The Effect of Transportation Infrastructure on Innovation: The Role of Market Access in the English Railway Boom*

Giorgio Ravalli

Latest Version

Abstract

What is the impact of transportation infrastructure on innovation? I study the historically unprecedented boom of the railway network in 19th Century England, which reduced average travel time between London and Birmingham by 80% between 1830 and 1911 and the average freight cost between any two districts in England between 1840-1861 by over 60%. Using "accidentally connected" locations I show that building a rail station caused patents per capita to increase by over 75% in a district. I find no evidence that this effect was due to the relocation of existing inventors. The positive railroad effect is driven by an increase in market access (a demand-induced innovation effect). Across 1823-1861 median market access approximately doubles; a doubling of market access increases innovation by over 50%. This mechanism is economically and statistically significant. I also find that the railroad induced greater knowledge flows through improved communication between inventors; this effect appears distinct from the market access channel.

Keywords: Innovation, Trade, Industrial Revolution, Transportation, Knowledge Flows
JEL codes: N73 O18 O31 O33 R40 F14

*I thank John Van Reenen and Jeremiah Dittmar for their guidance as well as Dave Donaldson for many helpful comments. I also thank Guy Michaels, Alan Manning, Steve Pischke, Daniel Sturm, Gaia Dossi, Thomas Sampson, Peter Lambert, Abhijit Tagade, and other CEP seminar participants as well as Tim Besley, Walker Hanlon, Donald Davis, and Jane Olmstead-Rumsey. For generously sharing their data, I thank Gaia Dossi and Davide Coluccia.

1 Introduction

I study how the construction of the British railway network increased innovation during the 19th century. The railway was one of the most transformative projects in history, revolutionizing domestic trade; its construction caused the average freight cost between any two districts in England to decline by a factor of 2.5 within two decades.[1] The impact of this transportation revolution is illustrated by the experience of James Nasmyth, an inventor and entrepreneur who is listed as an applicant for at least 16 patents related to his machine tools manufacturing firm in Salford, England. In his autobiography, Nasmyth cites the railroad's impact on his firm:

> *[T]he railway alongside enabled a communication to be kept up by rail with every part of the country ... After the machines or engines had been finished, it was the business of the same workmen to remove them from the workshops to the railway-siding alongside the foundry... (Nasmyth 1897)*

The railroad created new markets for Nasmyth; the great reduction in freight costs increased demand for his firm's goods. The effect of market demand on innovation is at the heart of a major strand of the economic growth literature. Innovation has long been believed to be demand-driven (Schmookler 1966) and from a firm's perspective, an increase in demand allows fixed cost R&D investments to be spread over a larger customer base (Sokoloff 1988). While prior work has established a positive relationship between market size and innovation (Acemoglu and Linn 2004; Aghion et al. 2024), the role of transportation networks in driving this dynamic remains unexplored, despite a nascent literature on how these networks affect innovation through other channels (Agrawal et al. 2017). This paper addresses that gap by examining the relationship between railway expansion and innovation in 19th-century England from a static trade perspective.

I present two main findings in this paper. First, I show that the establishment of a rail station causes a significant increase in local patenting activity. Second, I identify market access as a key mechanism to explain these returns to patenting.

Districts in England that receive a rail station experience a 77% increase in patent applications per capita between 1823-1861. By comparison, the effect on population is a 3% increase. This effect is persistent across several decades and robust to a wide range of sample restrictions and estimation procedures. Further, there is no differential impact by patent quality; eliminating patents in the bottom 75% by quality between 1823-1850 using the Nuvolari et al. (2021) Bibliographic Composite Index, which assesses the historical impact of patents, does not substantially change the results.

I follow the transportation network literature to deal with selection bias with respect to the rail treatment (Redding and Turner 2015). To focus on plausibly exogenous variation, I restrict the data to "accidentally connected" locations - districts between the major economic centers whose connection was the primary focus of the transportation projects. The "accidentally connected" locations are plausibly exogenously treated in that they are connected to the rail network because they happen to lie on a convenient path between cities. To further ensure exogeneity, I create a lowest cost path algorithm between districts and match this to the actual transportation network. If an "accidentally connected" town lies on the lowest cost path between cities it is reasonable to assume that the town was only connected by chance because of its location between places that were deemed economically important to the rail firms.[2] For the majority of the analyses in this

[1] I calculate this number in Section 5.

[2] See Chandra and Thompson (2000), Faber 2014, Hornung (2015), Banerjee et al. (2020), Bogart et al. 2022, and Andersson et al. 2023 for examples which incorporate some or all elements of this methodology.

paper, I restrict the treatment group to the districts which are both "accidentally connected" and which lie on the least cost path network, though I explore more alternate treatment definitions as robustness tests.

A rail network increases the size of the market to which firms can sell their goods by significantly lowering freight costs between districts. This expansion of market access boosts demand and enables firms to undertake costly R&D by enlarging their potential customer base. Notably, a third of the rail companies' revenue in this period came from freight, underscoring the role of railroads in changing domestic trade (Railway Commission 1848). I provide empirical evidence that this increase in market access is a key mechanism through which rail stations stimulate local innovation.

I estimate the effect that changes in market demand, driven by the railway network, had on local innovation. To do so, I create a transportation map of 19th century England using railroads, inland navigable waterways, and overland routes to measure each district's annual change in market access, where a district's market access is the sum across all other districts of the product of local GDP and the transportation cost between district pairs (following Donaldson and Hornbeck 2016). Annual changes in market access are determined by any railway construction which decreases the travel cost between any two districts. This means that rail construction does not necessarily need to be local to a district to influence its market access and changes in market access from "distant" rail construction is plausibly exogenous. Indeed, controlling for local rail construction does not greatly affect the market access coefficient estimate indicating that distant rail construction causes much of the increase in local patenting. The setting provides a relatively clean measure of changes to market access due to rail as there was no concurrent major expansion of infrastructure since the majority of the turnpike and canal networks were built in the previous century.

I find that a doubling of market access - approximately the median increase between 1823-1861 - causes an increase in local patenting per capita of 51%. Including the local rail station indicator in the regression and observing the change in magnitude of the coefficient implies that over 40% of the local station effect on patenting is explained by the increase in market access that a station brings. I also find that the local rail effect is driven by districts which are in the top quintile by the change in market access between 1823-1861. These districts have low market access in 1823, indicating that there are greater gains in connecting places with poor transportation infrastructure prior to the rail network. A back of the envelope calculation finds that the change in market access caused a 49% increase in aggregate patenting and a 5% increase in national TFP.

I exploit sector heterogeneity to empirically test the theory that greater market access causes an increase in innovation because firms which operate in small markets are unable to spread high fixed cost R&D expenditures to a large enough customer base and that a larger market allows for precisely this type of R&D spending (Sokoloff 1988). I examine sectors which I identify as having high fixed, but low variable costs or sectors which produce niche products. The latter gain from the rail network because they can ship their goods further and customers typically have to travel a greater distance to purchase a niche product compared with non-niche products (Brynjolfsson et al. 2003). In both cases I find that regressing patents on the interaction between market access and these sectors' labor shares produces large and statistically significant estimates, implying that these sectors are driving the returns to patenting. Additionally, market access increases concurrently with districts' labor share in these sectors, suggesting that firms in these sectors are benefiting from increased demand and increasing expenditures. I then use an input-output table for 1841 England created by Horrell et al. (1994) and disaggregate market access to the district-sector-year level and find that within sector patenting increases by 40%. Using an input-output table means that I am beholden to the sectors used in that table and identifying high fixed cost or niche product sectors is not viable as the

sectors are too broadly defined. I instead adopt the manufacturing definition used by Horrell et al. (1994) and find that market access increases patenting activity most in manufacturing sectors and in districts with a high share of manufacturing labor, in line with expectations that firms which produce goods benefit more from increased trade through the rail network than firms which produce services. Additionally, the returns to within sector patenting from an increase in market access are greatest for district-sectors with a high number of workers operating in that sector, implying that a district which does not produce (for example) textile goods does not gain textile patents because of an increase in market access.

It is possible that the market access effect I find is actually due to an increase in knowledge flows; previous papers have shown that transportation networks induce innovation by fostering knowledge flows between inventors by facilitating communication (Agrawal et al. 2017; Pauly and Stipanicic 2024). I adapt my market access variable to create a "knowledge flows" variable which measures a district's exposure to other districts' patents by substituting the stock of patents in a sector in place of local GDP.[3] The knowledge flows variable shows whether easier transportation between a district and a location with many inventors in the same field (rather than a large market) causes innovation in that district. By using within sector knowledge flows I mimic patent citations which, following Jaffe et al. 1993 are widely used in the literature to measure knowledge spillovers and which were not recorded during the period I study.

A doubling of this knowledge flows measure is associated with a 43% increase in within sector patenting per capita in a district. Large towns are disproportionately affected. Intuitively, a large enough existing knowledge base is a prerequisite for a district to benefit from incoming knowledge flows as there is a higher chance that a resident possesses the relevant knowledge to benefit from the incoming flows. Interacting the knowledge flows variable with the top quintile of workers in a district and sector yields a positive coefficient, indicating that most of the returns are in districts and sectors with a high number of workers with relevant expertise.

The market access and knowledge flows channels are distinct. Different districts receive the largest returns to patenting from the two measures: knowledge flows have a disproportionate effect on large towns and market access on districts with a high share of manufacturing labor. The reverse is not true, indicating that different districts are driving the effects. Though the variables exhibit high serial correlation that makes direct comparisons difficult, I find no correlation between the variables when comparing their change from the beginning of the period to the end. Additionally, the interaction between knowledge flows and starting point market access is insignificant (as is the reverse case), further pointing to a lack of relationship between the channels.

I do not find empirical evidence that the increase in patenting is due to a displacement of innovation caused by the railroads. The patenting rate of "not yet treated" and "never treated" districts does not significantly decrease after a district within 50KM receives a station, indicating that there is no outflow of patenting from the not treated into the treated districts. Similarly, the mean number of patents in the districts neighboring a treated district does not decrease after that district's treatment (and before the neighbors receive treatment). Using the 1851 Census, I proxy inventors (who are not identified as such in the Census) with highly skilled individuals, a group that includes entrepreneurs and engineers. Conditional on moving from their district at birth, these individuals are no more likely to move to districts with a rail station than are others. I also construct a structural model in the style of Kline and Moretti (2014) to explain changes to aggregate productivity following a large scale infrastructure investment and the results

[3]This measure is similar to what Pauly and Stipanicic (2024) use in measuring knowledge flows during the American Jet Age.

indicate that the productivity returns from railways do not cancel out at the aggregate level.

This paper builds on multiple strands of economic literature. First, it contributes to the literature on the impact of market size on innovation (Acemoglu and Linn 2004). Exporter firms have been shown to be more productive (Melitz and Redding 2014), with several papers finding evidence that access to new export markets causes firms to increase their productivity either through direct innovation or through investing in better technology.[4] In a similar vein, Aghion et al. (2023) and Aghion et al. (2024) find that French firms invest in innovation after a positive shock to market size. Related to this literature is research focusing on the impact of an influx of imports on local firm productivity and innovation (Bombardini et al. 2017; Bloom et al. 2016; Autor et al. 2020). My research is closely related to both the export and import strands of the literature but operates through a domestic trade channel rather than through an international one. This distinction has implications for the findings. Transportation networks facilitate trade because of a new technology or investment in infrastructure rather than tariff reduction, offering policymakers a distinct avenue for increasing productivity. Little empirical work has been done on how new infrastructure, and specifically, transportation networks, can cause such a demand-driven increase in innovation (Juhász and Steinwender 2023).[5] I use the definition of "market access" developed by Redding and Venables (2004) and Donaldson and Hornbeck (2016) to measure the rail network's returns to innovation that manifest through a trade channel and find strong, robust effects. This is the first paper to create an annual map of market access in England during this time period, which could be useful for future research on the rail network linking trade with fast changing dynamics, such as firm productivity.[6]

Second, this paper contributes to the literature on knowledge flows and on how increased communication between individuals leads to higher levels of innovation (Duranton and Puga 2004; Glaeser and Gottlieb 2009; Hanlon and Miscio 2017; Davis and Dingel 2019; Buera and Oberfield 2020). Specifically, this paper finds evidence that a decrease in communication costs fosters an increase in knowledge flows which causes an increase in within-sector patenting. In doing so, it builds on a nascent literature which spans a variety of settings: Agrawal et al. (2017) use the construction of highways in the 1980s in United States; Catalini et al. (2020) and Pauly and Stipanicic (2024) use changes in the American airline industry; and Hanlon et al. (2022) study the impact of decreased mail costs in 19th century England. This paper is the first to find direct evidence that knowledge flows fostered by a rail network increase innovation (Juhász and Steinwender 2023), though there is a growing literature on the effect of the railroad to diffuse knowledge.[7]

Third, this paper contributes to the literature regarding transportation infrastructure's general economic effects. Several papers have studied the impact of transportation networks on productivity (Donaldson and Hornbeck 2016; Donaldson 2018; Faber 2014; Baum-Snow et al. 2020). This paper is distinct in that it focuses more specifically on innovation. Transportation infrastructure has been found to increase local population (Gregory and Marti 2010; Bogart et al. 2022; Büchel and Kyburz 2020) and urbanization (Atack et al. 2010; Hornung 2015).[8] The literature also finds firm level effects: Atack et al. (2011) and Hornung (2015) find evidence that transportation networks are associated with increased firm size, possibly due to their role in increasing competition through market expansion, while other papers find evidence of higher

[4]For example: Lileeva and Trefler (2010), Damijan et al. (2010), Aw et al. (2011), and Bustos (2011).

