughout the 19$^{\text{th}}$ century, Britain has been exposed to a variety of technologies shaping productivity and the organization of firms. As emphasized in Juhász et al. (2020) in the case of France, the diffusion of technology was slow and required organizational changes. Hence, it created heterogeneity in technology adoption and production. This intuition is validated empirically: the business censuses in 1851 and 1861 document a significant mixture of small and very large firms in the manufacturing sector, while the other sectors exhibit

[1] See for more information https://sdgs.un.org/2030agenda, last visited Nov. 6th, 2023.

much more homogeneous firm size distributions. Around the same time, Britain experienced a large extension of its rail network (Bogart et al., 2022). The Second Industrial Revolution in Britain hence provides a unique case to test the interaction between exposure to technology and market integration. In this setting, we leverage sector-level differences in firms' heterogeneity together with geographic and time variation in the access to the rail to see how much the interaction of technology with market integration actually leads to intra-industry reallocation.

To measure intra-industry reallocation, we introduce a new measure of financial distress at the individual level: the universe of private bankruptcies in England and Wales. Combining Optical Character Recognition and text recognition algorithms, we collected and all public bankruptcy announcements from the period. Early on, the British bankruptcy law mandated that all insolvencies must be publicly announced in the *London Gazette* such that all creditors had the chance to make their claims heard. This practice continues until today, with new bankruptcy announcements being published on the London Gazette's homepage every week. We web-scraped all 41,000 issues of the gazette from 1788 until 1988 and identified over 420,000 individual bankruptcy cases. For each bankruptcy case, we geolocated the stated home address of the bankrupt and coded the bankrupt's occupation, assigning it to an economic sector. Our empirical analysis focuses on the time of the British railway expansion, 1851-1890. For this sample period, our dataset includes information on around 150,000 bankruptcy cases at the sector-geography-time level. Note that our data do not identify firm bankruptcies per se, but capture the financial distress of individuals: workers and capital owners. In addition, we use employment data from the British censuses in 1851, 1861, and 1881 at the same level of disaggregation to determine if these bankruptcies result from between or within sector reallocation.

Table 1 illustrates the underlying market mechanisms behind potential, jointly observed bankruptcy and employment outcomes. The upper-right and bottom-left cells illustrate observations in line with between-sector reallocation that we usually term structural transformation. While employment decreases in some market or sector (upper-right cell), labor moves into another market/sector, where employment hence increases (bottom-left cell). Similarly, we would expect that this market restructuring favors firms in the rising (bottom-left) market, while firms in the declining (upper-right) market face a higher risk of bankruptcy. Empirically, we would expect bankruptcies to follow the opposite pattern as employment shares: where employment decreases bankruptcies should become more likely and vice versa.

The upper-left cell of Table 1 describes the situation where one observes increasing bankruptcies but non-decreasing employment in a sector or market. This situation would describe an intra-sector reallocation. In such a case, the statistical pattern of bankruptcies does *not* follow the opposite direction as employment shares. This is, we would see an increase in employment amid a simultaneous increase in bankruptcies in the same sector or market. Such market dynamics would be explained by some (smaller) firms exiting the market and being re-employed at the surviving firms.

Table 1: Within versus Between sector reallocation

	Increasing Jobs	Decreasing Jobs
Increasing Bankruptcies	Within-Sector	Between Sectors
Decreasing Bankruptcies	Between Sectors	Within-Sector

We find that the British manufacturing sector exhibited this exact pattern during the Second Industrial Revolution. The manufacturing sector consisted of more heterogeneous firms than the other sectors during the expansion of the British rail network. Our results document that manufacturing employment increased by approximately 30 percent in places newly connected to the rail network. At the same time, however, the number of bankruptcies increased as well by 20 to 40 percentage points among employees in the manufacturing sector. We do not observe the same pattern for other sectors. These results are robust to different estimation methods and survive several robustness checks ensuring that none of our results are driven by specific places. To strengthen a causal interpretation of these results, we employ an inconsequential places approach, estimating a local average treatment effect for small places along the Least Cost Path between British major railway nodes.

Further extensions illustrate that our effect is dynamic. The first area connected to the rail experienced a larger increase in manufacturing employment than areas later connected to the rail. The last areas to be connected experienced as much as five times more bankruptcies in the manufacturing sector than places first connected to the rail. We further show that these dynamics are explained by market structure rather than creditors' incentives to recover some capital to reinvest in more dynamic places.

Our results contribute to three strands of the literature. First, they offer a reinterpretation of Melitz (2003)'s theory emphasizing how economic changes may trigger intra-sector reallocation. The effect of market integration on firm exits has been widely documented in the trade literature (Autor et al., 2016, 2020; Heblich et al., 2020). In this paper, we investigate an internal trade shock brought by the railway expansion. Beyond trade, some authors have emphasized the importance of technological change in driving wage inequality (Acemoglu and Restrepo, 2022). We consider these factors jointly and offer a new proxy for this within-sector reallocation: bankruptcies. In the words of Melitz (2003), bankruptcies capture “least productive firms exiting". For that aspect, this paper's first contribution relies on the development of this new measure of reallocation. With the number of personal bankruptcies, we directly identify the characteristics of the persons losing from this reallocation. 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 (Bernstein et al., 2019). This paper identifies the economic factors explaining these bankruptcies.

Second, this paper uses new data to identify the losers of technological change in general but also from the second industrial revolution in particular. Industrialization potentially reduces the demand for some skills and hence occupations (Goldin and Katz, 1998). Similarly, the gains of the Industrial Revolution were unevenly distributed (Crafts, 2022). Temin (1997) moreover shows that during the Industrial Revolution all sectors observed productivity gains but some sectors experienced more productivity gains than others. Juhász et al. (2020) present evidence of intra-sector reallocation in the case of cotton spinning in France. In their case, productivity was highly dispersed among firms and the less productive firms exited as mechanized cotton spinning developed. As Juhász et al. (2020) define intra-sectoral dynamics, our study adds a geographic dimension to such reallocation. Market integration increased intra-sectoral reallocation in a fast-evolving sector. This reallocation has political consequences (Caprettini and Voth, 2020). Yet, the mechanisms driving this destruction remain to be understood.

Third, our paper offers a new perspective on the geographic impact of railways: increasing productivity (Donaldson, 2018), increasing the diffusion of ideas (Tsiachtsiras, 2022), encouraging industrialization (Berger, 2019) and spurring economic growth (Donaldson and Hornbeck, 2016). The approach of this paper is similar to Bogart et al. (2022) and complements their estimate of the effect of the railways in 19$^{\text{th}}$ century England and Wales on urbanization and structural change. Our paper characterizes the nature of this structural change: it is biased against individuals in the sectors most impacted by market integration.

The next section sets the stage by discussion the bankruptcy law in 19th century Britain as well as the British railway expansion. Section 3 introduces our dataset and discusses our main Difference-in-Differences specifications before Section 4 discusses our main results. In Sections 5 and 6, we introduce our 2SLS specifications based on an inconsequential places approach along with additional robustness checks to our empirical analysis. Sector 7 then discusses various mechanisms that explain how the rail expansion led to the statistical pattern that we term intra-sector reallocation in the manufacturing sector before Section 8 concludes.

2 Historical background

2.1 Bankruptcy procedures in 19th century Britain

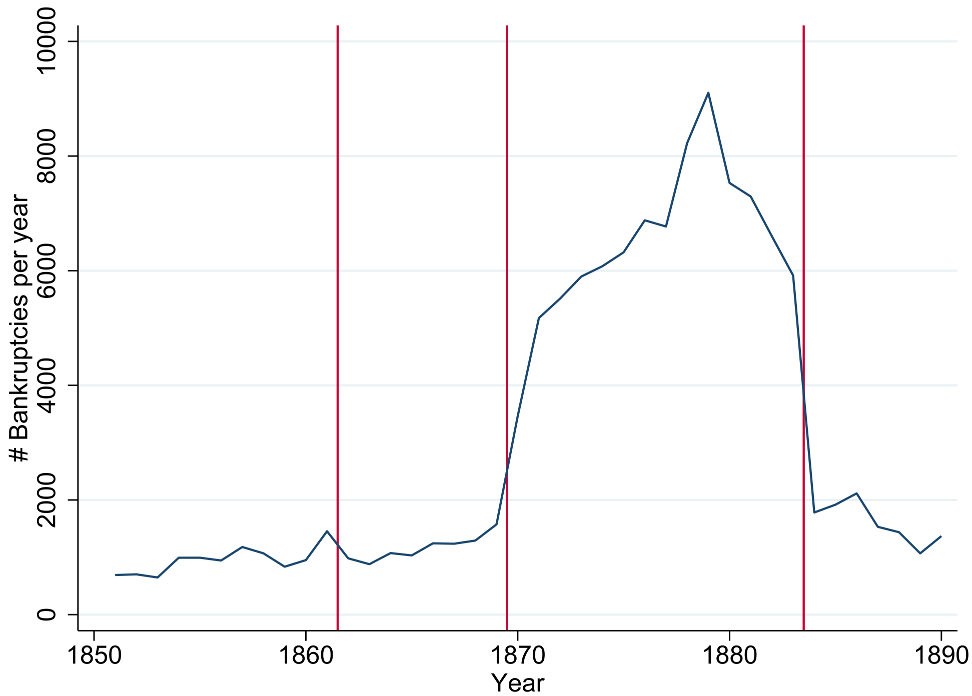
Bankruptcy procedures were at the forefront of political conversations throughout 19th century England (Lester, 1991). Debtors' prison illustrates well the consequences of bankruptcy, how complex the system was, and how important bankruptcies were in the collective image of 19th century England.[2] At the beginning of the 19th century, it was common for debtors that could not repay their debts to be sent to prison until their labour could repay their debt. Throughout the century, several reforms modernized both the procedure and the role of debtor's prison. From 1831, the procedure implied that officials would be appointed to collect and distribute the assets of bankrupts. Bankruptcy could then be initiated by both debtors and creditors. This doctrine of bankruptcy law called "officialism" was deemed inefficient by entrepreneurs and business elites. The system of officialism was costly and its ability to recover unpaid debt was slow and limited. The 1869 Bankruptcy and Debtor Acts massively changed this institution. After this series of reforms, debtors' prison was limited to debtors who were believed to have the financial means to repay their debt but did not do so. Moreover, the doctrine of officialism was repealed and a new system of bankruptcy management was put in place. In this case, if a majority of creditors agreed, they could proceed to the management of the bankruptcy themselves.

This new management of bankruptcies advantaged creditors. Recovery rates were higher as creditors had direct incentives to recover as much of the debt as possible. They also could avoid recovering small debts whose costs to recover were greater than the debt itself. Our dataset illustrates those changes. Figure 1 presents the evolution of the number of bankruptcies per year in the time frame of our study.

Three reforms occurred during the period of our study. The 1861 reform broadened the scope of the bankruptcy to all citizens and not only to those having a trading activity. The 1869 reform repealed officialism whereas the 1883 reform reintroduced it. Figure 1 evidences the importance of the bankruptcy regime to determine the number of bankruptcies. In Section 7.3, we leverage upon these differences in regimes to inform on the mechanisms potentially explaining more bankruptcies. Figure 1 shows the massive increase in the number of bankruptcies following the repeal of officialism. This shift shows how much creditors' incentives determine whether or not a bankruptcy takes place (through the official channel).

[2]Debtors are, for example, a common figure of Charles Dickens' work reflecting the author's father's own experience as an inmate in a debtor's prison.

Figure 1: The Evolution of Bankruptcies



Notes: This figure plots the aggregated number of bankruptcies per year based on new data collected by the authors. Red 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 the majority of creditors. In 1883, bankruptcy management returned to "officialism" where courts presided over bankruptcy cases).

2.2 The railway expansion

Between 1851 and 1881, 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 already expanded to Wales and the South-Western part of England. By the end of the 19^{th} century, the rail became the main mode of transportation for passengers and materials (Bogart et al., 2022).

The impact of the rail on the British has been at the center of academic debates for decades. Early scholars argued that the effect of the railway expansion on the economy was not clear and immediate in Britain, as opposed to other areas such as the U.S. (Mitchell, 1964). New Economic Geography models on the contrary emphasized the changes brought about by the rail (Lafourcade and Thisse, 2011). With decreasing transportation costs, the rail encouraged urbanization and structural change (Bogart et al., 2022). Similarly, railways fostered growth in Germany (Hornung, 2015) and in the US (Donaldson and Hornbeck, 2016). It also increased firms' productivity (Hornbeck and Rotemberg, 2019) and increased the diffusion of innovative ideas (Tsiachtsiras, 2022). Beyond these rather positive effects, a few studies show that the transformations generated by the rail also sometimes generated negative externalities and partially reduced life quality for some citizens (Waugh, 1956).

2.3 Sectors and firm heterogeneity

Our main hypothesis posits that the interaction of sector characteristics with market integration shapes market dynamics. When market integration comes together with a significant degree of market heterogeneity, we expect the occurrence of increased bankruptcies together with increased employment amid a within-market labor reallocation. Hence, market heterogeneity ultimately moderates the effect of market integration.

Table 2: Measures of heterogeneity across sectors – 1851

Measure	Manufacturing	Agriculture	Trade	Service
S.d	150.1	10.0	38.7	16.5
5^{th} largest/Median	334.3	115.7	150.0	35
Gini	0.77	0.55	0.66	0.66
GE(1)	2.05	0.62	1.31	1.03

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.

To substantiate this claim, we look into the business censuses of the 19^{th} century provided by the I-CeM project (Schurer and Higgs, 2023). For the years 1851 and 1861, these census tables include occupation descriptions for tens of thousands of firm owners. We combine text recognition algorithms in combination with an updated occupation dictionary (see below) 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, service, or agricultural sector based on the "History of Work (HISCO)" classification[3]. We were able to assign over 98% of firms in the business census to one of the four sectors that way.