[5]The only attempt at studying this question is Perlman (2017) which finds an increase in patenting following an increase in market access, but that this increase in mainly driven by changes in local rail infrastructure rather than construction of the entire rail network.

[6]While You et al. (2021) build a comprehensive, detailed map of market access in England, they only observe the years 1830 and 1911.

[7]See, for example, Andersson et al. (2023), Yamasaki 2017, and Américo 2022, which study the Swedish, Japanese, and Brazilian railroad networks respectively.

[8]Transportation networks have also been found to increase suburbanization (Baum-Snow 2007; Garcia-López et al. 2015).

firm productivity (Ghani et al. 2016; Gibbons et al. 2019; Holl 2016; Tang 2014). Michaels (2008) finds that highway connections increased demand for high skilled labor. In finding that the railway system provided large returns to patenting, this paper contributes to the literature which finds positive returns to transportation infrastructure. The trade channel I find could also be useful to more comprehensively measure the benefits of such transportation systems through a trade framework. Jaworski et al. (2023) estimate that removing the entire US highway system would decrease US GDP between 2.6%-3.6%, with a quarter of this figure due to higher trade costs; pinning down the returns to patenting could increase precision in these types of estimates.

The structure of the paper is as follows: Section 2 discusses the data used in this paper; Section 3 describes the identification strategy; Section 4 presents the main results; Section 5 discusses the potential mechanisms with a particular focus on the market access channel; Section 6 describes various robustness tests on the main results; Section 7 concludes. The Appendix includes additional tables, maps, and graphs as well as discussion of a model on aggregate productivity effects that follows Kline and Moretti (2014).

2 Data

2.1 Patents

I use patent applications as a proxy for innovation. I digitize patent data from Woodcroft (1854), which details all patent applications from 1823-1852. Each application includes the application date and a description of the patent, as well as the names, occupations, and addresses of the inventors. I exclude all patents which are filed from outside of England or which are filed by patent agents as the stated address corresponds with the agent or the Patent Office, rather than with the inventor.[9] Rail patents are also excluded from most analyses; they are identified by key words in the patent description (e.g. "rail" or "locomotive"). This is perhaps overly restrictive as it removes patents that are only broadly related to rail, such as engines that can be used for stationary or locomotive power, but it ensures that rail patents cannot enter into the analysis. I include these patents as a robustness check in Section 6 and find that they do not meaningfully alter the results.

I use the addresses from the patent data to match individuals to a geographic dataset (Satchell et al. 2023) which includes a shapefile denoting the boundaries of places at various geographic levels in England in 1851.[10] I aggregate to the district level (the second largest below county) as many of the addresses outside of London include only broad geographic areas.[11] There are 575 districts in England. Matching at the district level also allows me to use population data from the decadal Censuses prior to 1851, as the locations in these Censuses were at a district or union level. The mean area of a district is 86.97 KM^2.[12]

I supplement the patent applications with data collected and shared by Coluccia and Dossi (2023) which extends the time field through to 1891. I use this to extend the main analysis to 1861 and to show the persistence of the rail effect until 1891. Due to limits in the rail dataset, the main results are limited to the 1823-1861 time period.

The number of patent applications in a given district and year is low, with a mean of 1.13 patents. The data are even more sparse in the 1823-1852 period; there are 6,875 total patents across 575 districts and 30

[9] These are identified in the patent descriptions as "communications".
[10] The same patent can be matched to multiple locations if the applications listed multiple inventors.
[11] I am able to match 87.3% of the 1823-1852 patents in England in this way.
[12] By comparison, modern day Greater London is 1,572 KM^2

years (for a mean of .40 patents per district and year). The number of patents is disproportionately high in the cities. The 33 districts which comprise London and Birmingham account for 47% of these 6,875 patents. The Patent Law Amendment Act of 1852 greatly reduced the cost of filing a patent, and there was a large increase in applications following its implementation. I control for this by always including a full set of time dummies and further robustness tests present no evidence that the Patent Law affects my results. Most analyses in this paper utilize a Poisson regression model to deal with the sparsity and the accompanying non-normality of residuals in an OLS model.

I categorize each patent into one of 55 sectors.[13] For the 1823-1852 data, I use the patent description field to do so. The post 1852 data provided by Coluccia and Dossi (2023) contains a CPC industry field for the patents and I match these to the 55 sectors to have a uniform categorization across all years.

2.2 Rail

The rail data are obtained from multiple sources. First, I use the shapefiles provided by Henneberg et al. (2018a), Henneberg et al. (2018b), Henneberg et al. (2018c), and Satchell et al. (2018), to obtain precise rail line and station locations. These datasets are a snapshot of the stations and lines that existed in 1851 and 1861. In order to obtain precise opening years, I match the stations and lines in the shapefiles with Cobb (2015), which contains detailed maps of England with information on rail lines and stations' opening dates. I then match the rail lines to the districts contained in the geographic shapefile provided by (Satchell et al. 2023). The result is a shapefile that shows the exact year in which each district first received a rail line or station from 1825-1861.

I define a district as having a rail station if there is a station within the district's geographic bounds. Additionally, districts with a total area under 28.27KM^2 (equivalent to a radius of 3KM for a perfect circle) are considered to have access if there is a station within 3KM of the district's geographic centroid. This is to ensure that districts that are small in area are not excluded when the inhabitants of the district could reasonably walk to a nearby station.

Though the first rail station opened in 1825, the opening of the Liverpool and Manchester Railway in 1830 is considered the birth of the Rail Era in England as it was the first inter-city railway and the first to rely exclusively on steam power. Rapid expansion occurred in the 1840s in a period known as Rail Mania, during which there was high speculative interest in rail.

2.3 Other Data

I use the Natural England (2023) dataset, which is a shapefile that divides England into 250m x 250m squares with a land description of each square, to create simulated rail lines that act as a low cost path network between locations to address endogeneity concerns regarding placement of the actual rail lines. I explain the methodology in Section 3.

For years prior to 1851, I use decadal district level population data from the Census of Great Britain (1852). As the districts belonging to Gloucestershire and Staffordshire counties are missing from the available scans of some years, I remove all districts from these counties from the analysis, decreasing the total count of districts from 575 to 542. The first "modern" census was taken in 1851, and from 1851-1891 I use census microdata from (Schurer and Higgs 2023) to obtain more accurate decadal population estimates, occupations, and birth districts.[14]

[13] The full list of sectors can be found in the Appendix.

[14] The 1871 Census is mostly destroyed, leaving only data from the 1851, 1861, 1881, and 1891 Censuses.

Lastly, I use the return of gross annual value of property and profits measured through tax schedules for every parish in England in 1815, 1843, and 1860.[15] Scanned copies of the returns are held by the British Parliamentary Papers. Unlike the Poor Law land value measurements, the returns were collected by qualified officials sent from the House of Commons, making them less likely to be subject to the distorted values recorded in the Poor Law data.[16] The values are at the parish or union level and are aggregated or approximated to the district level using the Satchell et al. 2023 geographic dataset.[17]

3 Identification

A key aim of this paper is to identify the causal impact of a local rail station on innovation. The main econometric challenge is that rail stations are unlikely to be randomly assigned. This issue is endemic to the transportation network literature and I employ a common strategy referred to as the "inconsequential place approach" to deal with it (Redding and Turner 2015). The core concept is that rail lines are built to connect two places of interest. Districts which lay in between them can be considered "inconsequential" in that they are only connected to the network because the rail line has to run through them in order to connect the important districts. I present here the main empirical Poisson model used in the paper to describe the strategy in greater detail:

$$E(Patents_{i,t}) = exp(\alpha_i + \lambda_t + \beta D_{i,t}) \tag{1}$$

$D_{i,t}$ is an indicator variable which equals one for any year t in which a district i has a local rail station and zero otherwise and $Patents_{i,t}$ is the number of local patents per capita. α_i and λ_t are respectively district and time fixed effects.

The first step is to identify districts which rail firms likely sought to connect. A firm is likely to build a greater proportion of its stations in cities to maximize its potential customer base; population is therefore likely a key, observable determinant in station placement and I control for the imputed population by measuring innovation as patents per capita. I also remove all districts that are included as part of London and Birmingham, the two largest economic centers at this time which rail firms clearly sought to connect.[18] Additionally, the inclusion of year and district fixed effects in equation (1) allows me to control for any unobservable characteristics that rail firms might have used to make placement decisions.

Even with the inclusion of fixed effects and controlling for population, there is still the potential that growth trends of unobservable variables might differ. Districts which were expected to experience faster economic growth, which may be correlated with faster patent growth, would have been more profitable to the rail firms. If rail firms could correctly identify which districts were expected to grow faster, the treatment would not be randomly assigned, even with the inclusion of the fixed effects controls. To that end, I remove from the analysis all districts which were plausibly chosen by rail firms as important to connect and only retain districts which appear to have received rail connections "accidentally." The districts which are retained are those that lie between the "important" districts and are therefore only connected to the rail network because they happen to lie along the path that connects the "important" districts.

I designate a district as "important" - and therefore remove it from the analysis - if a rail firm clearly

[15] British Parliamentary Papers (1854a), British Parliamentary Papers (1854b), and British Parliamentary Papers (1860)

[16] See Purdy (1860) for a comparison of the two.

[17] Parish names do not always perfectly match with the parish names in Schurer and Higgs (2023). When using this dataset, I drop all districts for which the parishes that constitute them do not contain any matched data.

[18] Indeed, one of the first prominent rail firms was the London and Birmingham Railway.

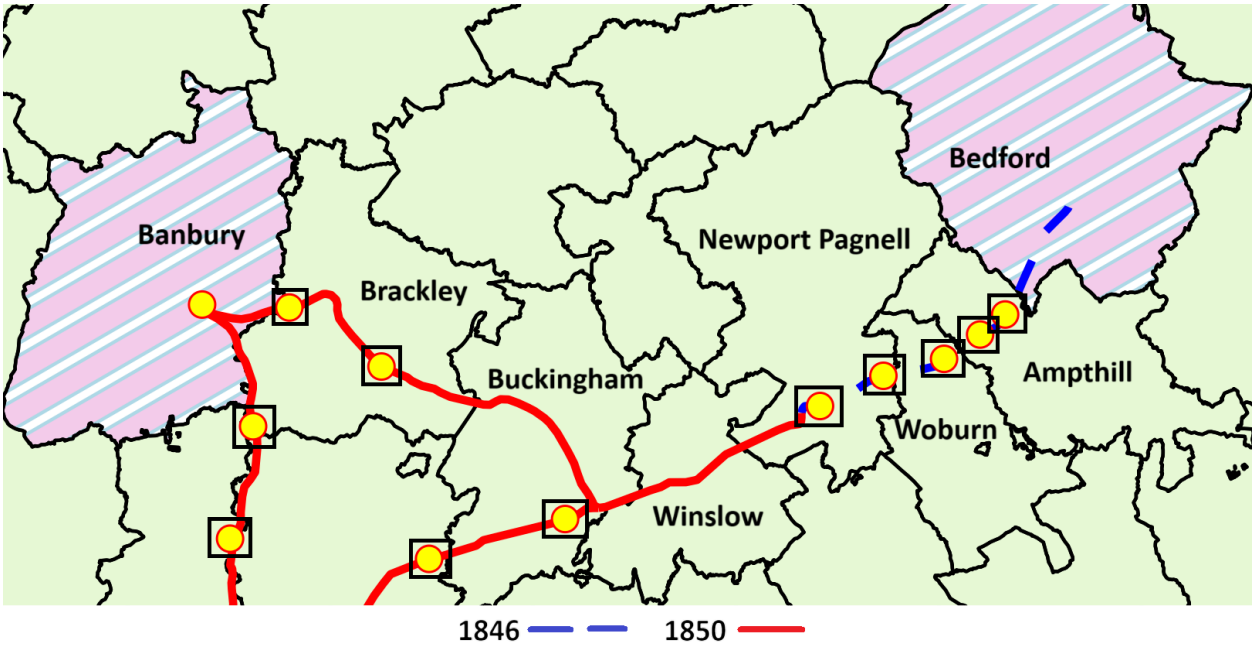


Figure 1: Example of a rail line that was opened in 1846 and 1850 to demonstrate how the start and end points of rail lines are identified. The opening years of the rail line segments are denoted at the bottom. Banbury and Bedford house line end points and so are excluded from the main treatment definition. Brackley, Buckingham, Winslow, Newport Pagnell, Woburn, and Ampthill all contain stations (circles marked by squares) and lie between the Bedford and Banbury end points and are therefore retained and considered treated.

sought to connect it to the network. I establish this intent using rail line placement. Any district which contains a rail line that is not connected on both ends is a clear start or end point of a rail line and therefore a district that a rail firm determined important enough to connect to the network. Given the evolving nature of the rail network, it is possible that these points served as temporary rail termini while the rest of the line was still under construction. To address this, only points which see no construction for the five years following their opening date are considered end points. Figure 1 shows a segment of the rail network in 1850. A line running from Bedford to Banbury was opened between 1846 and 1850. As the Bedford and Banbury districts house the ends of the rail line, they are removed; Brackley, Buckingham[19], Winslow, Newport Pagnell, Woburn, and Ampthill lie between these end points and have rail stations and are therefore considered treated districts.

It is possible that firms constructed lines to certain locations but then continued the line onwards. In such a case, there is no terminus and the district would be retained. To remedy this, I also remove all districts which are in the top quintile of rail kilometers per square kilometer of the district's total area in 1851. Districts with high rail density likely had multiple lines passing through it. I view this as an indication of intention for a particular destination by the rail firms because they are either competing to serve a district or using a location as a rail hub to connect otherwise separate lines.

While this methodology should in theory remove districts which were potentially identified by rail firms as exhibiting high expected economic growth, the possibility remains that the remaining treated districts were chosen in a non-random manner. These remaining districts may not have been chosen simply because they

[19]The station in Buckingham is not visible.

happened to lie between two of the "important" districts that were removed in the first step, but because the rail companies actively sought to connect them to the network as they were themselves "important." For example, a rail firm might have sought to connect London and Birmingham, but deviated from the straight line path between them in order to connect Oxford.

To address this, I follow the recent literature (Banerjee et al. 2020; Faber 2014; Atack et al. 2010) and I create a network of simulated rail lines which connect every station in England. The algorithm connects points with a straight line that moves along a grid of 250 x 250 meter squares of land identified and described by the Natural England (2023) dataset. The constructed line can only deviate from the straight line path to avoid a square that is defined as "unimproved land." This simulates the notion that rail firms chose to connect the "important" districts in the cheapest manner possible (by connecting them with straight lines and avoiding areas that required additional construction costs). Each simulated line segment connects only two points at a time. To preserve the time element of the dataset, two points can only be paired if they are in the same year. For example, in Figure 1 the 1846 line segment between Bedford and Newport Pagnell and the 1850 segment between Newport Pagnell and Banbury are simulated separately.

To avoid choosing the exact points on the actual line, I simulate each line five times, using five random points near the real start and end point of each line segment. For example, to simulate the Bedford-Newport Pagnell line segment, I choose a random point within a 1.1KM radius of the line end point in Bedford and a random point within a 1.1KM radius of the end point in Newport Pagnell. I repeat this four more times, and then choose the shortest simulated line to add an additional "cost saving" element to the process. Figure 2

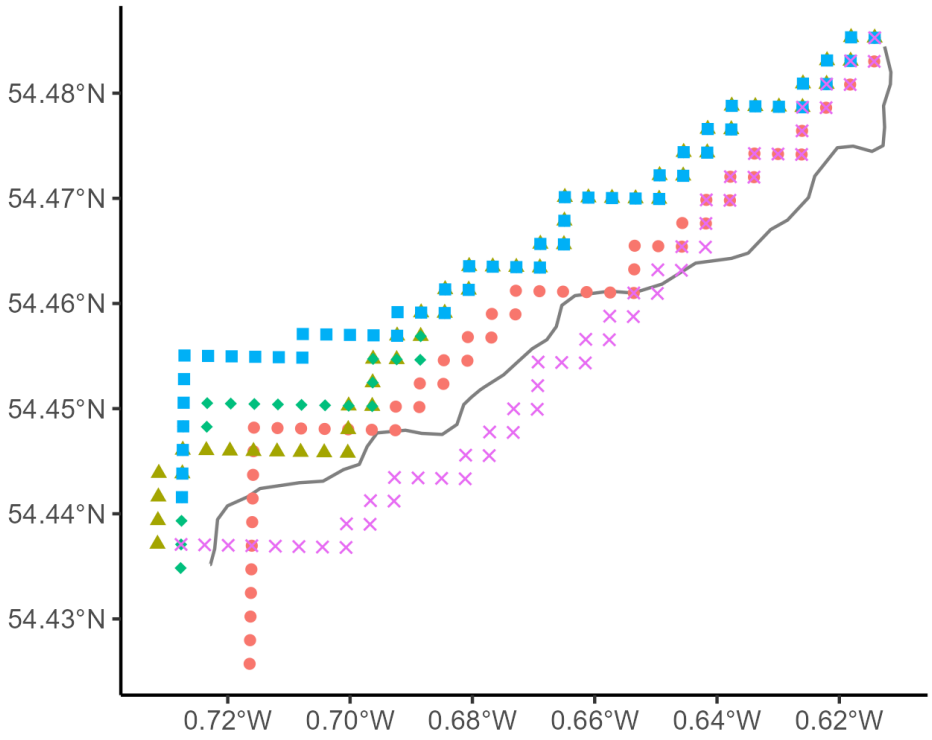


Figure 2: An example of the simulated lines. Shapes represent the paths of the simulated lines. Each shape/color represents one of the five simulated rail lines. The black line is the actual line they are simulating. The line composed of blue squares is the shortest line and is the one that is chosen to compare against the actual line.

Table 1: Summary Statistics Across Treatment Definitions

A. All Districts	Treated	Pre Treatment	Post Treatment	Never Treated
Mean Population	28,386.12	24,593.03	31,602.12	23,912.81
Mean Population Density	3,200.73	2,723.52	3,535.86	23,607.29
Mean Patents per District per Year	1.13	0.23	2.34	2.54
Total Districts	409	409	409	70
B. Ends and Hubs Removed	Treated	Pre Treatment	Post Treatment	Never Treated
Mean Population	18,884.56	17,927.07	18,884.56	14,897.61
Mean Population Density	220.72	209.14	220.72	971.98
Mean Patents per District per Year	0.26	0.08	0.26	0.13
Total Districts	176	176	176	42
C. Main Treatment Definition	Treated	Pre Treatment	Post Treatment	Never Treated
Mean Population	19,066.35	18,092.05	20,415.37	14,897.61
Mean Population Density	224.38	212.15	243.49	971.98
Mean Patents per District per Year	0.28	0.09	0.60	0.13
Total Districts	134	134	134	42

Notes: Summary statistics for the three treatment definitions discussed in Section 3. Panel A includes all districts. Panel B excludes treated districts which contain rail line end points and rail hubs as well as all districts which comprise London and Birmingham. Panel C uses the main treatment definition, which further restricts the previous definition to districts which are treated in the same year both in reality and according to the simulated lines network. The number of "Never Treated" districts decreases from the top panel due to 12 districts which house the end of a rail line but never receive a station and districts which comprise London or Birmingham. Districts which have zero patents between 1823-1861 are excluded as they are excluded from most regressions in this paper due to the use of district and year fixed effects.

shows an actual rail line in black and its five associated simulated lines. The line composed of blue squares is the shortest and therefore the line against which the black line is compared.

Once the shortest simulated line is chosen, I match the simulated line network with the real rail line network. Only districts which are treated in the same year in both the real and simulated networks are retained for the analysis. This is to ensure that the districts that remain in the analysis are only districts whose rail lines fall along the geographic path of least resistance in terms of rail construction. The group of remaining districts comprise the "main" treatment definition I use for most analyses in this paper.

Table 1 presents summary statistics for the three treatment definitions discussed above. These definitions are: "All Districts," which includes all treated districts; "Ends and Hubs Removed," which removes treated districts which house rail line end points and rail hubs as well as all districts which comprise London and Birmingham; and the main definition which excludes all treated districts that the previous definitions exclude in addition to districts whose simulated line network year of treatment differs from the actual treatment year. Each of the treatment definitions shows a large increase in the pre-treatment to the post-treatment mean number of patents. The relative difference in values between the "Never Treated" and the "Pre Treatment" columns decreases from the "All Districts" definition, but is still relatively sizeable and potentially problematic. This is due to a few districts within London which are near a station but never receive their own station inflating the "Never Treated" mean in Panel A before they are removed in Panel B (because they are part of London).

One concern with this methodology is that by removing districts from the analysis the results are not