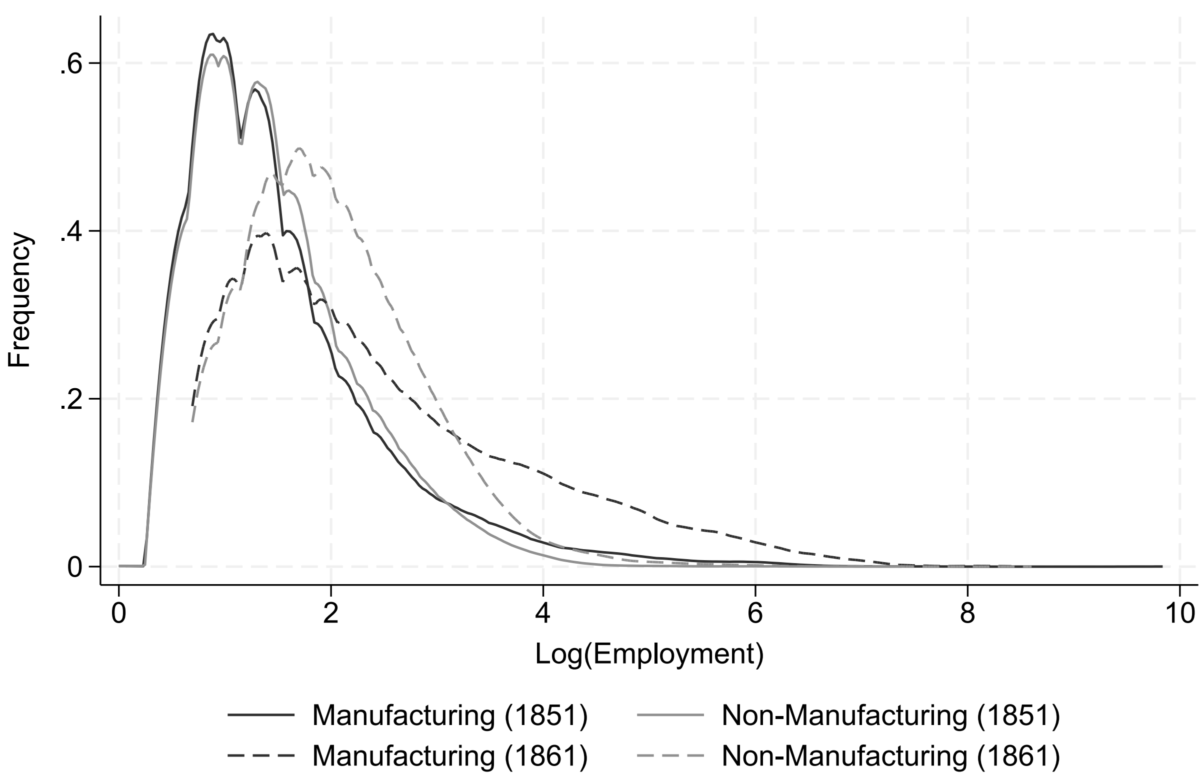
Table 2 compares several measures of heterogeneity across the four main sectors in 1851. This will later constitute the start of the sample for our main empirical analysis. The investigation shows that in the manufacturing sector, the standard deviation in the number

[3]See https://historyofwork.iisg.amsterdam/index.php for more information

of employees was 4 to 10 times bigger than in other sectors. If we take the ratio of the size of the fifth largest firm compared to the median, it is also at least double the ratio compared to the other sectors. The manufacturing sector also has the highest Gini coefficient and the highest general entropy score for the number of employees. In other words, the manufacturing sector seems more heterogeneous than other sectors. This heterogeneity seems, at least in part, explained by the existence of very large firms.

To see how the distribution of firms evolved over time, Figure 2 plots the firm size distribution for 1851 and 1861, the two censuses that included open items where firm owners could state their occupation. This decade is an important period for our study since this was the first one during which more than half of England and Wales was connected to the train network. In 1851, the plain lines show that the density function of the manufacturing sector is quite similar to the density function of other sectors. As expected, the right tail of the distribution is thicker for the manufacturing sector, reflecting the existence of some large firms in this sector.

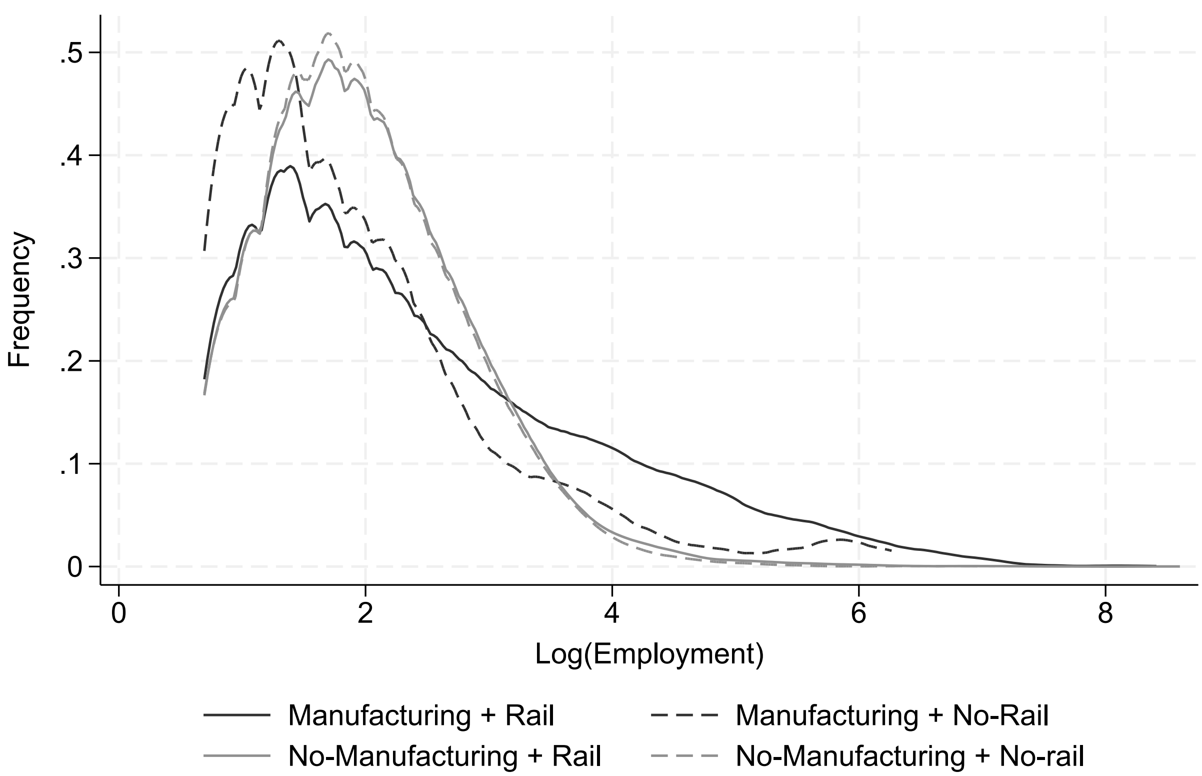
Figure 2: Firms' heterogeneity by sector in 1851 and 1861



Notes: This figure displays Kernel density functions for the number of employees based on data on business owners from the British 1851 and 1861 censuses. Bold lines display the densities from the 1851 census, dashed lines from the 1861 census. Dark grey lines display density functions for the manufacturing sector, and light grey lines for all other sectors.

The distributions diverge in 1861 as shown by the dashed lines. Both distributions shift to the right. However, they do not look alike anymore. The right tail of the manufacturing sector is now very different from the right tail of the other sectors. The kernel density of manufacturing firms hence became flatter and the number of smaller- to medium-sized firms decreased. Figure 3 further investigates the reason for this shift by distinguishing the distributions of firms by sector in 1861 with respect to their access to the rail network in 1851. The distribution of non-manufacturing firms (in grey) is exactly the same whether the firms are connected to the rail or not. On the contrary, the distribution of firms in the manufacturing sector depends on whether or not they were connected to the rail in 1851. In comparison to the non-manufacturing sector, the manufacturing firms not connected to the rail (black dashed line) are smaller but have a similar right tail. The manufacturing firms connected to the rail (black plain line) have a density function that is very different from the others. The right tail is thicker, meaning that the rail promoted the growth of many exceptionally large firms.

Figure 3: Firms Heterogeneity in 1861 and Rail presence



Notes: This figure displays Kernel density functions for the number of employees based on data on business owners from the British 1861 census. Bold lines display the density functions across locations with railway access, while dashed lines show density functions for locations without railway access in 1861. Dark grey lines display density functions for the manufacturing sector and light grey lines for all other sectors.

This exploratory evidence suggests that the interaction of heterogeneity and access to the rail impacted market structures. The rail encouraged the development of larger firms in the manufacturing sector. The consequence of this change in market structures remains to be understood. The shift in the distribution of firms could originate from the growth of incumbent firms or from the exit of some of the smaller firms. The following section develops our empirical strategy to investigate this question.

3 Empirical strategy

3.1 Data

We construct our main dataset at the grid cell-sector-decade level. Our spatial unit of observation is the hexagonal grid cell with an average area of 214 square kilometers. We observe each grid cell in each census year, i.e. in 1851, 1861, and 1881, and for each main sector, i.e. agriculture, manufacturing, trade, and services. Across the space-time-sector dimension, we collect data on bankruptcies and employment. We complement our dataset with additional data at the sector-year level, e.g. railway access and population.

Bankruptcy Data. We collect information on individual bankruptcy cases from publications in the London Gazette. Already early in the 18th century, British bankruptcy law required making insolvencies public such that potential creditors had the chance to make their claims official and be considered in the debt clearing process. For this purpose, the London Gazette contained a separate section that announced new bankruptcy adjudications and informed debtors on ongoing cases. 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. The first bankruptcy notice was published in the issue of June 5th, 1712. We accessed all digitized London Gazette issues from June 1778 until today via the official London Gazette homepage.[4] From 1778 until 1986, the publications of bankruptcy announcements followed a relatively fixed structure, which allows us to easily collect and encode individual cases.

To gather the individual bankruptcy announcements, we web-scraped scans of the 42,771 London Gazette issues published from 1788–1986 from the London Gazette homepage. These 42,771 issues include several supplemental publications that contain special information, but no bankruptcy announcements. We found 21,292 regular issues to include at least one

[4]For more information and to access the London Gazette issues, see https://www.thegazette.co.uk/. Unfortunately, issues from before 1778 were lost in a fire.

Figure 4: Examples of Bankruptcy Announcements

WHereas a Commiſſion of Bankrupt is awarded and iſſued forth againſt Joſeph Fernandes, late of Chelſea in the County of Middleſex, Wine-merchant, Dealer and Chapman, but now a Priſoner in the Fleet-Priſon, and he being declared a Bankrupt is hereby required to ſurrender himſelf to the Commiſſioners in the ſaid Commiſſion named, or the major Part of them, on the 18th Day of Auguſt next, at Twelve o'Clock at Noon, on the 19th Day of the ſame Month, and on the 9th Day of September following, at Eleven o'Clock in the Forenoon, at Guildhall, London, 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 Second Sitting to chuſe Aſſignees, and at the laſt Sitting the ſaid Bankrupt is required to finiſh his Examination, and the Creditors are to aſſent to or diſſent from the Allowance of his Certificate. All Perſons indebted to the ſaid Bankrupt, or that have any of his Effects, are not to pay or deliver the ſame but to whom the Commiſſioners ſhall appoint, but give Notice to Mr. Moſely, Shoe-lane, London.

(a) Bankruptcy Announcement 1788

FIRST MEETINGS AND PUBLIC EXAMINATIONS.

ROSENBERG, Lewis (formerly trading as the VICTORIA TIMBER COMPANY), of and carried on business at 43A, Durant-street, Hackney-road, London. TIMBER MERCHANT.
Court—HIGH COURT OF JUSTICE.
No. of Matter—437 of 1929.
Date of First Meeting—June 26, 1929. 11 a.m.
Place—Bankruptcy Buildings, Carey-street, London, W.C. 2.
Date of Public Examination—July 16, 1929. 11 a.m.
Place—Bankruptcy Buildings, Carey-street, London, W.C. 2.

(b) Bankruptcy Announcement 1929

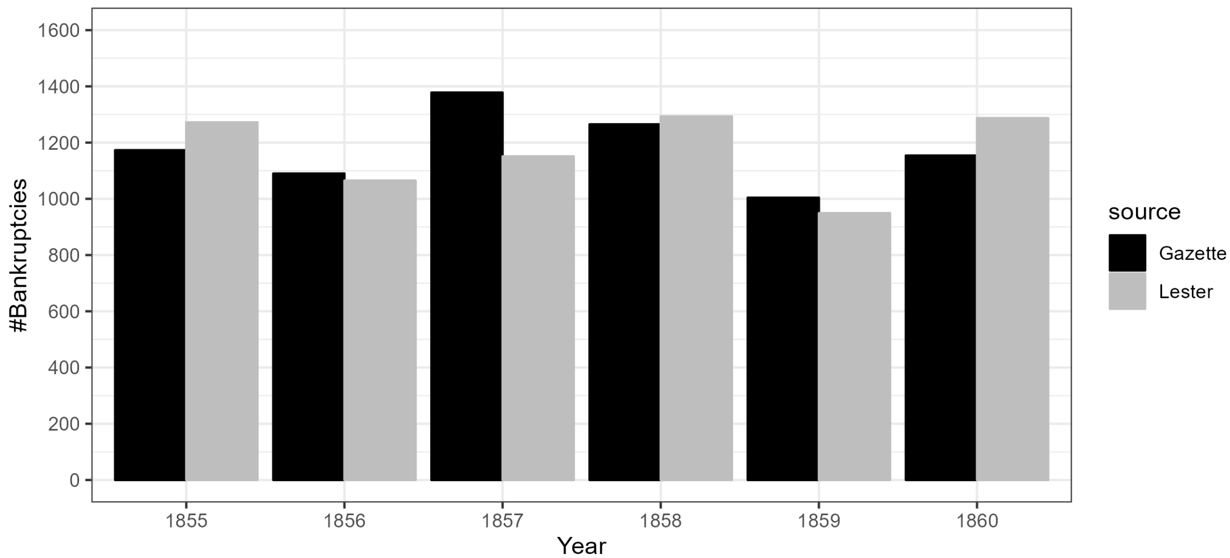
Notes: This figure illustrates the layout of the original London Gazette files based on which the bankruptcy data were collected. Figure (a) shows one of the oldest texts from 1788, and Figure (b) displays a newer entry from 1929. From these texts, our algorithm would collect the information on the bankruptcy's name, address, and occupation.

bankruptcy statement each. Figure 4 illustrates two examples of how the bankruptcy cases were announced in the London Gazette. To convert these images to data, first used Optical Character Recognition (OCR) software to convert the scans into a machine-readable text format, then and started the computational processing. We constructed various text recognition algorithms to detect individual bankruptcy announcements based on specific keywords. For each announcement, we extracted the bankrupt's name, address, and occupation. In a final computation step, we geolocated each address, usually at the city- or parish-level, and assigned the people's occupation to a sector by assigning History of Work Information (HISCO) codes according to people's occupation description. We describe the data collection process in more detail in Appendix B.

Figure 5 compares the yearly number of bankruptcies in our dataset to officially published statistics at the national level as collected by Lester (1991). Indeed, our coding follows the general trend very closely. This makes us confident that sampling bias is unlikely to affect our estimations other than by increasing standard errors due to random measurement error.

British Microcensus. We use British microcensus 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). The I-CeM project digitized full, individual-level census data for England and Wales in 1851, 1861, 1881, 1891, 1901, and 1911. Importantly for us, all census entries contain information on people's

Figure 5: 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.

occupations, for which the I-CeM project already coded the associated HISCO codes. Other control variables we add via the microcensus data are, among others, local population, age structures, gender ratios, and internal migration stocks. We assigned coordinates to all census observations based on the sub-district people stated in the survey and intersected the subdistrict coordinates with our grid cells. While the dataset also contains spatial information at the higher-resolution parish level, we restricted ourselves to the sub-district level because many historical parishes do not exist anymore today and names were not unique, such that an accurate geocoding of parishes across census waves was impossible.[5]

Additional Data. We complement our dataset with additional data sources that vary at the grid-cell level, over time or both. First, we use data on the locations of railway stations in England and Wales in 1851, 1861, and 1881 from Martí-Henneberg et al. (2017a,b,c). We spatially intersect the railway shapefile with our grid cell dataset, and assign to each grid cell the sum of stations in a given year. We further leverage data from Fernihough and O'Rourke (2014), which locate the British towns with access to coal. We calculate the distance of each grid cell's centroid to the closest town with coal access as a proxy for coal availability in a location. As final geographic covariates, we calculate the distance to London, the coast, and

[5]The spatial representativeness of subdistricts varies with population densities. Bigger cities like London or Liverpool consist of tens of subdistricts. Yet, in very rural regions, especially in Wales and Cornwall, the subdistrict density is rather low. As can be seen in Figure 6 below, we end up with some grid cells in these rural regions that do not contain any subdistrict coordinates. We therefore drop these grid cells from our estimations.

UK ports from every grid cell's centroid.

Our final dataset follows the structure of the British census waves at the grid cell-sector-decade level. This is, we observe each occupation sector in each grid cell for every decade from 1851 to 1881, with the exception of 1871 when no British census data are available. For each census period, we add the sector-level annualized number of bankruptcies as our main dependent variable. For this, we aggregate all bankruptcy cases in a sector and grid cell between two census periods, divide it by the number of years between the two census periods to control for the longer time span between 1861 and 1881, and assign this number to sector-grid cell observations in the year that begins the respective decade. We aggregate occupation sectors to the highest, 1-digit occupation category which divides occupations into four main occupation groups.[6] We assign some observations to another category "Other". These include pensioners, rentiers, or unemployed individuals, which we drop from our analysis. Our main estimations will focus on the census years 1851, 1861, and 1881 for which we have separate data on the spatial distribution of railway stations at each point in time.

3.2 Econometrics

The main econometric specification leverages the three dimensions of information we have on bankruptcies. To estimate the specific effect of a railway connection on a sector, we use the variation within areas becoming connected to the railway network. We use the Pseudo-Poisson-Maximum-Likelihood (PPML) estimator to account for the overdispersed distribution of our dependent variable. Our main specifications follow a Triple Difference-in-Differences intuition, where the treatment variable indicates having access to the railway network. Railway access varies over time, and we interact this indicator with sector fixed effects to estimate different effects across sectors. Our main specifications take the following form:

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

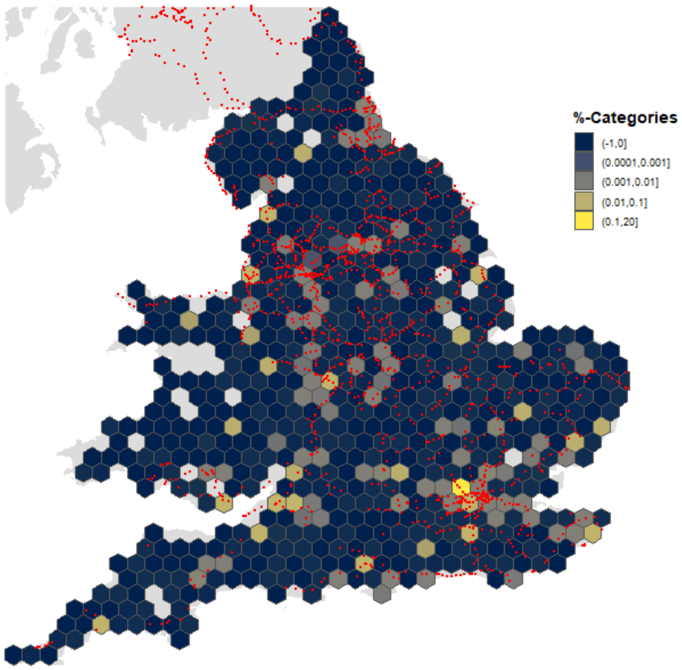
[6]The original HISCO coding divides occupations into 10 main categories. For our purposes, we re-arrange these ten groups slightly. First, we combine the groups "0" and "1", which both refer to "Professional Workers," with the "service" category. This category serves as the comparison group in most empirical analyses. 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"), "service" ("5"), and "sales" ("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 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 sales. In total, we re-arrange 10 occupation categories this way.

Our dependent variable, denoted by $Bankruptcies_{i,s,t}$, is the number of bankruptcies in some grid cell i, sector s, and year t. $\mathbb{1}Rail_{i,t}$ is a dummy variable equal to one once a grid is connected to the rail network. With ν_t, η_i, and θ_s we include year, location, and sector fixed effects, respectively. These fixed effects reduce our identifying variation to within geographic areas, within sectors, and within time periods. Our results can then not be explained by any geographic characteristics (such as the proximity to resources), by any sector characteristics (such as heterogeneity within a sector), or by global temporal shocks (such as economic crises). To minimize the remaining bias in our OLS estimations, the matrix $X_{i,s,t}$ adds a number of relevant control variables. Most importantly, we control for the level of employment in each sector-location-year to account for a potential scale effect in our more conservative specifications. We also control for the distance to coal, the distance to London, and the distance to the nearest port, each interacted with sector fixed effects. Distance to coal is an important control variable as it proxies for a location's propensity to industrialize (Fernihough and O'Rourke, 2020). Holding the distance to London constant is necessary to account for differences in the availability of investment capital, as London was the main financial center of the time. Finally, by controlling for the distance to the closest port, we account for locational differences in the exposure to international trade and migration. By interacting each of these three variables with sector fixed effects, we allow these confounders to have different impacts on bankruptcies or employment across sectors. The remaining unexplained variation in the number of bankruptcies or employment is denoted by the error term $\epsilon_{i,s,t}$.

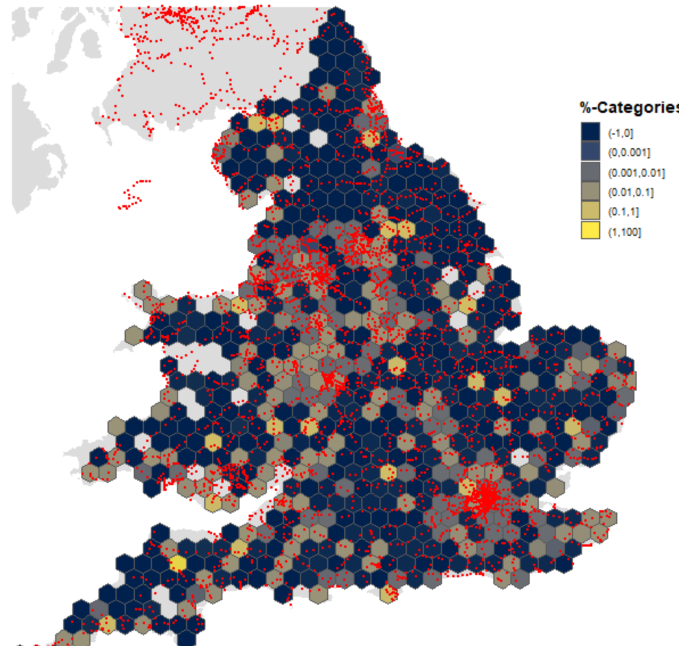
Our coefficient of interest is $\beta_{1,s}$, which captures the sector-specific reaction of the number of bankruptcies or employment to a connection to the railway network. Section **??** shows that both the temporality of the effect and its spatiality suggest that the effect we observe is causal.

Note that our specifications derive $\beta_{1,s}$ from the interaction of two baseline variables that are both collinear to the fixed effects or to the $\mathbb{1}Rail_{i,t}$ variable present in our model. This is, a location's changed access to the railway over time is controlled for by the location-time $\mathbb{1}Rail_{i,t}$, while sector characteristics are controlled for by the sector fixed effects ν_s and temporal shocks by θ_t. Hence, one must interpret $\beta_{1,s}$ as the *differential effect* the railway expansion has on a specific sector. In our baseline estimations, we will estimate the effect on Manufacturing with respect to other sectors. In the second step, we also estimate the elasticity of the number of bankruptcies for each sector. As emphasized by Juhász et al. (2020), some sectors reorganize following the arrival of new technologies. In this study, we investigate how the arrival of the rail had accelerated the re-organization of the one sector most likely affected by increased access to trade: manufacturing.

Figure 6: 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 colors indicate higher shares of bankruptcies. The red points indicate railway stations that were established at the beginning of the respective data sample. Light-grey cells are low-populated places and were omitted from the dataset because they do not contain a census sub-district, returning any census information.

4 Results

4.1 Illustration

Our empirical strategy relates the expansion of the British railway network to the occurrence of individual bankruptcies. Figure 6 illustrates our empirical analysis. The two maps show England and Wales covered by hexagonal grid cells, our unit of observation. Colors 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.

Figure 6 a) plots the extent of the railway by 1851 together with the aggregate number of bankruptcies from 1851-1860 relative to 1851 employment.[7] Figure 6 b) again illustrates the geographical correlation between the railway expansion and the occurrence of bankruptcies but for bankruptcies in the 1881-1890 period and the railway network in 1881.

Both maps illustrate the intuition behind our analysis very well. In 1851, the British railway system was still in its infancy, with 56% of cells having at least one railway station. Hence, almost half of the locations were not yet connected to the railway network, and the overall density of railway stations was low. Similarly, we only observe a small number of bankruptcies over the following ten years, with many grid cells not even experiencing one. Remarkably, however, the gridcells experiencing bankruptcies do almost all contain at least one railway station. In 1881, the railway network was much more advanced. According to our measure, more than 90% of cells now have at least one railway station. Not only does the overall number of bankruptcy cases increase; but we also see many grid cells lighting up now that did not exhibit any bankruptcy cases in the period before. And yet still, the bankruptcy cases closely trace the spatial extent of the railway network.

It should be noted that 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 extension between 1851 and 1881. The fixed effects are not collinear with the interaction $Rail_{i,t} \times Sector_s$ in the year 1851. Hence, our effects have to be interpreted as the effect of having a connection with the rail network and not as the effect of a station opening in the second phase of the expansion of the railway network. Table 6 further disentangles these two dimensions of our main estimator.

[7] Note that the map contains a number of gray cells, especially in the rural regions of Wales or Cornwall. These are grid cells that we dropped from the dataset because our coding of census sub-districts did not yield any matches for these grid cells in these very rural parts of Great Britain.

4.2 Baseline results

Table 3 presents our main results based on PPML regressions. Column 3.1 presents the coefficient from regressing the number of bankruptcies on the dummy variable capturing connection to the rail network. We do not yet include any controls or fixed effects, and let the railway affect bankruptcies across all sectors to the same extent. Our results mirror the image from Figure 6 and demonstrate that on average, being connected to the rail network is associated with a higher number of bankruptcies. According to this estimate, gridcells connected to the rail network experience on average around three times as many bankruptcies as non-connected cells.

Table 3: Main Results - The effect of the rail on bankruptcies

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	2.83***	2.71***	0.04	0.06	0.06	0.06	0.07	0.05
	(0.21)	(0.21)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s		0.40***	0.40***	0.40***	0.41***	0.40***	0.38***	0.43***
		(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
Observations	8185	8185	7293	7293	7293	7293	7293	7293
Pseu. R^2	.0765	.0852	.827	.827	.827	.827	.827	.827
Sector FE			✓	✓	✓	✓	✓	✓
Geo FE			✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
$\text{Sector Employment}_{i,s,t}$				✓	✓	✓	✓	✓
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 annualized number of bankruptcies that occurred between two census periods. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. The main control variables are the number of people employed in sector s as well as the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

In Column 3.2, we add an indicator variable for the manufacturing sector together with an interaction term of both explanatory variables to differentiate the effect of the rail by sector. The interaction term suggests that the relationship between the railway expansion and the number of bankruptcies is most predominant in the manufacturing sector. The coefficient implies that a connection to the rail network would increase bankruptcies in the manufacturing sector by around 49 additional percent in comparison to other sectors.[8] Across the following columns, we progressively add fixed effects for sector, year, and location, as well as our different control variables. From Column 3.3 to 3.8, the interaction between the rail

[8]According to $\Delta = (e^{0.4} - 1) \times 100$.

dummy variable and the manufacturing sector variable remains significant and positive, while the coefficient for the rail variable turns insignificant once we account for location fixed effects. Accordingly, the average effect of the railway on bankruptcies is zero once we account for location-specific characteristics. However, all fixed effects and control variables do not change our estimate for the effect of rail access on bankruptcies *in the manufacturing sector*. For this sector, our estimates suggest a robust increase in bankruptcy incidence by around 49 percent. Indeed, neither controlling for employment at the sector-geo-time level nor adding the high-dimensional fixed effects changes the estimate significantly. This suggests that the effect we observe goes well beyond just a sector size effect and is also not solely driven by between-sector reallocation. Similarly, the effect is not explained by the specificities of the manufacturing sector in areas close to coal, ports, or London, which were probably more likely to be connected to the rail.