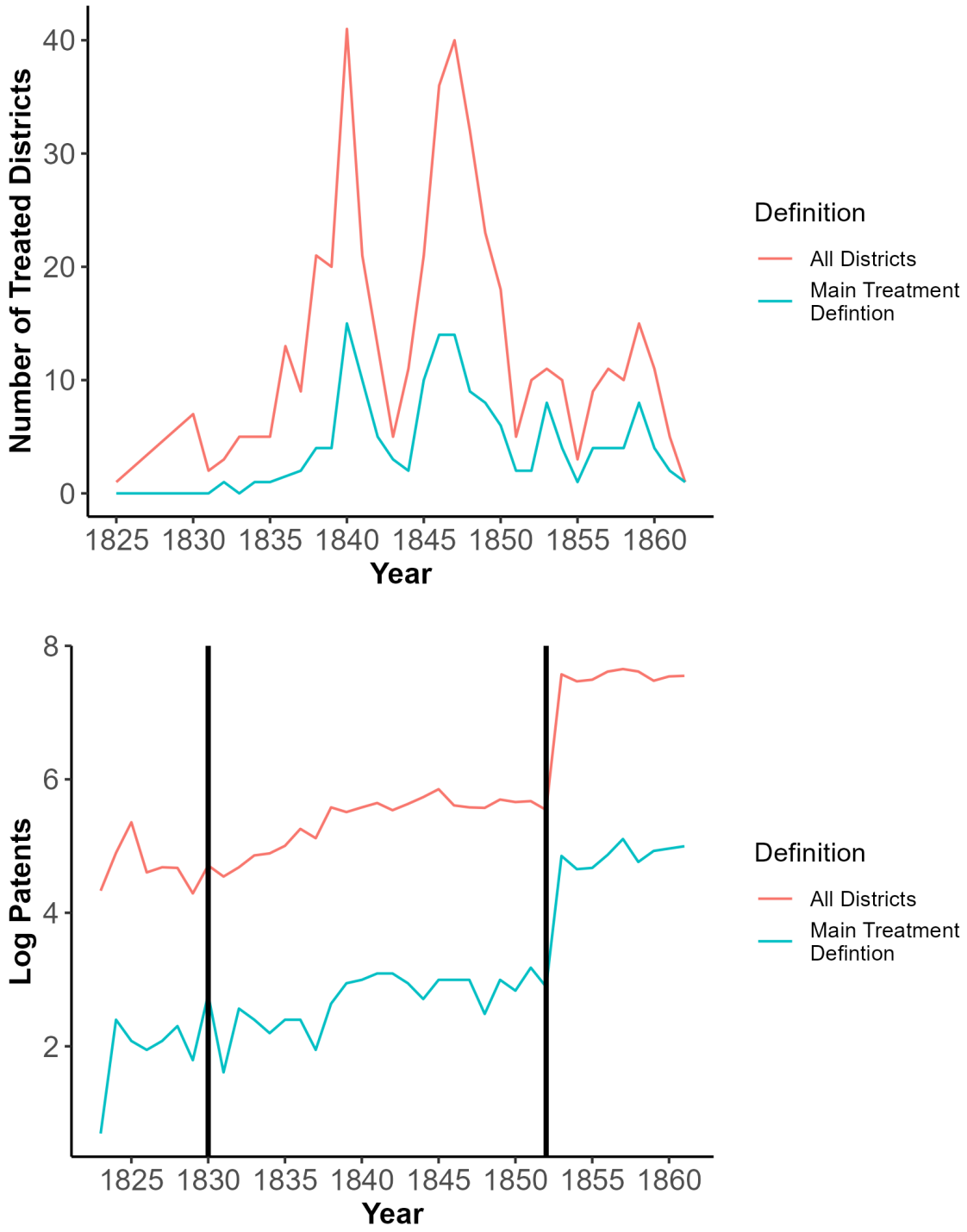


Figure 3: The top graph shows the number of districts first receiving a rail station in each year by treatment definition. The red line includes all districts and the blue line includes only districts that are either never treated or those which constitute the main treatment definition as defined in Section 3. The first station opened in 1825 and the first inter-city line opened in 1830. The bottom graph shows the log of patents applications by inventors by sample. Vertical lines at 1830 and 1852 represent the beginning of the Rail Age and the 1852 Patent Law Amendment Act respectively.

representative of the nation. This is potentially a valid point; urban centers are "important" for rail firms and so most are removed. The decrease in the mean population across treatment definitions is evident in Table 1. Nevertheless, I find that the composition of the data which comprises the main treatment definition is fairly similar to that which comprises all districts. The total number of districts treated in each year follows a similar trend across treatment definitions as does the number of log patents. This suggests that the main treatment definition is not altering the time component of the panel in a meaningful way. The main definition includes a higher proportion of agricultural patents, as one might expect given that many of the urban centers are removed. The composition of the other 55 sectors however is quite similar. A table comparing the patent sectoral composition across definitions is in the Appendix.

It is possible that firms' rail placement decisions had a political dimension that is not covered by this methodology. Parliamentary approval was required for each line and politicians with an interest in the rail lines were able to steer projects (Esteves and Geisler Mesevage 2021). To ensure that I am not removing lines which deviate from the least cost path due to political interference I use project approval from an apolitical team of engineers as a robustness check. This is discussed in greater detail in Section 6.

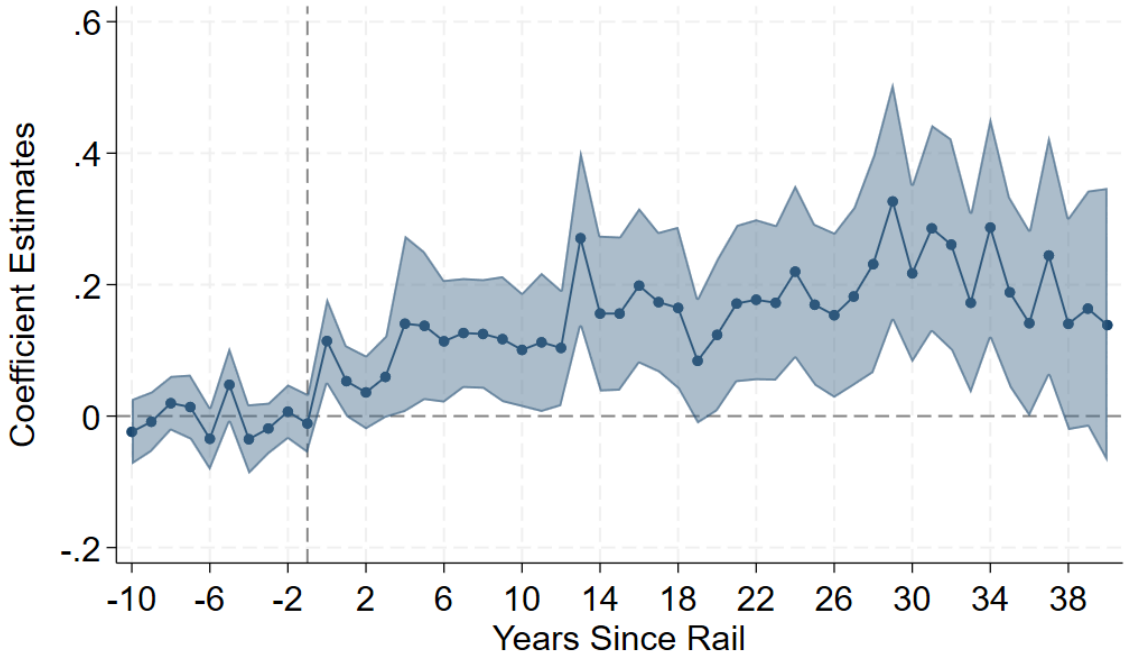


Figure 4: Event study plot with patents per capita as the dependent variable and 1823-1891 as the time period. Districts are treated in the year in which they receive a rail station. Because of limitations to the rail data, districts which are not yet treated by 1861 are considered as "never treated" districts. I use the difference-in-difference estimator proposed by Callaway and Sant'Anna (2021) to correctly estimate the "never treated" control group.

4 Main Empirical Results

4.1 Causal Effect of Rail

As a starting point to the analysis, I plot an event study graph using the difference in difference estimation strategy proposed by Callaway and Sant'Anna (2021) where districts which are not treated by 1861, the last year for which I have annual rail data, are considered "never treated." I estimate the effect of a rail station on a district's patents per capita during the period 1823-1891. The event study corresponds to Table 2, Panel A, column (5).

The event study plot shows a clear and durable increase in patenting activity following a district's connection to the rail network. The pre-treatment period is flat with no period statistically different from zero, indicating a lack of evidence that the pre-trends of the treated districts behave differently compared with the control. A similar graph with population as the dependent variable likewise shows no pre-trends. The post-treatment period shows an immediate jump at time zero - the year in which a district receives a rail station. It is plausible that patent applications were affected quickly for two reasons. First, the complexity of inventions was lower in the 19th century, requiring less time for an idea to come to fruition. Second, patent applications did not require significant details to be included in the initial application. Per the National Archive, "specifications were required and patentees had several months to submit these after the patent was sealed"(The National Archives 2023).

A concern with using patents as the dependent variable is that inventors could rely on secrecy and a small market to shield their innovation and then patent with the arrival of a rail station (to protect against rival entrepreneurs in the larger market). The long-running positive returns to patenting in Figure 4 make this implausible as the patenting effect would die shortly after the initial period post rail.

I estimate equation (1) using a Poisson model across 1823-1861 and present the results in Table 2. Errors are clustered by district to deal with potential serial correlation of the residuals over time. I present results where the dependent variable is the number of patent applications per capita (using an exposure variable for the imputed population) and when it is the number of patent applications. A positive coefficient for the former indicates that patent applications are growing faster than the population. I interpret the coefficients as the average effect of the treatment on the treated in the post-treatment period (ATT). In all specifications, districts are only retained in the estimation results if there is at least one positive value in the dependent variable because of the inclusion of district and time fixed effects.

The baseline result in column (1) indicates that having a rail station increases local patenting per capita by 77%.[20] The results are quite similar when excluding the population exposure variable in Panel B. This is possibly because the estimated increase of population caused by a local rail station is only 2.9%, an almost negligible figure when compared with the increase in patenting.

In column (2) I estimate the model using the broader "Accidental" treatment definition which includes districts otherwise excluded due to the simulated line network. The estimated coefficient of .479 is smaller than the baseline coefficient of .571 in column (1), though reasonably similar. This suggests that the simulated rail line network I create is not significantly changing the results, even as it removes 42 treated districts.

In column (3) I use the patent quality index created by Nuvolari et al. (2021). The authors' index is designed to evaluate the relative importance of patents between 1700-1850, assigning a value to each patent based on inventors' biographies and the history of technology. I limit the dependent variable to patents in the top 25% of the Nuvolari et al. (2021) index within the 1823-1850 period. This removes most of the

[20] I find the implied elasticity by using the transformation of (exp(coefficient)-1)*100.

Table 2: Effect of Rail Network Access on Innovation

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Patents per Capita						
Rail	0.571***	0.479***	0.807*	0.394**	0.173***	1.033***
	(0.139)	(0.130)	(0.487)	(0.164)	(0.039)	(0.275)
Analysis	Main	No End Pts.	Quality	Long Run	C & S	SDID
Observations	6864	8502	1300	12144	12144	2464
Districts	176	218	50	176	176	176
Years	1823-1861	1823-1861	1823-1850	1823-1891	1823-1891	1823-1861
Mean of Dep. Var.	0.247	0.237	0.075	0.759	.255	1.007
Panel B: Patents						
Rail	0.576***	0.480***	0.859*	0.343*	0.625**	0.338***
	(0.147)	(0.136)	(0.485)	(0.180)	(0.254)	(0.095)
Analysis	Main	No End Pts.	Quality	Long Run	C & S	SDID
Observations	6864	8502	1300	12144	12144	2464
Districts	176	218	50	176	176	176
Years	1823-1861	1823-1861	1823-1850	1823-1891	1823-1891	1823-1861
Mean of Dep. Var.	0.247	0.237	0.075	0.759	0.759	.229

Notes: Poisson estimation of equation (1) where the dependent variable is the number of patent applications per capita in Panel A and patent applications in Panel B. The per capita rate is calculated using an exposure variable for population. The main independent variable is an indicator equal to one if a district has a rail station in year t. District and time fixed effects are included in all columns. The main results are displayed in column (1). Column (2) uses the "Ends and Hubs Removed" treatment definition which excludes rail line end points and hubs, but does not use the simulated line network to further restrict districts. Column (3) only includes patents in the top quartile by patent quality as the dependent variable. The patent quality dataset only runs to 1850. Column (4) extends the time period to 1891. Districts treated after 1861 are considered untreated for the entire period due to constraints with the rail data. Column (5) reports the estimates of a Callaway and Sant'Anna (2021) style model only using "Never Treated" districts as the control. Column (6) reports the estimates of a synthetic difference-in-difference model proposed by Arkhangelsky et al. (2021). Due to the low patent count, in column (6) periods are aggregated to three years to ensure a more reasonable pre-trend comparison between the treated and control groups and patent counts and population are averaged across the three years. Districts with zero patents in every year between 1823-1861 are removed from the analysis in all columns due to the inclusion of district and time fixed effects. Errors are clustered by district in all columns. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

patents that had little historical impact. I find that the increase in patenting activity was not limited to low quality patents; a connection to the rail network caused local patenting to increase by 124%. The error term however is higher than in the baseline result in column (1). This is likely due to a large decrease in observations. The dataset provided by Nuvolari et al. (2021) ends in 1850, which significantly curtails the number of observations. Patenting was more expensive prior to the 1852 Patent Law Amendment Act and there are fewer patents in this period. Nevertheless, the large coefficient provides assurance that the effect on top quality patents is likely positive.

For the baseline result, I restrict the data to 1823-1861, the last year in the rail dataset. I extend the time frame to include an additional thirty years in column (4) by considering all districts not yet treated by 1861 as untreated until 1891. The coefficient of .391 is somewhat lower than in the shorter time frame, though it still demonstrates a sizeable, long lasting effect from rail treatment. A likely reason for the decrease in

the coefficient is that additional districts were treated during this time period, increasing patenting activity in the control group. Indeed, by 1881 (the only additional year of rail data) the vast majority of districts are treated. Ignoring new rail construction beyond 1861 is an even larger issue when exploring the market access mechanism as market access depends on construction of the entire network. Restricting the data to 1861, as I do for most estimates, avoids this issue.