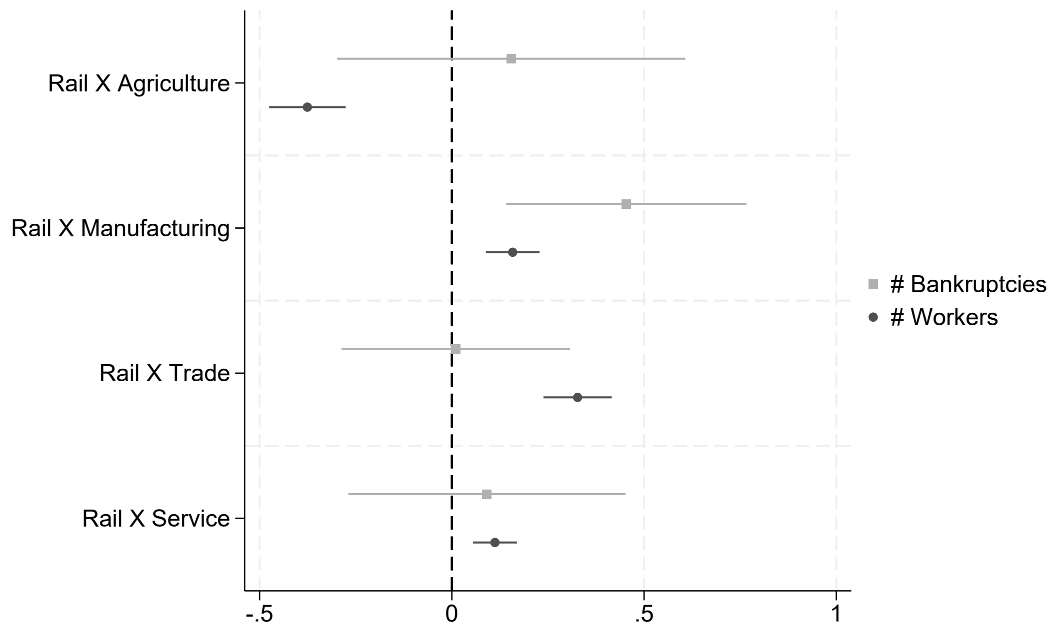
Table 4: Main Results - The effect of the rail on Employment

			Dependent Variable: #Employed$_{i,t,s}$					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$ (Rail$_{i,t} > 0$)	1.11***	0.83***	-0.48***	-0.27***	-0.17***	-0.27***	-0.31***	-0.13***
	(0.10)	(0.09)	(0.05)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)
$\mathbb{1}$ (Rail$_{i,t} > 0$) × Manufacturing$_s$		0.59***	0.59***	0.59***	0.38***	0.59***	0.67***	0.29***
		(0.07)	(0.07)	(0.07)	(0.07)	(0.06)	(0.07)	(0.06)
Observations	8496	8496	8496	8496	8496	8496	8496	8496
Pseu. R^2	.0691	.226	.864	.901	.919	.902	.904	.92
Sector FE			✓	✓	✓	✓	✓	✓
Geo FE			✓	✓	✓	✓	✓	✓
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 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. The main control variables are the total population in the cell as well as the straight-line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Standard Errors in parentheses are clustered at the grid cell level, * $p<0.1$, ** $p<0.05$, *** $p<0.01$

The increased bankruptcies observed in the manufacturing sector can have two origins. Either they result from a decline in sector activity, or they are the product of labor reallocation in the wake of a deeper market integration. To test which of these hypotheses prevail, Table 4 estimates the baseline regression using sectoral employment as the dependent variable. In Column 4.1, we observe that places connected to the rail network on average have higher employment. Column 4.2 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 4.3 to 4.8). The coefficient for the interaction of the manufacturing sector dummy variable with the rail dummy variable is significantly positive across all columns. Accordingly, the industrial sector of a connected location has a number of employees that is 34% to 95% higher than unconnected cells. If anything, a connection to the rail network does not seem to trigger a decline in the manufacturing sector but rather increases its dynamism. As in Bogart et al. (2022), we find that the expansion of the rail triggered a reallocation towards the manufacturing sector and no movement out of it.

Figure 7: Coefficient plot – Railways' 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 gray) 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 Appendices B.1 and B.2.

It then appears that the manufacturing sector experienced both an increase in bankruptcies and an increase in the number of employees along the development of the railway network. The effects documented in Table 3 and 4 are however relative to other sectors connected to the rail. To better grasp the between and within sector reallocation, Figure 7 shows the coefficients from interaction terms with the Rail dummy variable and an indicator variable for each sector. This figure summarizes which sectors underwent between-sector or within-sector reallocation. Looking at the estimates in dark grey, we can see that a connection to the rail decreased the number of workers in the agricultural sector while increasing the number of workers in all other sectors. We see that the largest effect the rail has is in increasing employment in the trading sector – which makes sense given the new opportunities the rail opens in this specific sector. A connection to the rail also increased employment in manufacturing and in Service. Structural change from the primary to secondary sector explains the increase in manufacturing employment whereas at the same time employment in service also increased to sustain larger production units (Katz and Margo, 2014). The estimates of the effect of rail on employment echo the findings of Bogart et al. (2022). The rail encouraged urbanization and reduced the activity in the primary sector. The estimates in light grey show the effect of a rail connection on bankruptcies for the different sectors. As in Table 3, the coefficient for Rail × Manufacturing is significantly positive. Estimates for other sectors are not different from zero. The rail hence increased the number of bankruptcies only in the manufacturing sector, while at the same time also increasing manufacturing employment. The manufacturing sector is hence the only sector that exhibits the dynamics we would expect under within-sector reallocation as outlined in Table 1 above.

These results are robust to a wide set of tests presented in Section 6 and detailed in Appendix C. In the next section, we test the assumptions for causality of our estimates.

5 Identification

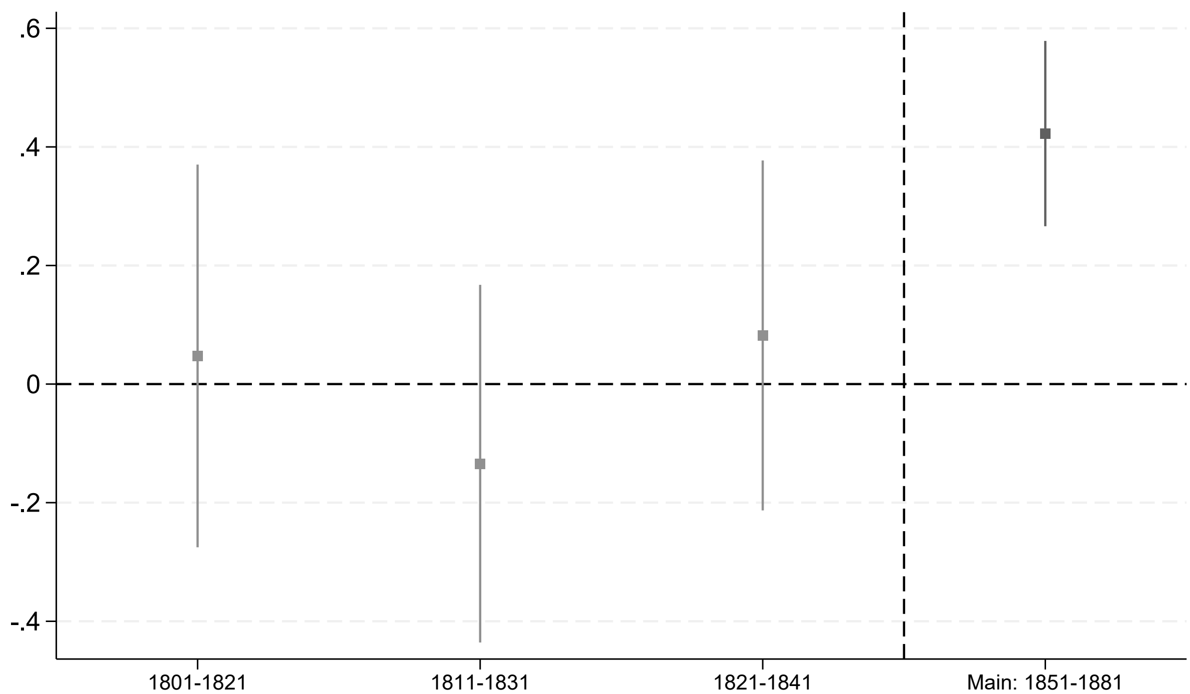
Our estimators add several fixed effects and control variables that could correlate with both the connection to the rail network and the number of bankruptcies in the manufacturing sector. The identifying variation excludes geographic features, variation over time, and sector characteristics. Despite this restrictive set of fixed effects and control variables, one may argue that the interaction of the rail connection variable with the manufacturing sector variable may reflect other dynamics that would vary over time-location and affect specifically the manufacturing sector. To circumvent this potential pitfall, we document how much the timing of the effect and its spatiality show that our estimates can be considered as causal.

5.1 Time Dimension – Pre-treatment placebos

Our panel estimations follow the logic of triple Difference-in-Differences estimations, as we exploit variation across place, time, and sector. We are interested in the coefficients of the interaction between railway expansion and an indicator variable for the manufacturing sector. Our identifying assumption is hence that the potential outcomes of employment and bankruptcies in the manufacturing sector would have been the same across locations with and without rail access in case the railway would not have been built. This assumption is close to a parallel trends assumption in a difference-in-differences framework.

We test the plausibility of this assumption by looking at 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, use the railway expansion from 1851-1881, but use bankruptcies in the periods 1801-1821, 1811-1831 and 1821-1841 as the dependent variable. Figure 8 presents the estimates by year in a coefficient plot. None of the coefficients of our placebo pre-treatment estimations is significant. Their standard errors are large and the point estimates are always close to zero. Moreover, the coefficients do not exhibit any specific upward or downward trends. Hence, the increased bankruptcy trends in the manufacturing sector among places connected to the rail from 1851 onwards were not yet present in the 50 years before the rail was actually constructed.

Figure 8: Testing Parallel Pre-Trends – Coefficients $\text{Rail}_i \times Industry_s$ 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 run the same regressions as for our main results, but regressing bankruptcies that occurred in these earlier periods on the placebo rail expansion variables.

These placebo estimates hence show that the manufacturing sector in places connected to rail from 1851 onwards did not experience more bankruptcies before 1851. Accordingly, the location of the rail from 1851 to 1881 does not pick up any long-term geographic patterns that would be specific to the manufacturing sector. They do not as well capture long-term geographic characteristics that would apply to all sectors since we have added location fixed-effects in our estimations.

5.2 Space dimension – IV estimates

Our second test leverages upon an exogenous variation of the connection to the rail. In the main specification, location fixed effects directly control for locations' different exposure to the railway construction, but not for potential characteristics of the manufacturing sector in those places. We use a Least Cost Path (LCP) approach similar to Bogart et al. (2022) to instrument the access to the rail network.[9] To generate the time variation in our instrument, we then create different buffers around the LCP that are equal to the distance between the LCP and the median distance of stations to the LCP for this given year. We actually let this threshold vary between the 40^{th} percentile and the 60^{th} percentile. The intuition of the instrument comes from the way the railway evolved during 19^{th} century. The main network was concluded around 1850. From that point onwards, the network mainly branched out to reach smaller cities close to the existing network. We will then use this instrument in reduced form regressions as an alternative treatment indicator that is arguably exogenous to local, sector-specific characteristics.

Figure 9 shows the first stage of this instrument. The x-axis presents the percentile of distance used in the construction of the instrument, while the y-axis shows the coefficient β_5 of the following equation:

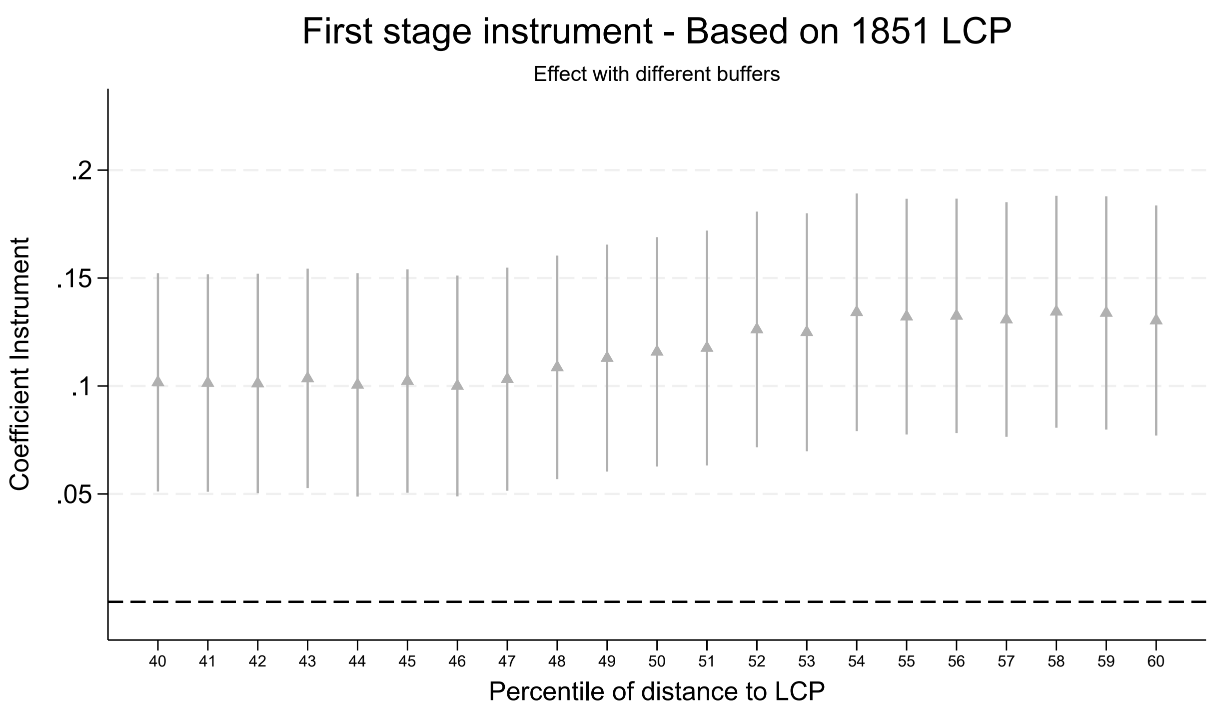
$$\begin{aligned} \mathbb{1}Rail_{i,t} = \beta_1 DistCoal_i + \beta_2 DistPort_i + \beta_3 DistLondon_i + \\ \beta_4 RailNodes_i + \beta_5 \mathbb{1}(DistLCP_i < Xpct_{i,t}) + \nu_t + \epsilon_{i,s,t} \end{aligned} \tag{2}$$

We control for distance to the main determinants of industrial activities: Coal, Port, London, and the rail nodes in our sample. This way, our instrument captures the inconsequential

[9] Following Bogart et al. (2022), we select the 100 biggest towns in 1850 as natural railway nodes, i.e., as towns that with almost certainty would have been among the first towns 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, taking into account the bilateral distance together with the variation in construction costs due to elevation (which required building tunnels) and rivers (which required building railway bridges). We use these LCPs to code grid cells' propensity to be connected to the railway. The main assumption here is that if a location lies along the LCP between two nodes, the railway lines must go through this location, which automatically increases the location's likelihood of receiving a railway station.

locations that happened to be on the way between two nodes but were not themselves targeted by the extension of the network. We also control for year fixed effects to be sure that our instrument does not capture other shocks in time.

Figure 9: First-stage results - LCP and extensions

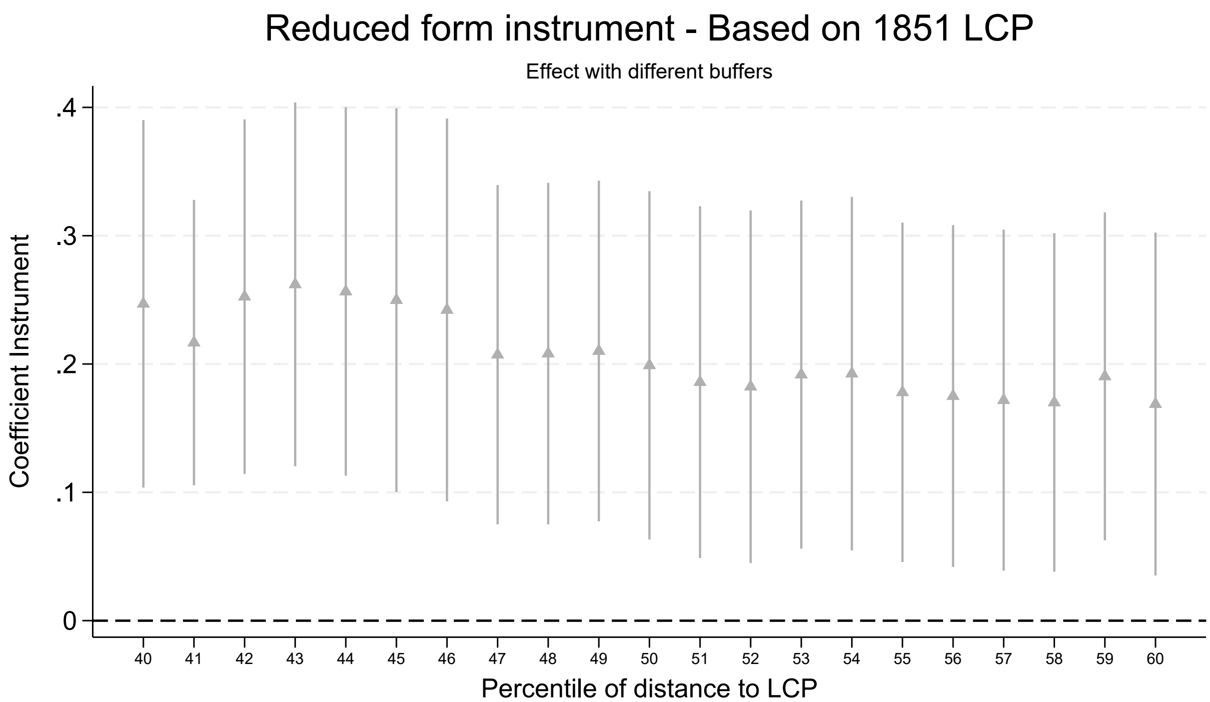


Notes: Results of the estimation of equation 2. Coefficients of the instrument for different values of the percentile considered in the extension of the network (X= 40 to 60^{th} percentile). Standard errors are clustered at the location level. Lines depict 95% confidence intervals.

All the coefficients of the instruments are positive irrespective of the size of the buffer chosen around the LCP. It should be noted that the smaller the percentile used for the LCP the more false negatives we will have in the construction of the instrument. Conversely, the larger the percentile the fewer false negatives the instrument will reach but the more false positives will be included. The Figure shows that the right balance between the two types of error seems to be around the 50^{th} percentile of distance. In that case, being within the distance buffer of the LCP increases the likelihood of being connected to the rail by 10 to 15 percentage points.

Our instrument hence is a good exogenous predictor of a connection to the rail network. Figure 10 further shows that this instrument also predicts more bankruptcies in the manufacturing sector in reduced form estimates. Accordingly, a higher predicted likelihood of being connected to the rail network increased the number of manufacturing bankruptcies by around 20%. Our inconsequential places approach reaches results similar to our baseline results. The magnitude of these coefficients is hard to compare to our baseline estimates since they capture an intention-to-treat effect and not the average treatment effect.

Figure 10: Reduced form results - LCP and extensions



Notes: Coefficient of reduced form estimations using our instrument in equation 1 instead of the rail measure. The instrument is constructed using the percentile in the distance of stations to the 1851 LCP. Standard errors are clustered at the location level. 95% confidence intervals are shown.

Taking stock, our baseline results appear at the actual date of development of the rail and not earlier. Moreover, when we instrument for the main spatial variation we leverage upon in our estimations, we still find that locations that got a connection to the rail because they were along the way between two important places did also observe a surge in bankruptcies in the manufacturing sector following the arrival of the rail. In Appendix D, we test whether our instrument actually identifies locations getting *new* access to the railway network. Our regression evidence robustly shows that the correlation between the number of stations and our instrument is not significant. Yet, our reduced form estimates remain strongly significant when we control for the number of railway stations as we show in the Appendix Table D.1. Taken together, this emphasizes that our instrument identifies the effect of new market integration on bankruptcies, but does not work through the intensity of the connection.

Table 5: Main Robustness tests

Dependent Variable:	$\mathbb{1}(\text{Bank}_{i,t,s}>0)$	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	LPM	Inter FE 1	Inter FE 2	3w Cluster	No Nodes	Big Cells	Small Cells
$\mathbb{1}$	0.00		0.12	0.05	0.10	0.20	0.05
$(\text{Rail}_{i,t} > 0)$	(0.02)		(0.16)	(0.13)	(0.18)	(0.24)	(0.23)
$\mathbb{1}$	0.05**	0.44***	0.21**	0.43***	0.46***	0.57***	0.30***
$(\text{Rail}_{i,t} > 0) \times \text{Manufacturing}_s$	(0.02)	(0.09)	(0.09)	(0.11)	(0.10)	(0.14)	(0.11)
Observations	8185	6027	7293	7293	5841	4024	8816
Pseu. R^2		.843	.835	.827	.765	.711	.693
Sector FE	✓	✓	✓	✓	✓	✓	✓
Geo FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
$\text{Population}_{i,t}$	✓	✓	✓	✓	✓	✓	✓
$\text{Coal}_i \times \text{Manufacturing}_s$	✓	✓	✓	✓	✓	✓	✓
$\text{Port}_i \times \text{Manufacturing}_s$	✓	✓	✓	✓	✓	✓	✓
$\text{London}_i \times \text{Manufacturing}_s$	✓	✓	✓	✓	✓	✓	✓
$\text{Geo FE}_i \times \text{Time FE}_t$		✓					
$\text{Sector FE}_i \times \text{Time FE}_t$			✓				

Notes: Table reports results from OLS (Column 1) and PPML regressions at the grid cell-sector-census year level. The dependent variable in Column (1) is an indicator variable for at least one bankruptcy occurring at a given grid cell, sector, and census period. In all other columns, the dependent variable is the annualized number of bankruptcies that occurred between two census periods. 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. The main control variables are the total population in a grid cell as well as the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Columns (5)–(7) alter the sample for our analysis. In Column (5), we drop all grid cells that were atop the population distribution in 1850 and hence constituted important railway nodes. In Columns (6) and (7), we double (half) the average area of the grid cells on which we base our sample. Column (4) 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.

6 Robustness

We test the robustness of our estimations in different ways. Table 5 presents the main testa. Further tests are presented in the Appendix. In Table 5, we observe that the results remain robust to using a linear probability model to estimate the probability that places connected to the rail experience at least one bankruptcy in a given sector (Column 5.1). We see that the rail does not change the probability of having at least one bankruptcy for other sectors. It does however increase the extensive margin for the manufacturing sector by 5 percentage points. Results are also robust when we add Geo $_i$ × Year$_t$ Fixed effects and Sector$_i$ × Year $_t$ Fixed effects (Column 5.2 and 3) and when we cluster standard errors at the three levels of variation (Column 5.4). Our results also do not hinge on rail nodes. They remain significant and of similar magnitude when we exclude railway nodes from the sample (Column 5.5). The definition of the size of the grid we use for the estimation also does not drive our results. The results are similar if we multiply the grid cells' area by two or if we divide their area by two (Columns 5.6 and 5.7).

Beyond these first tests, we provide more details on the different robustness checks we have performed in Appendix C. Our results are virtually identical if we use a linear probability model throughout the different specifications of Table 3 (Appendix C.1). The results are also exactly the same if we control for different measures of local economic shocks 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 (Appendix C.2). The estimates also remain of the same magnitude when we exclude the 5% most populated cells, the 5% least populated cells, rail nodes, and all of them together (Appendix C.3). When the independent variable of interest is changed to the number of train stations in a grid cell, the results are also consistent with our baseline estimates (Appendix C.4). Results are further unchanged when we drop the "reform years" (1869-1883) from our sample which are years during which bankruptcies were simplified and creditors-led (Appendix C.5).

7 Understanding how rail access increased bankruptcies

In this section, we will try to pinpoint the exact mechanism driving the specific rise in bankruptcies in the manufacturing sector. This section provides two sets of results refining our baseline approach. First, we will add another layer of heterogeneity by comparing the manufacturing sector in cells already connected to the rail in 1851 with cells that got connected during our sample period. Second, we investigate whether the effect of rail on bankruptcies in the manufacturing sector was different during the "Reform years" as these incentives to file for bankruptcy were swapped during these years (Lester, 1991).

7.1 Heterogeneity in time, Heterogeneity in productive advantage

The mechanism we have emphasized so far suggests that rail connections particularly increased the number of bankruptcies in the manufacturing sector because this sector was more heterogeneous than others. By construction, the exposition to this heterogeneity was not the same for all firms in the manufacturing sector across all cells. The first cells connected to the rail however were not exposed to high competition and probably benefited from higher connections to intermediary goods and additional markets. The manufacturing sector in the last connected cells on the contrary likely experienced a competition shock. Our baseline effect encompasses the effect for both groups. Table 6 distinguishes these two effects. Columns 6.1 and 6.3 consider a railway connection in 1851 as the treatment variable whereas Columns 6.2 and 6.4 add Geo $\times$ Sector FE to the baseline estimation. In Columns 6.2 and 6.4 the interaction $\mathbb{1}\ (\text{Rail}_{i,t} > 0) \times \text{Manufacturing}_s$ then captures the effect of getting connected to the rail network (within Sector-Geo ID variation to access).

In Table 6, Columns 1 and 2 estimate the effect of these two types of variation on the number of bankruptcies whereas Columns 3 and 4 use the number of employees as the dependent variable. In Column 1, the coefficient attached to the interaction of the Rail network in 1851 with the manufacturing sector is equal to 0.12. Accordingly, being connected to the rail in 1851 increased the number of bankruptcies by 13% in the period 1851-1881. In Column 2, the within sector $\times$ geo variation in railway implies that getting connected to the rail in 1861-1881 increased the number of bankruptcies by 72 percent. The magnitude of the estimates is then more than six times that of a first mover connection. This difference in the estimates shows how dynamic the effect we estimate is. First movers did not suffer from the same competition when they got connected to the rail. This led to fewer bankruptcies. Meanwhile, the manufacturing sector in first connected cells could evolve and then face higher competition coming from the later connected. The later connected suffered the most as a consequence. Columns 6.3 and 6.4 confirm this intuition using the number of employees as the dependent variable. The coefficient for the manufacturing sector in cells connected to the 1851 network (Column 6.3) is ten times that of the coefficient for the within variation (Column 6.4). In cells connected to the rail in 1851, the manufacturing sector had 36% more employees in the manufacturing sector. The development of a connection in 1861-1881 increased the number of employees in the manufacturing sector by only 3%.