The staggered expansion of the rail network presents a potential issue in that the empirical specification dictates that treated districts are compared against all districts which are not treated in year t. This includes both "Never Treated" districts and "Not Yet Treated" districts. It is possible that a rail station might have heterogenous effects across time. This would entail a larger effect for districts that were treated in some years than in other years. To remedy this, I estimate the model using the estimation strategy proposed by Callaway and Sant'Anna (2021), which only compares treated districts with never treated districts for the extended 1823-1891 period. The results are in column (5) of Table 2. Unlike the Poisson estimates in columns (1)-(4), these estimates are of a linear model. To compare, I divide the coefficient by the mean, which produces an implied semi-elasticity of 68%, somewhat similar to the implied elasticity in column (4) of 48%.

I utilize the "Synthetic Difference in Difference" (SDID) estimation strategy proposed by Arkhangelsky et al. (2021) to verify that the difference-in-difference parallel trend assumption is not violated. The SDID estimator optimally weights the data to ensure that the parallel trends assumption holds. Due to the low count patent data, I aggregate each time period to three years and take the mean patent and population values. A district is shocked in period t if it receives a station in any of the three years that comprise that period. This aggregation allows for a more meaningful comparison of the pre-trends between the control and treated groups. Like the Callaway and Sant'Anna (2021) estimator, the SDID estimator is linear; the implied elasticity of 105% is somewhat larger than the baseline estimate in column (1) but still reasonably close, indicating that the weighting does not meaningfully affect the results, even as it produces a control group with pre-trend lines that appear close to parallel with the treated group's pre-trends.[21]

4.2 Comparison of Results with the Literature

A few recent papers also find that gaining railroads produces large, positive effects on innovation. Andersson et al. (2023) find that Swedish municipalities were 15-30 percentage points more likely to patent relative to their baseline measure, similar to the 35% elasticity I find when estimating the extensive margin effect, while Perlman (2017) finds that over the two decades following a rail connection, patenting per capita in an American counties doubles.

While my results are in line with other papers' findings on the effect of railroad networks, these findings are much larger than what researchers have found in other settings. Agrawal et al. (2017) and Bottasso et al. (2022) find that that a 10% increase in the amount of highways in a region causes a 1.7% - 4% increase in regional patenting over a five year period. Those papers both use a single five year change as the dependent variable and measure the treatment as an intensive margin effect rather than an extensive effect, complicating comparisons. Nevertheless, interpreting the results as an average treatment of the treated implies that the impact to innovation is much larger in the setting I study than in either of those papers. The upper bound of their estimates is still half the size of the baseline estimate in Table 2. Pauly and Stipanicic (2024) study the airline industry and find that an increase in knowledge access caused patenting to increase by 3.5% annually.

[21] Parallel trend graphs utilizing SDID weighting can be found in the Appendix.

One plausible reason for the difference in magnitude between the railroad literature and the non-railroad literature is the setting. National railroad networks were built a century before the airline and highway networks studied by other researchers. A railroad made overland travel dramatically cheaper for both goods and people; highways provided more marginal gains as they complemented the existing railway networks. Another possibility is that railroads increase innovation through multiple channels: by increasing market access and by increasing knowledge flows as I discuss below. The literature on the returns to innovation from non-railroad transportation networks has not focused on market access as a source of the increase; rather these papers all focus on the knowledge channel. Because of this, it is difficult to say whether a market access driven increase in innovation is ignored by these papers or unique to railroads.[22]

5 Mechanism

I now present evidence that much of the returns to patenting can be explained through a demand channel: an increase in market access. A rail network lowers freight costs and therefore increases trade and a firm's potential market size. An increase in a firm's market size causes an increase in the firm's expected profits which induces greater R&D investment by the firm because it can spread the fixed costs of that investment across larger returns (Sokoloff 1988; Melitz and Redding 2023). I explore whether this setting provides evidence that an increase in market access increases local innovation.

For this mechanism to be plausible inventors must act as profit maximizing agents, rather than "hobbyist inventors," and they must be able to take advantage of the decrease in freight costs from the rail network. Using full names and patent years, I am able to match 51 patentees to the Oxford Dictionary of National Biography (2024). Of these, 35 (76%) are unambiguously patenting in this period as firm owners or managers within the sector of their firms. These inventors would not gain from a decrease in freight costs if the rail lines were used only to carry goods between the cities, ignoring the freight needs of the small towns in between them. The first intercity line, the Liverpool and Manchester Railway, was built primarily to carry freight between Manchester and the port of Liverpool as the potential of rail with respect to moving people had not yet been fully understood (Wolmar 2008). The 31 mile line included 7 stations between these end points, suggesting that firms built rail lines to load and unload freight between the end points as well as the major destinations they connected. Additionally, the 1845-1846 Railway Returns shows that one third of firms' revenue came from freight, a significant amount suggesting that rail provided a sizeable boost to domestic trade more generally (Railway Commission 1848).

5.1 Market Access

5.1.1 Defining Market Access

In order to measure the market size available to firms within a district, I adopt the definition of market access developed by Donaldson and Hornbeck (2016). Market access for a district i is the summation of the product of local GDP j multiplied by a transportation cost function, where j is any district in England not equal to i. As the transportation cost between districts decreases, market access increases.

[22] Pauly and Stipanicic (2024) include a control for market access as a robustness check, but the authors do not report the coefficient.

$$MA_{i,t} = \sum_{j=1}^{J} GDP_j * {c_{ij,t}}^{-\theta}$$
$$c_{ij,t} = distance_{ij} * cost_{ij,t} \tag{2}$$

The cost function $c_{ij,t}$ represents the cost of transportation between districts i and j and depends on the distance and transportation options between those districts in each year. I calculate the optimal route between the centroids of each district pair using Dijkstra's algorithm where a journey can occur on roads, railroads, and navigable waterways. The district centroids, railway stations, and points along the navigable waterways[23] are the nodes between which the optimal journey progresses; the waterways, railways, and overland straight line routes are the edges which connect these nodes. The algorithm uses any of these three transportation technologies depending on what is most efficient for each stage of the journey. Because of this, $c_{ij,t}$, and by extension $MA_{i,t}$, can change even if there is no new construction of local rail infrastructure; if the districts between i and j are conveniently connected by rail in year t, the cost of transportation between i and j will decrease accordingly as freight can now move along rail for that part of the journey.

I use the 1851 navigable waterway network data from Satchell and Shaw-Taylor (2018) to map canals. Using static data for the entire time period could cause market access to be incorrectly measured. However, much of the canal network was constructed in the 18th century. Construction slowed even further in the 19th century as the rail network was developed. Between 1800-1829, there was an annual average of 5.0 canal projects approved by Parliament. In the subsequent 30 year period, there was an average of 2.7 approved canal projects and 57.8 railway projects (Casson 2009).

I assign a value of .36 shillings per ton per mile to move freight along railroads and .54 shillings per ton per mile to move along canals. These values are derived from the freight cost estimates of transportation between Liverpool and Manchester in 1832 (UK Parliament 2024). Movement along roads is assigned a value of 4.99 shillings per mile per ton, a value which I derive from the average 1820-1827 carriage freight rates reported by Bogart (2005).[24] Lastly, I assign a cost of 10.78 shillings to load and unload freight.[25]

I assign the trade elasticity parameter θ a value of 8.22 following Donaldson and Hornbeck (2016). I approximate local GDP using the 1815 district land value (British Parliamentary Papers 1854a).[26] Utilizing values from a year that predates the rail network eliminates endogeneity concerns arising from the rail network inducing local changes in income levels and land values.

5.1.2 Summary Statistics

Rail decreased freight costs dramatically in this period. $c_{ij,t}$ is the freight cost between district pairs across time. Figure 5 shows the mean of $c_{ij,t}$ across time. As only the rail data is time varying, the change in freight cost is only due to new rail construction. The average cost of shipping a ton between any district pair in England decreased by 60% from 1840 - the start of the first major rail construction wave - to 1861.

[23] Lacking data on canal docks along the navigable waterways, I assign the nearest point on a waterway to the centroids of all districts through which a waterway passes as a waterway node.

[24] The costs that Bogart (2005) reports are incurred by carriages along the turnpike network. I use this cost to approximate the cost along the straight line path between districts, canals, and rail stations. Given the ubiquity of the turnpike network by 1830 (Rosevear et al. 2017), this is unlikely to represent a large deviation from the actual overland freight costs.

[25] I derive this value from the proportional cost between the overland freight costs and loading and unloading costs in Donaldson and Hornbeck (2016), which in turn derives these costs from Fogel (1964) which estimates freight costs for the US economy in the 19th century.

[26] As a robustness test, I also use the 1821 district level population as a proxy for GDP.

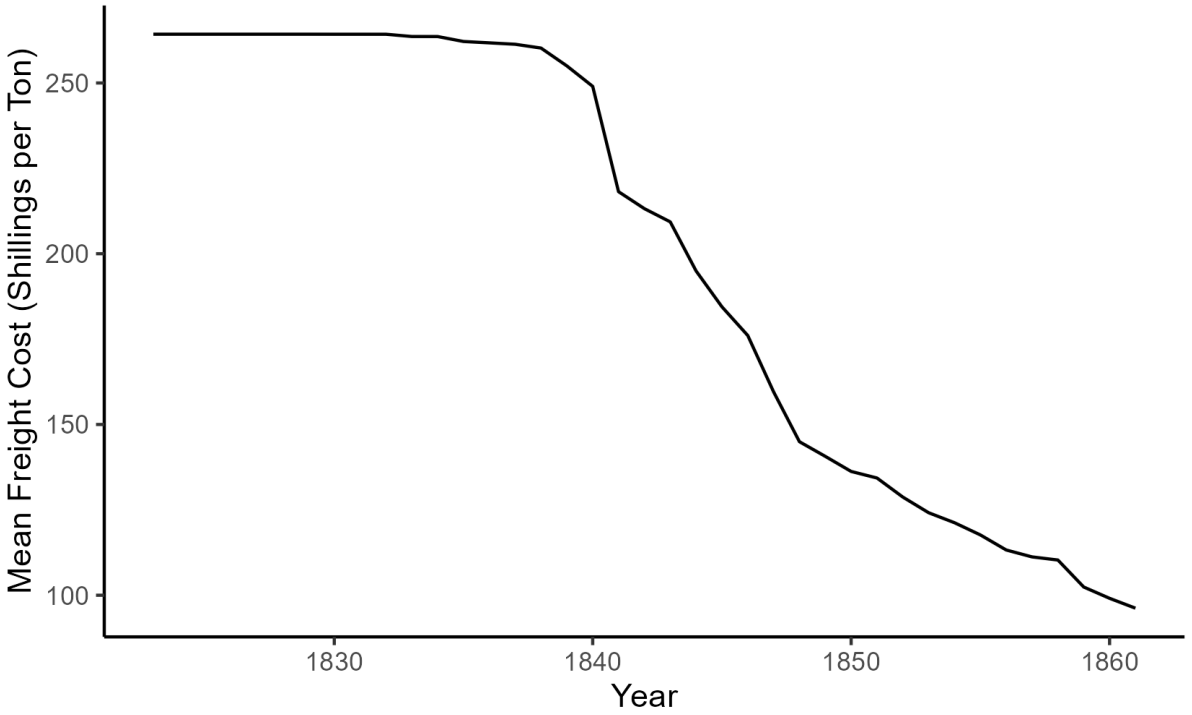


Figure 5: Mean change in $c_{ij,t}$ across time where $c_{ij,t}$ is the average cost of moving a ton of goods along the most efficient path between any district pair i and j.

For most districts, market access is a very small number; the median value of market access in 1823 is 2.79e-08. For all regressions I take the log of market access and therefore measure the change in market access. Table 3 reports summary statistics: the median market access in 1823 (column 1), the median market access percent change between 1823-1861 (column 2), and the median change in the year in which treated districts receive a local rail station (column 3). The difference in the values of columns 2 and 3 between all districts and districts which are included in the main treatment definition is mainly due to the population composition of the treatment definitions. The districts excluded by the main definition skew more urban. Districts with larger populations will receive a smaller increase in market access than the districts around them because they do not receive the benefit of connecting with their own large market.