Interestingly, the positive effect on employment and on bankruptcies is still present using

Table 6: First Mover and the effect of railway extension

Dependent Variable:	$\#\text{Bankruptcies}_{i,t,s}$		$\#\text{Employed}_{i,t,s}$	
	(1)	(2)	(3)	(4)
$\mathbb{1}$ ($\text{Rail}_{i,1851} > 0$) × Manufacturing_s	0.12**		0.31***	
	(0.06)		(0.06)	
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s		0.54***		0.03*
		(0.09)		(0.02)
Observations	7293	6095	8496	8496
Pseu. R^2	.827	.827	.921	.992
Sector FE	✓	✓	✓	✓
Geo FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Geo × Sector FE		✓		✓
Sector $\text{Emp}_{i,t,s}$	✓	✓		
$\text{Pop}_{i,t}$			✓	✓
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 variables are the annualized number of bankruptcies that occurred between two census periods (Columns (1) and (2)), and the number of employed people in a census year (Columns (3) and (4)). The main explanatory variables are indicators for a grid cell i having at least one rail station recorded in 1851, and in census year t, both interacted with an indicator variable for the manufacturing sector. The main control variables are the number of people employed in sector s, the total population, as well as the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01.

either of these variations. This suggests that the rail triggered a reallocation towards manufacturing even for the late-connected locations. Those late-connected locations moreover experienced more bankruptcies than the early-connected ones. The nature of the integration of markets both groups faced hence was not the same.

7.2 Heterogeneity, Large Firms, and Employment Status

Our baseline estimations identify variations in competition and firms' heterogeneity via the sector of activity. Despite the imperfection of firm level data in the period, it is possible to directly capture within-sector heterogeneity by looking at the characteristics of firms getting connected to each other via the rail. Table 7 presents estimates of an interaction of the rail dummy variable with different firm size indicators. Based on the British business census for 1851, we identify those firms that belong to the top decile of the national employment distribution across firms. We then construct additional variables that identify the employment in these exceptionally large firms, located in the same cell and belonging to the same sector as our unit of observations. Next, we construct different market access measures by counting the large-firm employment in grid cells within 100km, over 100km but within 250km, and over 250km away that have railway access and operate in the same sector as our unit of observation.

Column 7.1 already tests the argument using the employment of the largest firms in the same cell. The interaction with the rail variable bears a negative sign and implies that cells that host large firms and get connected to the rail were less likely to experience an increase in bankruptcies. 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 small firms to bankruptcy. Once the cells with large firms got connected to the rail, there were hence no small firms that would potentially suffer from a connection to the rail.

Table 7: Rail, Existing Market structure and Bankruptcies

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Large Firms $_{i,t,s}$	-0.40***				-0.39***
	(0.03)				(0.03)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Large (Dist<100km) $_{i,t,s}$		0.02			-0.09**
		(0.03)			(0.04)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Large (100<Dist< 250km) $_{i,t,s}$			0.09***		0.13**
			(0.02)		(0.05)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Large (Dist>250km) $_{i,t,s}$				0.08***	0.04
				(0.02)	(0.04)
Observations	7293	7293	7293	7293	7293
Pseu. R^2	.827	.827	.827	.827	.827
Sector FE	✓	✓	✓	✓	✓
Geo FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
$\text{Sector Employment}_{i,s,t}$	✓	✓	✓	✓	✓
$\text{Coal}_i \times \text{Manufacturing}_s$	✓	✓	✓	✓	✓
$\text{Port}_i \times \text{Manufacturing}_s$	✓	✓	✓	✓	✓
$\text{London}_i \times \text{Manufacturing}_s$	✓	✓	✓	✓	✓

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 periods. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. The main control variables are the number of people employed in sector s as well as the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Columns 7.2 to 7.4 then test the effect of a connection to the rail interacted with the employment in large firms in connected cells nearby. It considers multiple mutually excludable buffers to construct measures of exposure to larger firms. In Column 7.2, firms that are within a 100km radius do not impact the number of bankruptcies in a newly connected cell. They logically were already quite accessible using other modes of transportation. Yet, large-firm employment in cells located between 100 and 250km from a treated cell is associated with a significant increase in bankruptcies. In Column 7.5, we include all the indicators together and observe that large firms that were far enough not to be part of the local market but close enough to become competitors with the rail are the firms that increased the number of bankruptcies the most.

Table 8: The effect of the rail on Employment status

	Dependent Variable: $\#\text{SelfEmployed}_{i,t,s}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	1.17***	-0.07*	0.08***	0.06**	0.08***	0.12***	0.07***
	(0.10)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s	-0.06	-0.06	-0.23***	-0.14***	-0.23***	-0.33***	-0.16***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
Observations	8185	8185	8185	8185	8185	8185	8185
Pseu. R^2	.145	.864	.909	.916	.909	.912	.917
Sector FE		✓	✓	✓	✓	✓	✓
Geo FE		✓	✓	✓	✓	✓	✓
Year FE		✓	✓	✓	✓	✓	✓
$\text{Sector Employment}_{i,s,t}$			✓	✓	✓	✓	✓
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 annualized number of bankruptcies that occurred between two census periods. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. The main control variables are the number of people employed in sector s as well as the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

We further document how much the interaction of market integration and sector heterogeneity has affected market structures. To do so, we use data on employment status from the business census. These data record whether a worker is self-employed or employed in a firm. Table 8 presents the results of the estimation of Equation 1 using the number of self-employed as the dependent variable. According to the mechanism emphasized in Melitz (2003), the smallest firms should exit as they cannot afford to pay the cost to enter the trading sector. In this table, the coefficient attached to the interaction of the manufacturing dichotomous variable and the rail variable is negative and significant at the one-percent or five-percent level once we control for employment at the sector level (from Column 3 onwards). Our estimates hence imply that beyond its effect on employment, a connection to the rail also changed the employment composition. Overall, a connection to the rail decreased self-employment by around 15 to 20 percentage points. As the number of self-employed decreased in the manufacturing sector in areas connected to the rail, the size of firms increased. This pattern reflects the organizational changes firms underwent with the arrival of the rail. We hypothesize that these organizational changes toward larger firms brought new tasks and, hence, new occupations. Larger firms became more complex in terms of occupational composition. We test whether a connection to the rail increased the complexity of firms in the manufacturing sector. In Equation 1, we replace the dependent variable with a Herfindahl index based on the number of workers working in different occupations within each occupation category. Normally, as new occupations appear, the Herfindahl index decreases as the "market concentration" on specific tasks becomes less.

Table 9 presents the results. The results echo baseline results quite well. Once we control for sector employment (Column 4 onward), the coefficient attached to the rail variable is insignificant. Accordingly, a rail connection does not impact the *overall* Herfindahl index, i.e., it does not lead to more complex production processes or firms. However, the coefficient attached to the interaction of the rail variable with the manufacturing variable is always negative and significant. In places connected to the rail, employment in the manufacturing sector got less and less concentrated on specific occupations.

Table 9: Manufacturing sector, diversity of occupation and rail

		Dependent Variable: Herfindahl Occupations$_{i,t,s}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}$ (Rail$_{i,t} > 0$)	-0.08***	-0.02	0.03	0.01	-0.00	0.01	0.00	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\mathbb{1}$ (Rail$_{i,t} > 0$) $\times$ Manufacturing$_s$		-0.23***	-0.16***	-0.17***	-0.12***	-0.17***	-0.15***	-0.06*
		(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)	(0.03)
Observations	5954	5954	5952	5952	5952	5952	5952	5952
Adj. R^2								
Sector FE			✓	✓	✓	✓	✓	✓
Geo FE			✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
Sector Employment$_{i,s,t}$				✓	✓	✓	✓	✓
Coal$_i$ $\times$ Manufacturing$_s$					✓			✓
Port$_i$ $\times$ Manufacturing$_s$						✓		✓
London$_i$ $\times$ Manufacturing$_s$							✓	✓

Notes: Table reports results from OLS regressions at the grid cell-sector-census year level. The dependent variable is the Herfindahl index based on the number of workers in each occupation within a given sector in a given space in a given time. The main explanatory variable is an indicator for a grid cell i having at least one rail station recorded in census year t. The main control variables are the number of people employed in sector s as well as the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Standard Errors in parentheses are clustered at the grid cell level, * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

These results taken together clarify the organizational changes brought by the rail in the manufacturing sector. These changes impact market structures and the labor market. Firms got larger and more complex. This was harder for firms exposed to a larger competition because of the rail. The consequences for labor were then harder for laggards exposed to already organized firms.

7.3 Structural change and bankruptcies – Creditors' demand for capital or debtors' insolvency?

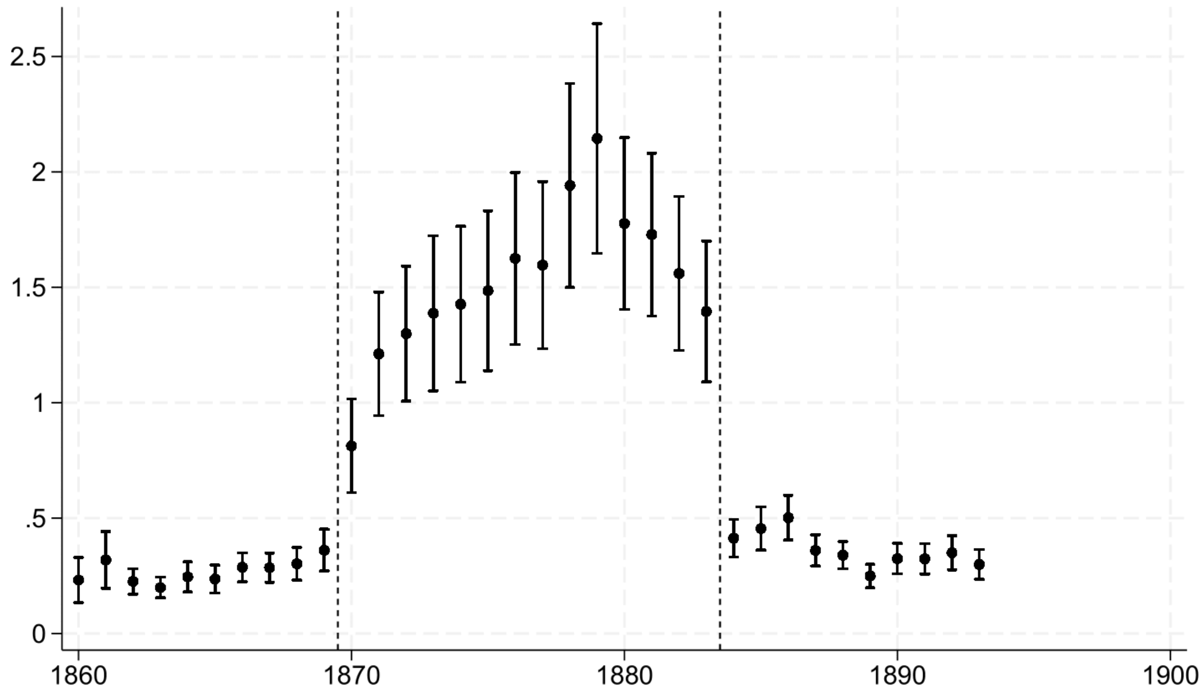
The argument developed so far rests a lot on market structure and on the cost of market integration for heterogeneous firms. A counterargument would be that with the arrival of the rail, some investors might have been harsher towards their debtors due to better re-investment alternatives. In that case, investors would have triggered bankruptcies to get part of the assets of the debtors.

To test this alternative explanation, we use the changes produced by two reforms of bankruptcy laws in 1869 and 1883. In 1869, England repealed the "officialism" doctrine for bankruptcy (Lester, 1991). Before this reform, bankruptcies were managed by local courts, often took a long time to be resolved and their outcome was uncertain. We hypothesize that during this period the "reinvestment" motive of creditors to file bankruptcy was limited. After the 1869 reform, bankruptcies were managed by creditors if a majority of them agreed. This procedure advantaged creditors and increased their incentives to file for bankruptcies for quick reinvestment. In 1883, England went back to the "officialism" doctrine.

Figure 11 illustrates the first fact about the repeal of officialism. It presents the annual number of bankruptcies around the reform. Creditors indeed appreciated the reform as we see two discontinuities at the time of the two reforms repealing and re-introducing officialism. The financial constraints of debtors could not have changed so dramatically overnight. These reforms created variation in creditors' incentives to file bankruptcies. Bankruptcies before the 1869 reform and after the 1883 reform can be thought of as an imprint of the financial situation of debtors. In between the two reforms, bankruptcies capture both the financial situation of debtors and creditors' interest. After the 1883 reform, the number of bankruptcies returns to the pre-1869 reform level lending more credence to our interpretation that the surge in bankruptcies in the 1869-1883 period was mainly due to the repeal of officialism and variation in creditors' incentives to file for bankruptcies.

If our effect would be explained by creditors' incentives, then we would expect a connection to the rail to increase the number of bankruptcies even more when "officialism" was

Figure 11: The 1869 and 1883 Reforms – Time discontinuities to study the motives for bankruptcies



Notes: Figure represents the annual distribution of bankruptcies across grid cells. Points represent the yearly mean values, lines indicate the 95% confidence intervals. Dashed lines indicate two significant reforms to the bankruptcy law: In 1869, the "officialism" system was repealed and the majority of creditors were allowed to manage bankruptcy cases independently. The reform in 1881 repealed the 1869 reform, and Britain returned to the officialism system.

repealed. We test this hypothesis using the dataset of bankruptcies at the yearly frequency. Table 10 uses a triple-interaction $\mathbb{1}$ $(\text{Rail}_{i,t} > 0) \times \text{Manufacturing}_s \times \text{Reform}_t$ on top of our previous estimations. This triple interaction investigates whether the shift in incentives towards creditors impacts the main effects. Table 10 shows that this triple-interaction does not turn significant. Accordingly, the main results are driven by the financial situations of debtors more than by the motivation of creditors to trigger the bankruptcies of some creditors on the edge.

Table 10: Incentives and the 1869/1883 reforms

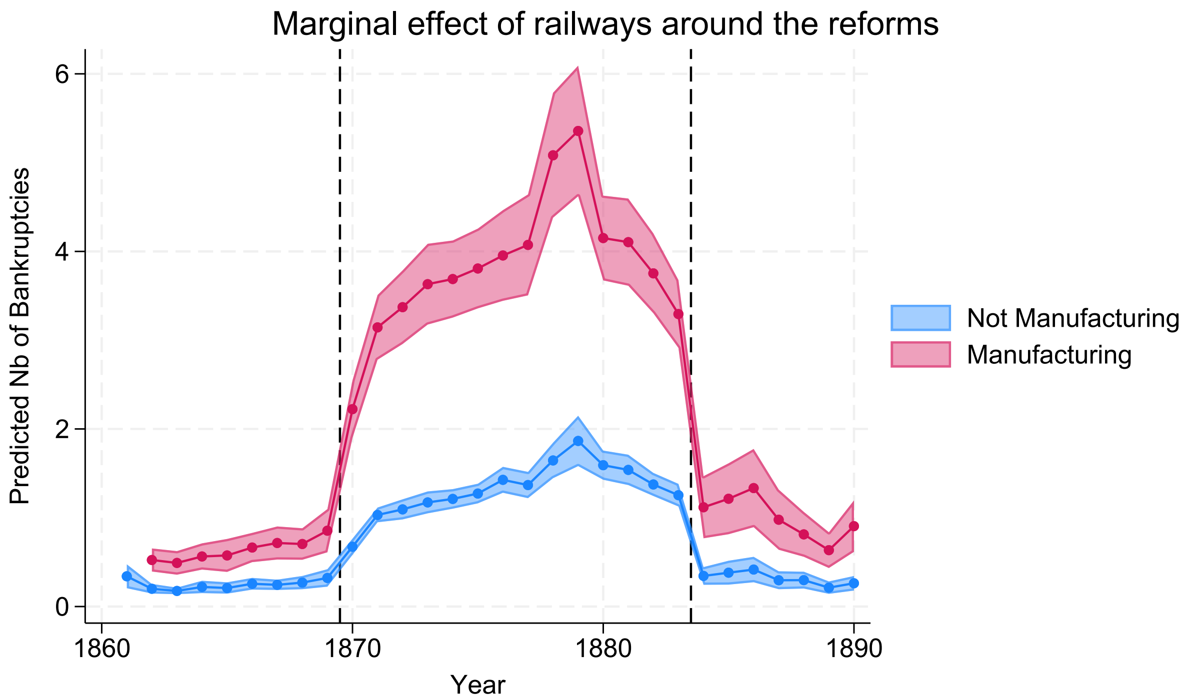
	Dependent Variable: #Bankruptcies$_{i,t,s}$	
	(1)	(2)
$\mathbb{1}$ (Rail$_{i,t} > 0$) $\times$ Man$_s$ $\times$ Ref$_t$	0.02	0.02
	(0.04)	(0.04)
$\mathbb{1}$ (Rail$_{i,t} > 0$) $\times$ Manufacturing$_s$	0.23***	0.21**
	(0.08)	(0.08)
Observations	111960	111960
Pseu. R^2	.822	.822
Sector FE	✓	✓
Year FE	✓	✓
Geo FE	✓	✓
Sector Emp$_{i,t,s}$	✓	✓
Coal$_i$ $\times$ Manufacturing$_s$		✓
Port$_i$ $\times$ Manufacturing$_s$		✓
London$_i$ $\times$ 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 periods. 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. The main control variables are the number of people employed in sector s as well as the straight line distance to the nearest city with coal deposits, to the nearest port, and to the city of London. Standard Errors in parentheses are clustered at the grid cell level, * p< 0.1, ** p< 0.05, *** p< 0.01

Figure 12 plots the intuition behind the estimates of Table 10. During the "Officialism" era, the effect of rail in the manufacturing sector is approximately twice to three times as large as for other sectors. The gap between the blue line for other sectors and the red line for the manufacturing sector increases during the repeal of officialism. In Figure 12, the coefficients attached to the manufacturing sector are also twice to three times larger than those attached to the effect of the rail in other sectors. Hence the effect of the reform does not appear when adding the triple-interaction terms.

This section documented the mechanisms behind our first compound effect. Within-sector reallocation is dynamic. As a consequence, being connected to the rail first does not have the same effect as getting this connection later. The largest losers are the last to integrate into the market. This integration into the market increases the number of bankruptcies because it shapes market structures and affects the least adaptable to the new market structures. We found no evidence that our effects are driven by the behavior of investors.

Figure 12: Rail and bankruptcies – Coefficients over time



Notes: Figure reports the marginal effect of regressing Equation 1, interacting the $\mathbb{1}Rail_{i,t} \times Sector_s$ interaction additionally with year dummies. The dependent variable is the annualized number of bankruptcies, the control variables follow the main specifications in Table 3.

8 Conclusion

This paper extends the geographic dimension to explain who loses from structural change within geographic units becoming integrated. Previous studies suggest that the least productive firms suffer from disruptive technologies (Juhász et al., 2020) and from market integration (Melitz, 2003). This article combines the two approaches to show that both factors interact with each other. The intuition is simple: the competitive disadvantage of technology non-adopters is not relevant if markets are not integrated. Similarly, market integration does not trigger firm exits if there are no heterogeneous costs associated with this integration (Melitz, 2003). In other words, our results directly reconcile the economic geography literature on intra-market reallocation and the literature on the effect of technology adoption without a geographic dimension.

The extension of the railway in England and Wales during the 19$^{\text{th}}$ provides us with a perfect setting to test the complementarities between these theories. Some sectors experienced dramatic technological changes whereas others did not. We hypothesize that these differences in the technology available to each sector can theoretically be associated with the heterogeneous costs faced by firms once trade costs decrease. In the meantime, the expansion of the railway provides us with variation in trade costs over space and over time. Our estimates assess how many individuals in each occupational class exited due to the shock generated by the railway expansion.

The railway expansion generated more bankruptcies in the manufacturing sector. At the same time, we document an increase in employment in the manufacturing sector due to railway access, which follows the interpretation of within-sector reallocation. We further illustrate that the manufacturing sector stands out by exhibiting significantly higher firm-size heterogeneity than the other sectors. Various extensions emphasize that this heterogeneity is the main driver of the within-sector reallocation we observe in the manufacturing sector.

These results 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., 2020). They also shed new light on the factors potentially explaining how spatial and sectoral inequality may interact today (Autor et al., 2020). This research emphasizes that despite positive aggregate effects technology and trade generate losers. This redistribution has important (political) consequences (Frey et al., 2018; Caprettini and Voth, 2020; Autor et al., 2020). Future research could build on these new results to better understand how to mitigate the redistributional consequences of growth.

References

Acemoglu, D. and P. Restrepo (2022). Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica 90*(5), 1973–2016. 2, 4

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. 4, 38

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. *The Quarterly Journal of Economics 135*(2), 645–709. 38

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, 4

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

Berger, T. (2019). Railroads and rural industrialization: Evidence from a historical policy experiment. *Explorations in Economic History 74*, 101277. 5

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

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. 3, 5, 7, 20, 21, 23

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

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, 38

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. *The Journal of Finance 63*(2), 565–608. 5

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. 38

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. 38

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

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics 131*(2), 799–858. 5, 7

Fernihough, A. and K. O'Rourke (2014). Coal and the european industrial revolution. The institute for international integration studies discussion paper series, IIIS. 13

Fernihough, A. and K. H. O'Rourke (2020). Coal and the European Industrial Revolution. *Economic Journal 131*(635), 1135–1149. 15

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. 38

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

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

Heblich, S., S. J. Redding, and Y. Zylberberg (2020). The distributional consequences of trade: evidence from the repeal of the corn laws. *Work in progress and incomplete*. 4

Hornbeck, R. and M. Rotemberg (2019). Railroads, reallocation, and the rise of american manufacturing. Technical report, National Bureau of Economic Research. 8

Hornung, E. (2015). Railroads and Growth in Prussia. *Journal of the European Economic Association 13*(4), 699–736. 7

Juhász, R., M. P. Squicciarini, and N. Voigtländer (2020). Technology adoption and productivity growth: Evidence from industrialization in france. Technical report, National Bureau of Economic Research. 2, 5, 15, 37, 38

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. 21

Lafourcade, M. and J.-F. Thisse (2011). New economic geography: the role of transport costs. *A handbook of transport economics*, 67–96. 7

Lester, V. M. (1991). *Insolvency and reform of English bankruptcy law, 1831-1914*. Ph. D. thesis, University of Oxford. 6, 12, 13, 28, 34

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

Martí-Henneberg, J., M. Satchell, X. You, L. Shaw-Taylor, and E. Wrigley (2017a). 1851 england, wales and scotland railway stations shapefile. 13

Martí-Henneberg, J., M. Satchell, X. You, L. Shaw-Taylor, and E. Wrigley (2017b). 1861 england, wales and scotland railway stations shapefile. 13

Martí-Henneberg, J., M. Satchell, X. You, L. Shaw-Taylor, and E. Wrigley (2017c). 1881 england, wales and scotland railway stations shapefile. 13

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *econometrica 71*(6), 1695–1725. 2, 4, 31, 37

Mitchell, B. R. (1964). The coming of the railway and united kingdom economic growth. *The Journal of Economic History 24*(3), 315–336. 7

Nye, J. V. (1987). Firm size and economic backwardness: A new look at the french industrialization debate. *The Journal of Economic History 47*(3), 649–669. 38

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. 38

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

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. 8, 12

Temin, P. (1997). Two views of the british industrial revolution. *The Journal of Economic History 57*(1), 63–82. 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 (December 8, 2022)*. 5, 8

Waugh, E. L. (1956). Railroads and the changing face of britain, 1825–1901. *Business History Review 30*(3), 274–296. 8

A Appendix A – Detecting Bankruptcies

For example, the earliest issues starting in 1788 and going until 1861 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. One exception is a short intermezzo in 1919 when they published notices of first meetings, intermediate meetings, and final meetings in separate tables at the end of an issue. Finally, from 1920 until 1986 the London Gazette went back to the structured text format illustrated in Figure 4 (b). Only after 1986, did lawyers and solicitors take over the management of bankruptcy cases and published announcements in their own, individual way. We, therefore, focus our systematic data collection on the 1788–1986 period when bankruptcy announcements followed systematic and easy-to-code patterns.

To extract individual announcements from an issue, we wrote various algorithms that, depending on the announcement pattern of a given time period, identified the start of a new announcement. For example, in the early issues from 1788–1861, 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.[10] 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.[11]

Our algorithm detected a total of 422,769 bankruptcy cases, i.e. on average 19.9 bankruptcy announcements per issue, with a median of 14 announcements per issue. For each bankruptcy case we detected, we extracted the first 300 letters after the start of 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*, such that detecting commas as breakpoints helped structure the text. Using these comma-break points as general hints for where to look for certain infor-

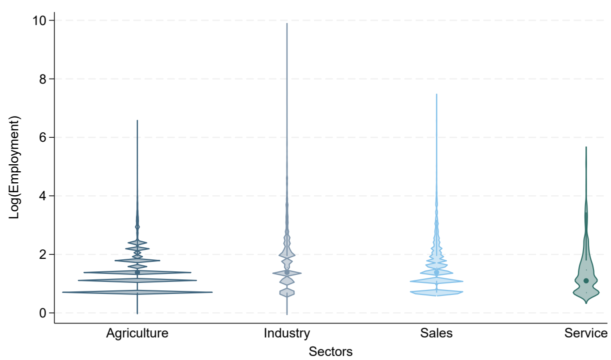
[10]The actual pattern switch occurred with the new bankruptcy act in the issue 22,564 from November 12th, 1861. While the overall pattern remained stable across announcements, the individual solicitors who published the announcements would vary the text pattern somewhat, e.g. using past tense ("was 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 good as all variations in our algorithm and went back to issues with an unusually low number of detected announcements to look for pattern variations that we might have overlooked before.

[11]For the short period when the announcements were published in a table format, we accessed the Google Vision API to detect the table structure accurately and directly transfer the relevant information into a digital table format.

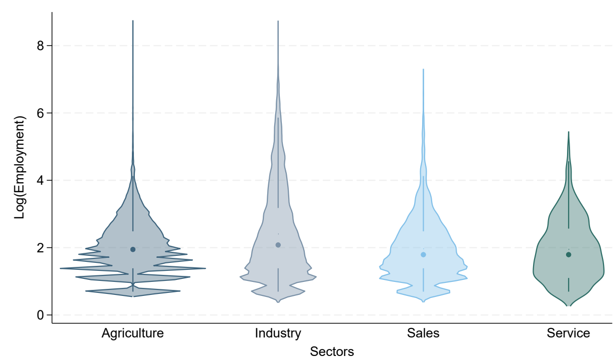
mation, we ran the specific text subsets against lists of city-, county-, borough-, and parish names as well as a list of (historical) census occupations respectively to detect matches.

Due to the occasionally bad quality of the scans, this required a lot of 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. For many locations, we were able to link them to the coordinates for a specific parish or city, some we could only geocode at the county level. To make use of the occupation titles, we assigned them to 6-digit historical international classification of occupations (HISCO) codes as defined by the International Institute of Social History Amsterdam.[12] Despite the pre-and post-processing steps, we were not able to acquire full information for all bankruptcy cases that our algorithm collected. Overall, we were able to geocode 373,555 bankruptcy cases (of this, 343,091 cases to the city- or parish level) and assign HISCO codes to 373,010 cases.