For districts included in the main treatment definition, market access doubles between 1823-1861. Market access increases by 35% in the year of treatment compared with the previous year. Districts with the highest starting level market access see disproportionately smaller increases. This largely holds across the entire distribution of starting level market access. There is also a large gap between districts in the top quintile by market access change between 1823-1861 and those in the bottom 80%. The gap suggests that the broader increases in market access are driven by this top quintile. I find evidence that the districts in the top quintile by market access growth are driving the effect on local patenting below.

Figure 6 shows that there is a constant increase in market access following a district's connection to the rail network. As the network grows, travel between districts becomes cheaper which increases market access. The period prior to the district's connection year is marked by little to no change in market access. This is likely because overland freight costs were 13 times higher on roads than on railroads - travel to a nearby district's rail station may well have been prohibitively expensive and therefore unlikely to shift market access.

Table 3: Summary Statistics for the Market Access Variable

Definition	Median MA 1823	MA Ch. (%) 1823-1861	MA Ch. (%) Rail Treatment
All	2.79e-08	73.42	11.25
—Top Population	2.98e-07	24.60	0.23
—Top 1823 MA	6.57e-05	0.23	0.03
—Top MA Growth	3.22e-09	1,598.66	171.23
Main	1.95e-08	95.40	34.87
—Top Population	1.82e-08	217.77	42.28
—Top 1823 MA	1.74e-07	25.13	10.58
—Top MA Growth	7.45e-09	1,316.52	270.28

Notes: Summary statistics for market access using various subsets of the data. The top half of the table uses data from all districts and the bottom half from districts which are included in the main treatment definition. I restrict these districts to the top quintile by 1831 population, 1823 market access, and 1823-1861 market access growth. Column 1 reports the median market access in 1823; column 2 reports the median market access change between 1823-1861; column 3 reports the median market access change in the year in which a district receives a local rail station. Only districts treated between 1823-1861 are included in column 3.

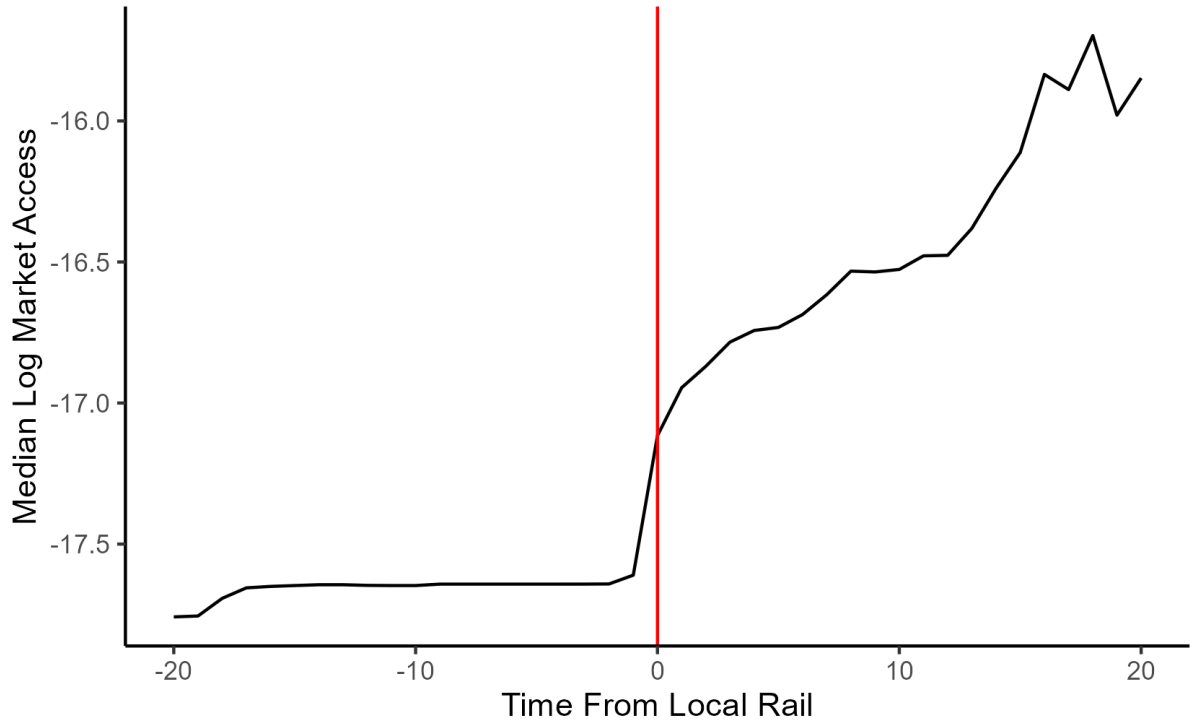


Figure 6: Log market access averaged across time from rail station construction in a district. Time zero is the year in which a district received a rail station. Only districts which constitute the main treatment definition are included.

Table 4: Effect of Market Access on Innovation

	(1)	(2)	(3)	(4)	(5)
Log Market Access	0.511***	0.427***			0.145**
	(0.130)	(0.141)			(0.066)
Rail		0.326**	0.120	1.120***	
		(0.164)	(0.190)	(0.311)	
Observations	6864	6864	6864	6864	18642
Districts	176	176	176	176	478
Years	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861
Definition	Main	Main	Main	Main	All

Notes: Columns (1), (2), and (5) report the results of a Poisson estimation of equation (3) where the dependent variable is the log of patents per capita in district i and year t and the main independent variable is log market access ($MA_{i,t}$). Population is measured as an exposure variable. Column (2) adds the rail indicator $D_{i,t}$ which is equal to one for a district i for all years in which i is connected to the rail network. Column (5) includes all districts - all other columns use the main treatment definition. Columns (3) and (4) estimate equation (1) where the main independent variable is the rail indicator $D_{i,t}$. In column (3), all districts which are treated but which receive an increase in market access between 1823-1861 that is in the top 20% of all districts are assigned as control districts for all years and the bottom 80% are considered treated. In column (4) districts in the bottom 80% are assigned as the control. In both columns, the treated districts receive treatment in the year in which they receive a rail station. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

5.1.3 Effect of Market Access on Patenting

To estimate the effect of market access on patenting I modify equation (1) by replacing the difference in difference estimator $D_{i,t}$ with $log(MA_{i,t})$:

$$E(Patents_{i,t}) = exp(\alpha_i + \lambda_t + \beta * log(MA_{i,t})) \tag{3}$$

The estimation results are presented in Table 4 and the baseline result in column (1) indicates a patenting elasticity of .51 with respect to market access. Including the local rail station term $D_{i,t}$ as an independent variable in column (2) only somewhat reduces the baseline estimate to an elasticity of .42. This result demonstrates that the increase in market access is not driven solely by local rail construction. Further, the coefficient for $D_{i,t}$ decreases to .326 from .571 in column (1) of Table 2. This implies that 43% of the local rail station effect is explained by the change in local market access, indicating that market access is a major channel.

Local rail stations produced much larger returns to patenting in districts which experienced large market access increases. Dividing districts by their overall 1823-1861 market access growth, I estimate equation (1) where all districts in the bottom 80% by market access growth are considered as treated districts in all years following the construction of a local rail station and those in the top 20% are considered as control districts (as are never treated districts). The results are statistically insignificant and with a relatively small coefficient. The reverse of this exercise - considering the top 20% as the treated and the bottom 80% as the control - produces a statistically significant coefficient that is twice as large as the baseline case in Table 2. This exercise acts as a heterogeneity test of the market access results more broadly in that districts with a higher level of market access growth experience greater patenting growth. Additionally, these results indicate that rail treatment only increases local patenting if there is a concurrent large increase in market access; a

local rail station must serve an important function for it to cause an increase in patenting.

5.1.4 Endogeneity

A feature of the market access definition is that it does not rely only on local rail construction for variation. Variation arises from changes to the entire transportation network linking districts i and j. Decisions regarding rail construction in district j or any district between i and j will increase market access in district i and are plausibly orthogonal to changes in district i.

Column (2) of Table 4 reports the coefficient estimate of log market access when controlling for changes in the local rail status of district i. The control only slightly lowers the coefficient on log market access, indicating that it is indeed the rail network beyond district i rather than the construction of a local station that causes much of the increase in local patenting.

If changes in districts' market access are orthogonal to changes in those districts' railroad infrastructure, it may no longer be necessary to restrict locations to those which comprise the main treatment definition. Including all districts in England produces a smaller coefficient of .15 (column 5). This result is less than a third of the size of the baseline coefficient. Much of this discrepancy can be attributed to the fact that the main treatment definition includes few cities and a smaller proportion of districts close to canals. Removing districts in the top decile by total 1815 income level (i.e., large local markets) or with immediate access to the main inland navigable waterway[27] increases the coefficient to .33.[28] It is plausible that places that are isolated from large markets and with poor transportation options can engage in high "catch-up growth" in patenting due to the big jump in market access.

5.1.5 Sector Heterogeneity

The theory underlying these empirical results is that a larger market spurs innovation because it spreads R&D costs over a larger customer base (Sokoloff 1988). I empirically test this theory by examining sector heterogeneity. Figure 7 shows the coefficient estimates of regressing patents on the interaction between market access and the 1851[29] labor share of two sector groups: high fixed, low variable cost sectors and niche product sectors.

Districts that are more exposed to high fixed cost, low variable cost sectors experience a larger increase in patenting than other districts as they gain more from the increase in R&D cost spreading and lowered transportation costs.[30] The coefficient estimate is large and statistically significant, implying that the high fixed, low variable costs sectors are driving the returns to patenting.

A niche product is more likely to require shoppers to travel greater distances than a general good because it is sold at fewer locations (Brynjolfsson et al. 2003). An increase in market access is analogous to a decrease in this higher transportation cost for consumers. I define niche sectors using the 1851 national labor shares. Any sector for which only five districts make up at least 25% of the sector's national labor force are considered niche. As in the case with high fixed, low variable cost sectors, the coefficient estimate of the interaction between the 1851 niche sector labor share and market access is statistically significant.

[27] Access is defined as the waterway passing within the geographic boundary of a district.

[28] Excluding all districts with any waterway access, rather than just access to the main waterway, produces an almost identical result, but the regression includes 47 fewer districts.

[29] This is the earliest year of labor data in the English Census.

[30] I designate the following sectors as high fixed cost, low variable cost: metallurgy, energy and gas manufacture, distilling, infrastructure, photography, chemistry, and all textile sub-sectors.

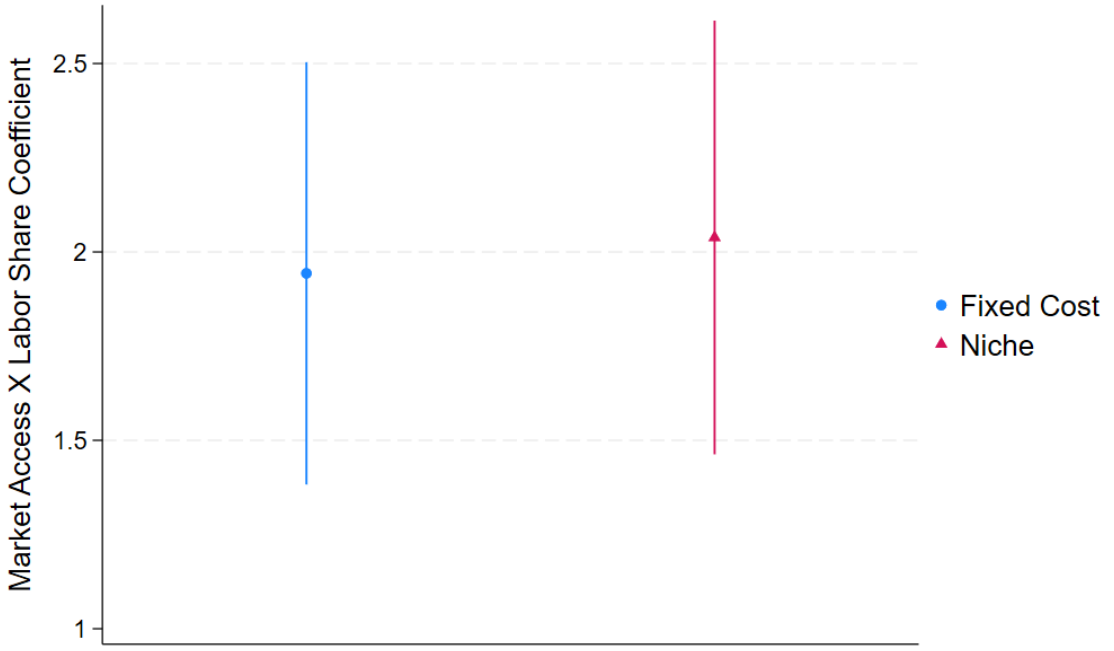


Figure 7: Coefficient estimates of equation (3) where the dependent variable is the number of patents and the main independent variable is log $MA_{i,t}$ and the interaction between log $MA_{i,t}$ and the 1851 share of labor in high fixed, low variable cost sectors (left) or niche sectors (right). Only the interaction coefficient estimates are included in the graph. Only districts which are included in the main treatment definition are included. Errors are clustered by district.