Figure B.1: Violin's plot – Firms' size by sector



(a) Violin plots of Firms' size by sector (1851)



(b) Violin plots of Firms' size by sector (1861)

B Appendix B – Supporting Evidence

Table B.1: Estimates by sector – Bankruptcies

	Dependent Variable: #Bankruptcies$_{i,t,s}$				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Agr_s	-0.22	-0.13	-0.25	-0.16	0.15
	(0.22)	(0.21)	(0.22)	(0.23)	(0.23)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s	0.46***	0.47***	0.46***	0.45***	0.45***
	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Trade_s	0.10	0.09	0.10	0.09	0.01
	(0.17)	(0.18)	(0.17)	(0.16)	(0.15)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Service_s	0.05	0.04	0.04	0.09	0.09
	(0.18)	(0.17)	(0.18)	(0.19)	(0.18)
Observations	7293	7293	7293	7293	7293
Pseu. R^2	.827	.827	.827	.827	.83
Sector FE	✓	✓	✓	✓	✓
Geo FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Sector $\text{Employment}_{i,s,t}$	✓	✓	✓	✓	✓
Coal_i $\times$ Manufacturing_s	✓	✓	✓	✓	✓
Port_i $\times$ Manufacturing_s	✓	✓	✓	✓	✓
London_i $\times$ Manufacturing_s	✓	✓	✓	✓	✓

Notes: Results from PPML regressions. Clustered Standard Errors in parentheses, * p< 0.1, ** p< 0.05, *** p< 0.01

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

Table B.2: Estimates by sector – Employment

	Dependent Variable: $\#\text{Employed}_{i,t,s}$				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Agr_s	-0.80***	-0.67***	-0.80***	-0.76***	-0.38***
	(0.08)	(0.11)	(0.08)	(0.06)	(0.05)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s	0.32***	0.21***	0.32***	0.35***	0.16***
	(0.04)	(0.03)	(0.04)	(0.04)	(0.04)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Trade_s	0.48***	0.53***	0.49***	0.40***	0.33***
	(0.06)	(0.06)	(0.06)	(0.05)	(0.05)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Service_s	0.12***	0.22***	0.13***	0.03	0.11***
	(0.05)	(0.04)	(0.04)	(0.03)	(0.03)
Observations	8496	8496	8496	8496	8496
Pseu. R^2	.908	.927	.912	.914	.943
Sector FE	✓	✓	✓	✓	✓
Geo FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
Sector $\text{Employment}_{i,s,t}$	✓	✓	✓	✓	✓
Coal_i × Manufacturing_s	✓	✓	✓	✓	✓
Port_i × Manufacturing_s	✓	✓	✓	✓	✓
London_i × Manufacturing_s	✓	✓	✓	✓	✓

Notes: Results from PPML regressions. Clustered Standard Errors in parentheses, * p< 0.1, ** p< 0.05, *** p< 0.01

C Appendix C – Robustness checks (Main estimates)

Table C.1: OLS estimations

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$						
	(1)	(2)	(3)	(4)	(5)		
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	0.27***	0.24***	-0.13***	-0.11***	-0.11***	-0.10***	-0.10***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) × Manufacturing_s		0.14***	0.14***	0.11***	0.11***	0.09***	0.06***
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	8185	8185	8185	8185	8185	8185	8185
Adj. R^2	.0452	.0537	.69	.708	.708	.709	.71
Sector FE	✓	✓	✓	✓	✓	✓	✓
Geo FE	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓
$\text{Population}_{i,t}$	✓	✓	✓	✓	✓	✓	✓
Coal_i × Manufacturing_s	✓	✓	✓	✓	✓	✓	✓
Port_i × Manufacturing_s	✓	✓	✓	✓	✓	✓	✓
London_i × Manufacturing_s	✓	✓	✓	✓	✓	✓	✓

Notes: Results from PPML regressions. Clustered Standard Errors in parentheses, * p< 0.1, ** p< 0.05, *** p< 0.01

Table C.2: Local shocks

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$			
	(1)	(2)	(3)	(4)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	0.00	0.05	0.02	-0.03
	(0.16)	(0.16)	(0.16)	(0.16)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s	0.43***	0.43***	0.44***	0.44***
	(0.08)	(0.08)	(0.09)	(0.08)
$\text{Unemployment}_{i,t}$	0.99**			1.09**
	(0.41)			(0.45)
$\text{Migrants}_{i,t}$		0.66		0.26
		(1.07)		(1.06)
Male $\text{pop}_{i,t}$			-0.26**	-0.28***
			(0.11)	(0.10)
Observations	7293	7293	7293	7293
Pseu. R^2	.828	.827	.828	.828
Sector FE	✓	✓	✓	✓
Geo FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
$\text{Population}_{i,t}$	✓	✓	✓	✓
Coal_i $\times$ Manufacturing_s	✓	✓	✓	✓
Port_i $\times$ Manufacturing_s	✓	✓	✓	✓
London_i $\times$ Manufacturing_s	✓	✓	✓	✓

Notes: Results from PPML regressions. Clustered Standard Errors in parentheses, * p< 0.1, ** p< 0.05, *** p< 0.01

Table C.3: Excluding Cells with low/high levels of populations

		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
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$)	-0.23	0.05	-0.23	0.10	-0.08
	(0.14)	(0.16)	(0.14)	(0.18)	(0.20)
$\mathbb{1}$ ($\text{Rail}_{i,t} > 0$) $\times$ Manufacturing_s	0.43***	0.43***	0.43***	0.46***	0.47***
	(0.08)	(0.08)	(0.08)	(0.10)	(0.10)
Observations	7025	7284	7016	5841	5736
Pseu. R^2	.744	.827	.744	.765	.608
Sector FE	✓	✓	✓	✓	✓
Geo FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
$\text{Population}_{i,t}$	✓	✓	✓	✓	✓
Coal_i $\times$ Manufacturing_s	✓	✓	✓	✓	✓
Port_i $\times$ Manufacturing_s	✓	✓	✓	✓	✓
London_i $\times$ Manufacturing_s	✓	✓	✓	✓	✓

Notes: Results from PPML regressions. Clustered Standard Errors in parentheses, * p< 0.1, ** p< 0.05, *** p< 0.01

Table C.4: Robustness – Extensive margin instead of intensive margin

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\text{Log}(\text{Rail}_{i,t})$	1.43***	1.40***	-0.26**	-0.27**	-0.27**	-0.27**	-0.27*	-0.30**
	(0.10)	(0.10)	(0.13)	(0.14)	(0.14)	(0.14)	(0.14)	(0.15)
$\text{Log}(\text{Rail}_{i,t}) \times \text{Manufacturing}_s$		0.06***	0.07***	0.07***	0.07***	0.07***	0.08***	0.15***
		(0.02)	(0.03)	(0.02)	(0.02)	(0.02)	(0.03)	(0.04)
Observations	8185	8185	7293	7293	7293	7293	7293	7293
Pseu. R^2	.31	.318	.827	.828	.828	.828	.828	.828
Sector FE			✓	✓	✓	✓	✓	✓
Geo FE			✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
$\text{Sector Employment}_{i,s,t}$				✓	✓	✓	✓	✓
$\text{Coal}_i \times \text{Manufacturing}_s$					✓			✓
$\text{Port}_i \times \text{Manufacturing}_s$						✓		✓
$\text{London}_i \times \text{Manufacturing}_s$							✓	✓

Notes: Results from PPML regressions. Clustered Standard Errors in parentheses, * p< 0.1, ** p< 0.05, *** p< 0.01

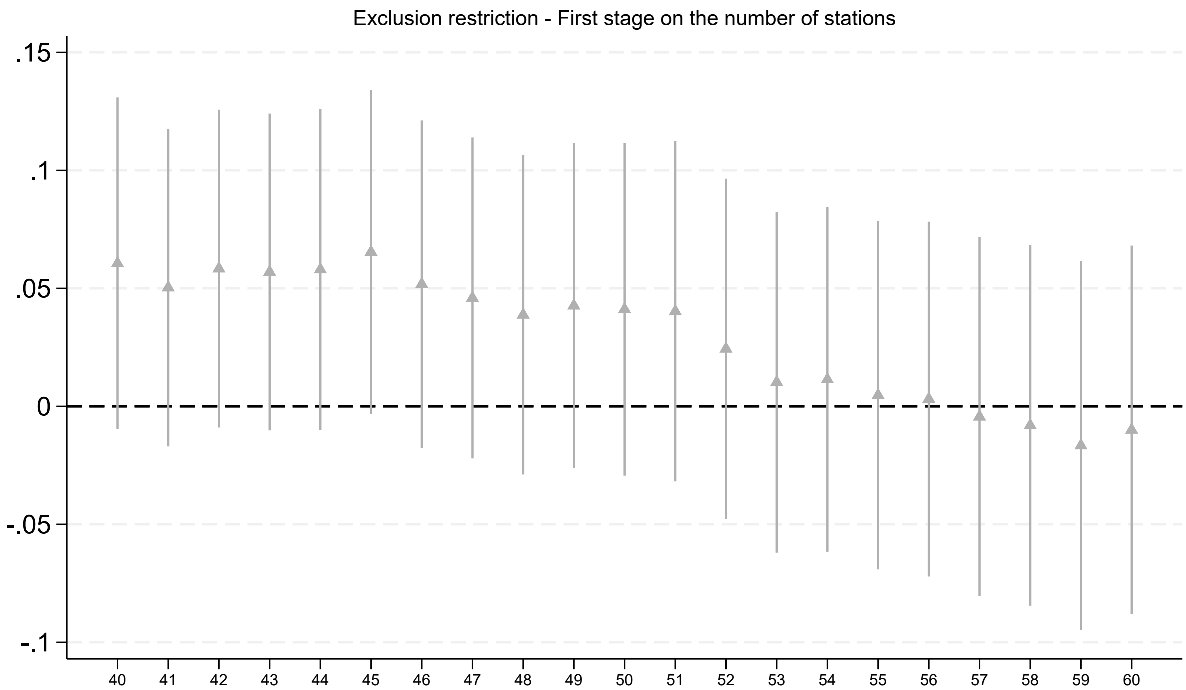
Table C.5: Excluding Reform years

	Dependent Variable: $\#\text{Bankruptcies}_{i,t,s}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\mathbb{1}\ (\text{Rail}_{i,t} > 0)$	2.45***	2.33***	0.04	0.04	0.05	0.04	0.01	-0.00
	(0.23)	(0.24)	(0.23)	(0.23)	(0.23)	(0.23)	(0.22)	(0.22)
$\mathbb{1}\ (\text{Rail}_{i,t} > 0) \times \text{Manufacturing}_s$		0.46***	0.46***	0.46***	0.40***	0.45***	0.54***	0.57***
		(0.12)	(0.12)	(0.12)	(0.12)	(0.11)	(0.11)	(0.10)
Observations	8185	8185	6473	6473	6473	6473	6473	6473
Pseu. R^2	.0595	.0626	.701	.701	.701	.702	.701	.702
Sector FE			✓	✓	✓	✓	✓	✓
Geo FE			✓	✓	✓	✓	✓	✓
Year FE			✓	✓	✓	✓	✓	✓
$\text{Sector Employment}_{i,s,t}$				✓	✓	✓	✓	✓
$\text{Coal}_i \times \text{Manufacturing}_s$					✓			✓
$\text{Port}_i \times \text{Manufacturing}_s$						✓		✓
$\text{London}_i \times \text{Manufacturing}_s$							✓	✓

Notes: Results from PPML regressions. Clustered Standard Errors in parentheses, * p< 0.1, ** p< 0.05, *** p< 0.01

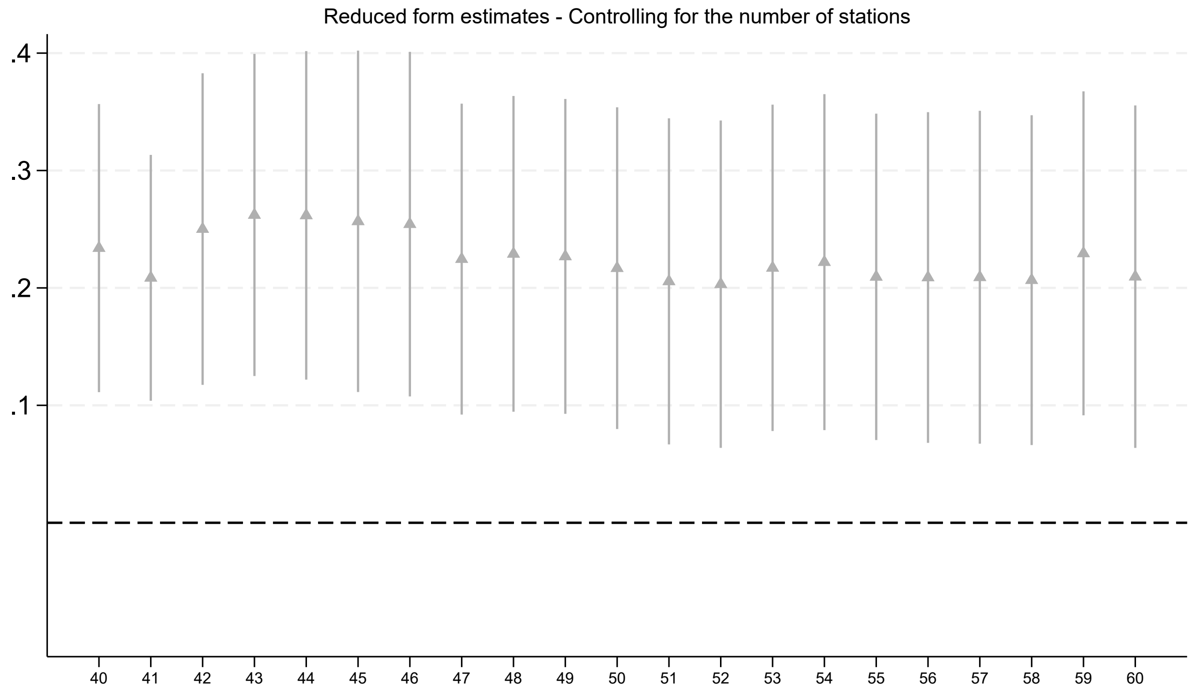
D Appendix D – Robustness (Identification)

Figure D.1: Placebo First Stage - Number of stations



*Notes:*Results of the estimation of equation 2 with the dependent variable being the number of stations in a gridcell. Coefficients of the instrument for different values of the percentile considered in the extension of the network (X= 40 to 60th percentile). Standard errors are clustered at the location level. Lines depict 95% confidence intervals.

Figure D.2: Reduced form while controlling for the number of stations



Notes: Coefficient of reduced form estimations using our instrument in equation 1 instead of the rail measure. These estimations also add the number of stations as a control variable. The instrument is constructed using the percentile in the distance of stations to the 1851 LCP. Standard errors are clustered at the location level. 95% confidence intervals are shown.