These sectors also exhibited a small increase in labor, suggesting that there was an increase in demand for their products and is consistent with the the firms increasing expenditures. Figure 8 shows a scatter plot of districts' log change in market access between 1851-1861 and the concurrent log change in the high fixed cost or niche labor share. The relationship is generally positive though with a small coefficient.

To further examine sector heterogeneity in the market access term, I construct a sector level definition of market access using an input-output table for 1841 England compiled by Horrell et al. (1994). The drawback is that I am forced to use the 13 sectors defined by Horrell et al. (1994), rather than the 55 sectors I defined to precisely allocate patents. There is no clear designation of high fixed cost or niche sectors among the Horrell et al. (1994) definitions. Instead, I turn to the authors' designation of manufacturing sectors as sectors which are plausibly more affected by an increase in market access.[31] Manufacturing firms can reasonably be expected to gain disproportionately from the rail network as these firms can more easily ship their goods via rail than firms in other sectors, such as "gas and water."

I denote the sectors from the input-output table as $\overline{s}$. Sectors are mapped from the 55 sectors I assigned to each patent and occupation in the census micro-data to the 13 sectors defined by Horrell et al. (1994). Each district-sector receives input values from all other district-sectors which are then assigned as weights to local GDP, which is proxied by the 1851 local labor force:

[31]These sectors are: construction, food, drink, and tobacco, gas and water, metal goods, metal manufacture, mining and quarrying, other manufacturing, soap, candles, and dyes, and textiles, clothing, and leather goods.

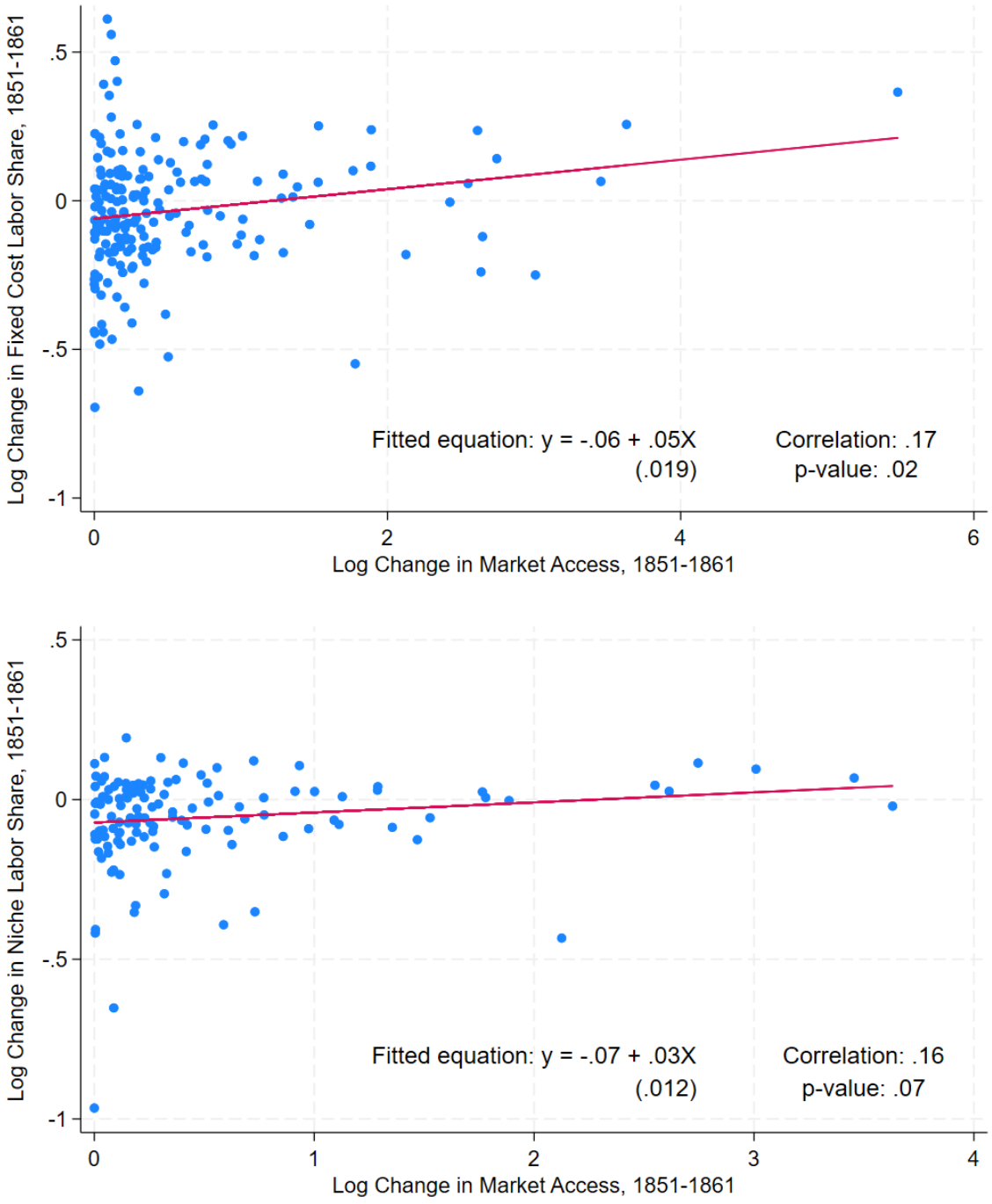


Figure 8: Scatter plot of the relationship between districts' log change in market access between 1851-1861 and the 1851-1861 log change of a district's high fixed, low variable cost labor share in the top graph and the log change in the niche labor share in the bottom graph. Only districts which are included in the main treatment definition and which have at least 100 workers in the given sector groups in both 1851 and 1861 are included. The pairwise correlation coefficients and its p-values are included in the bottom right hand corner of each graph and the fitted line equations are included at the bottom.

$$MA_{i,\overline{s},t} = \sum_{j=1}^{J} \sum_{\overline{k}=1}^{\overline{K}} Labor_{j\overline{k}1851} * (\frac{input_{\overline{k}\overline{s}}}{\sum_{\overline{k}=1}^{\overline{K}} input_{\overline{k}\overline{s}}}) * c_{ij,t}^{-\theta} \quad (4)$$

Market access for district i, sector $\overline{s}$, and year t is defined as the sum of sector $\overline{k}$ labor within each district j not equal to i multiplied by the share of total inputs into sector $\overline{s}$ that $\overline{k}$ constitutes and by the cost function $c_{ij,t}^{-\theta}$, which is unchanged from equation (2).

I use labor data from the 1851 Census to obtain local labor in district j. Since this is the first census to include occupational responses, it is the earliest year in which I can obtain these values. I restrict the regressions to 1851-1861 in order to minimize endogeneity concerns. Column (1) of Table 5 reports the coefficient estimate obtained when restricting the baseline regression of $MA_{i,t}$ on patenting to these years in order to provide a comparison when using $MA_{i,\overline{s},t}$.

I modify equation (3) by incorporating sectors to both the left and right hand sides of the equation such that the log of $MA_{i,\overline{s},t}$ is regressed on the log of patents in district i and sector $\overline{s}$ to obtain the effect of $MA_{i,\overline{s},t}$ on within sector patenting.

$$E(Patents_{i,\overline{s},t}) = exp(\alpha_i + \lambda_t + \beta * log(MA_{i,\overline{s},t})) \quad (5)$$

The estimation results in column (2) of Table 5 indicate that the within sector elasticity of patenting per capita with respect to market access is .40. This is marginally larger than the sector-less baseline

Table 5: Heterogenous Effect of Market Access on Patenting Across Sectors

	(1)	(2)	(3)	(4)	(5)
Log Market Access	0.349**	0.401***	0.281***	0.286***	0.164***
	(0.161)	(0.144)	(0.080)	(0.086)	(0.049)
Log Market Access X 1851 Lab. Sh.			0.291***		
			(0.095)		
Log Market Access X Manuf. Sector				0.219*	
				(0.119)	
Log Market Access X Manuf. 1851 Lab. Sh.					0.672*
					(0.345)
Observations	1896	20856	20856	20856	20856
Districts	158	158	158	158	158
Years	1851-1861	1851-1861	1851-1861	1851-1861	1851-1861

Notes: Column (1) estimates equation (3) where the time period is restricted to 1851-1861 where patents are not at the sector level and the main independent variable is $MA_{i,t}$. All other columns estimate equation (5) which uses log $MA_{i,\overline{s},t}$ as the main independent variable and log patenting per capita in district i, sector $\overline{s}$, and time t as the dependent variable. Columns (3)-(5) include an interaction term between log $MA_{i,\overline{s},t}$ and an indicator; the indicator in column (3) is equal to one if the 1851 sector $\overline{s}$ share of labor in district i is in the top quintile by all districts' share of sector $\overline{s}$ labor and zero otherwise; the indicator in column (4) is equal to one if $\overline{s}$ is a manufacturing sector; the indicator in column (5) is equal to one if district i is in the top quintile by districts' 1851 manufacturing share of labor. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

estimate of .35 in column (1). The larger coefficient provides confirmation that the incorporation of the input-output table yields additional explanatory variation as the inputs act as weights on local GDP, though the small magnitude of the increase is likely because the 19th century English economy was not particularly inter-connected. Indeed, the input-output table is quite sparse.

Column (3) reports the interaction between log market access and an indicator equal to one if the 1851 sector $\overline{s}$ share of labor in district i is in the top quintile by all districts' sector $\overline{s}$ share of labor and zero otherwise. The positive coefficient indicates that patents were disproportionately affected by the increase in market access in sectors for which there were enough workers to ramp up production. Seen another way, if London produces no output of textiles, an increase in market access does not produce an increase in textile patents.

In column (4) I estimate the interaction between log $MA_{i,\overline{s},t}$ and an indicator equal to one if sector $\overline{s}$ is a manufacturing sector to estimate the differential effect on patenting of manufacturing against all other sectors and find that manufacturing sectors experience larger returns to patenting. Repeating the exercise with districts in the top quintile by their 1851 manufacturing labor share shows (in column 5) that these "manufacturing towns" drive the increase in patenting.

5.1.6 Core Robustness Exercises

The results in this section are robust to a variety of tests, which are presented in the Appendix. Adjusting the parameter θ to either end of the range proposed by Donaldson and Hornbeck (2016) only slightly lowers the coefficient estimates for the main results in Table 4, and the results remain statistically significant. Using the 1821 population of district j as a proxy for GDP, rather than the 1815 land values, when defining market access similarly does not meaningfully change the results. Using stacked decadal differences instead of a difference-in-difference estimation framework produces a larger coefficient, but the results are still roughly similar. Systemically removing each of the ten census divisions (each of which represents a geographic area of England) does not greatly alter the estimation results, indicating that no single area is driving the results.

5.1.7 Aggregate Effects

A simple back of the envelope calculation shows that the change in market access due to rail caused a 48.55% increase in aggregate patenting. I calculate this by taking the median change in market access between 1823-1861 (95%) and multiplying by the baseline coefficient of .511 in column (1) of Table 4. Using figures from all districts, rather than those in the main treatment definition implies a 10.65% increase in aggregate patenting (from a 73% increase in market access and a .145 estimation coefficient). Using data from OECD countries between 1973-1996, Furman et al. (2002) find that the elasticity of TFP with respect to patents is .11. Multiplying 48.55% and 10.65% by 11% respectively produces 5.34% and 1.17% increases in aggregate TFP. These are naive estimates designed to produce an idea of aggregate, static results and make no attempt to estimate general equilibrium effects. Additionally, the elasticity of TFP could well be different in 19th Century England than in the latter half of the 20th century.

5.2 Other Mechanisms

This paper focuses on the market access channel to explain the increase in patenting from a local rail station. However, other channels may also explain some of the local station effect. I present evidence here that an increase in knowledge flows arising from easier communication due to the railways is a plausible

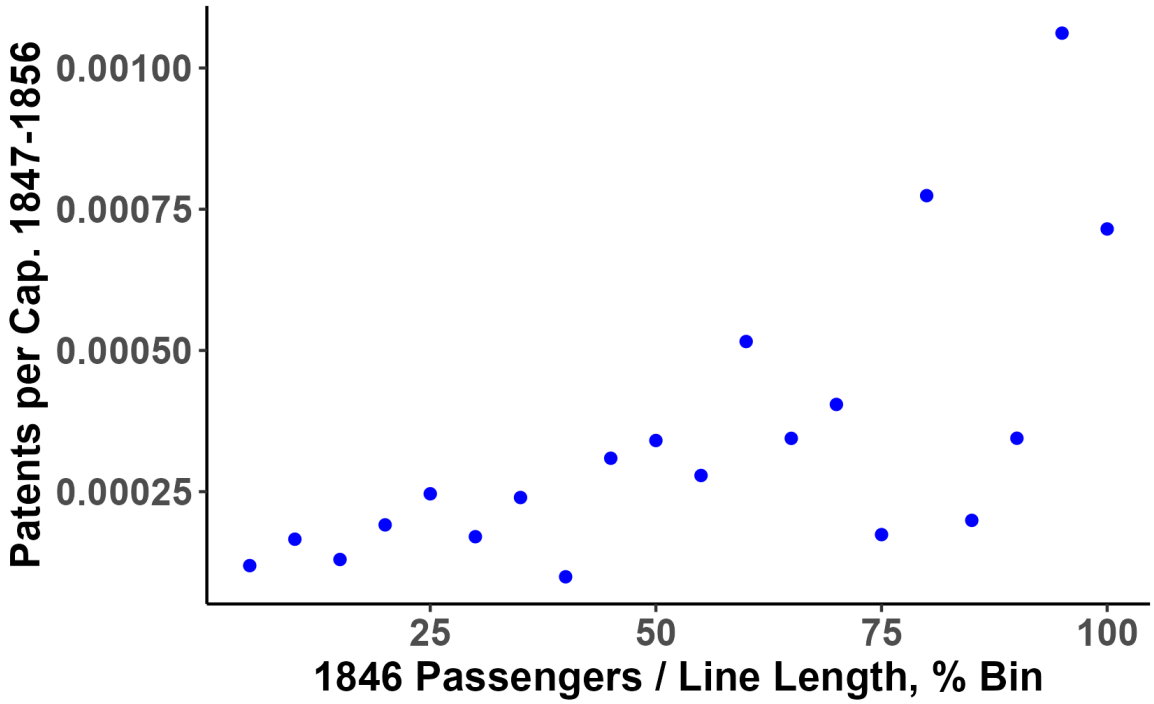


Figure 9: The x-axis shows the number of passengers in the 12 months period ending in June 1846 in a district divided by the total length of the rail line in each district. Districts are aggregated into five percentage point bins. The y-axis shows the mean number of patents per capita between 1847-1856 for each bin. The data exclude districts treated after 1846.

mechanism and that it is likely distinct from the market access channel. I also show that I find no evidence that the increase in local patenting is due to inventors moving to districts with rail stations.

5.2.1 Knowledge Flows

Previous research on the effect of railroads on innovation has focused on the diffusion of technology, rather than an increase in innovation (Juhász and Steinwender 2023). However, research on other types of transportation networks has found knowledge flows to be a mechanism driving local increases in innovation (Agrawal et al. 2017; Bottasso et al. 2022; Pauly and Stipanicic 2024). I present evidence for the existence of this channel in a railroad setting to establish that this is a plausible mechanism for the results described in Section 4.

A transportation network can increase knowledge flows because the network, in lowering travel costs, decreases the cost of face to face communication. This then induces innovation (Storper and Venables 2004). Rail construction had an immediate impact on individuals' movement. The first intercity rail line between Manchester and Liverpool opened in 1830 and was almost 40% faster than the average speed on turnpike roads. Within three months of the line's opening, over half of the stagecoach companies along this route ceased running, even as the rail line carried over 50% more daily passengers than the stagecoach companies had combined (Wolmar 2008). Following the 1844 Railway Regulation Act, affordable tickets were mandated by law on every route (Shaw-Taylor and You 2018).

Before establishing a formal definition of knowledge flows to measure the railroad's impact, I first highlight three pieces of evidence that support the notion that the railroad decreased communication costs. First,

Table 6: Effect of Local Rail on Wealthy Districts and Districts with No Prior Patenting

	(1)	(2)	(3)	(4)
Rail	0.588***	0.652***	0.571***	0.342**
	(0.152)	(0.160)	(0.139)	(0.174)
Rail X Top Income		-0.606**		
		(0.289)		
Rail X No Patents				0.852***
				(0.261)
Observations	5304	5304	6864	6864
Districts	136	136	176	176
Years	1823-1861	1823-1861	1823-1861	1823-1861
Definition	Main	Main	Main	Main

Notes: Poisson regression estimation of equation (1) where the dependent variable is patents per capita in year t and district i and the main independent variable is an indicator equal to one if district i has a rail in year t. In column (2) the rail indicator $D_{i,t}$ is interacted with an indicator equal to one if district i is in the top quintile by 1843 income per capita. In column (4) the rail indicator is interacted with an indicator equal to one if district i had no patents in any year between 1823-1839. Columns (1) and (2) exclude districts treated before 1843 to ensure that the income variable does not reflect changes due to rail. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

unlike poorer districts, districts that were in the top quintile by income per capita saw no benefit from a local rail station. This finding suggests that wealthy individuals could engage in distant communication even when it was costlier in the years before rail. Second, districts that were isolated from innovation - those with no patents between 1823-1839 - saw the biggest increase in patenting due to a local rail station. Individuals in these districts would have had to incur additional costs to travel outside of their towns to find inventive minds with whom they could discuss new ideas. Third, districts that were served by lines that carried more individuals saw a higher rate of future patenting. If ideas are produced by communication between individuals, it is reasonable that an increase in the total amount of potential communication between unique individuals leads to a higher rate of inventions.

Many papers follow Jaffe et al. (1993) and use patent citations to measure knowledge flows (for example, Agrawal et al. 2017; Atkin et al. 2022; Pauly and Stipanicic 2024). Patent citations were not recorded in the time period I study so I adapt my definition of market access to construct a variable to measure knowledge flows:

$$KF_{i,s,t} = \sum_{j=1}^{J} Patents_{j,s} * {c_{ij,t}}^{-\theta}$$
$$c_{ij,t} = distance_{ij} * cost_{ij,t} \tag{6}$$

The key difference between this definition and the market access definition described by equation (2) is that rather than including the sum of local GDP (as proxied by local land values) in other districts j, I include here the number of sector s patent applications in those districts. This proxies their inventive capacity so that cheaper travel between i and j produces an increase in knowledge flows into sector s only

Table 7: Knowledge Flows Summary Statistics

Definition	Median KF 1839	KF Ch. (%) 1839-1861	KF Ch. (%) Rail Treatment
All	5.91e-16	2,915.23	338.94
All - Top Population	1.35e-14	531.49	28.17
All - Top 1839 KF	1.1e-12	11.36	1.55
All - Top KF Growth	3.65e-19	1,085,759.00	1,411.42
Main	2.81e-16	4,235.80	437.07
Main - Top Population	1.77e-16	9,013.53	476.92
Main - Top 1839 KF	7.06e-14	104.89	14.33
Main - Top KF Growth	4.15e-19	721,561.10	1,180.49

Notes: Summary statistics for the knowledge flows variable using various subsets of the data. The top half of the table uses data from all districts and the bottom half from districts which are included in the main treatment definition. I restrict these districts to the top quintile by 1831 population, 1839 knowledge flows, and 1839-1861 knowledge flows growth. Column 1 reports the median knowledge flows in 1839; column 2 reports the median knowledge flows change between 1839-1861; column 3 reports the median knowledge flows change in the year in which a district receives a local rail station. Only districts treated between 1839-1861 are included in column 3.

if district j has inventors with the relevant knowledge. The sector level brings this definition of knowledge flows closer to patent citations and provides greater precision to the estimation. A distiller who becomes more easily connected to a district filled with textile inventors is unlikely to receive any knowledge flows that will improve the distilling process. However, connecting the distiller with other inventive distillers may foster relevant knowledge flows. $Patents_{j,s}$ is the cumulative total sector patents in a district between 1823-1838. Using a 16 year time frame allows for enough variation across districts despite the sparsity of the patent data.[32] I use 1839 as the cutoff as this is the year in which the first major rail expansion begins. This cutoff limits districts' exposure to the treatment in the calculation of $KF_{i,t}$ as the rail network was still in its infancy. Regressions which include $KF_{i,t}$ use 1839 as the starting point to limit this potential endogeneity. I estimate the following equation and report the results in Table 8

$$E(Patents_{i,s,t}) = exp(\alpha_i + \lambda_t + \beta * log(KF_{i,s,t})) \tag{7}$$

Table 8 column (1) shows that the within sector patenting elasticity with respect to knowledge flows is .43. Aggregating the data to the sector $\overline{s}$ level (which reduces the number of sectors from 55 to 13) in column (2) produces a coefficient nearly twice as large. While this could be due to the decrease in sparsity of patents for many sectors, there is also a clear loss of precision with regards to the knowledge flows due to the more aggregated level.

Including sector fixed effects decreases the coefficient from .427 to .185 in column (3). The sparsity of the data could be driving the decrease in magnitude; many sectors have few patents for the entirety of the period throughout all of England. The more aggregated sector $\overline{s}$ level produces a smaller difference in coefficients of .828 (column 2) and .576 (column 4).[33]

Including all districts of England in column (5) reduces the effect of $KF_{i,s,t}$ on within sector patenting to .093. This is unsurprising as knowledge appears to flow from cities and towards large towns and the districts which comprise the main treatment definition are less urban.

[32] 277 of the 541 districts in England have at least one patent during this period.

[33] Additionally, restricting the sector s data to only sectors in the top half by count of patents between 1823-1838 eliminates most of this difference, producing coefficients of .336 without sector fixed effects and .240 with fixed effects.

Table 8: Effect of Knowledge Flows on Patenting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Knowledge Flows	0.427***	0.820***	0.185**	0.576***	0.077***		0.179**
	(0.062)	(0.097)	(0.075)	(0.220)	(0.019)		(0.075)
Rail						0.466***	0.299*
						(0.148)	(0.169)
Observations	203573	49933	188209	49933	566812	188209	188209
Districts	167	167	167	167	465	167	167
Years	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861
Definition	Main	Main	Main	Main	All	Main	Main
Sector FE			Y	Y	Y	Y	Y

Notes: Results of the Poisson estimation of equation (7) where the dependent variable is the log of sector s patents in district i and year t and the main independent variable is log $KF_{i,s,t}$. Columns (3) - (7) include sector fixed effects. Columns (2) and (4) are estimated at the more aggregated sector $\bar{s}$ level. Columns (5) includes all districts while all other columns include only districts which comprise the main treatment definition. Column (6) includes the rail treatment variable $D_{i,t}$ as the main independent variable and column (7) includes both $KF_{i,s,t}$ and $D_{i,t}$. As $KF_{i,s,t}$ is calculated using patents from 1823-1838, I restrict the data to 1839-1861 in all columns. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 9: Effect of Knowledge Flows on Patenting - Population and Labor Interactions

	(1)	(2)	(3)
Log Knowledge Flows	0.185**	0.108**	0.086*
	(0.075)	(0.050)	(0.052)
Log Knowledge Flows X Top Pop.		0.190**	
		(0.090)	
Log Knowledge Flows X Top Lab.			0.139***
			(0.053)
Observations	188209	188209	188209
Districts	167	167	167
Years	1839-1861	1839-1861	1839-1861
Definition	Main	Main	Main
Sector FE	Y	Y	Y

Notes: Results of the Poisson estimation of equation (7) where the dependent variable is the log of sector s patents in district i and year t and the main independent variable is log $KF_{i,s,t}$. Column (2) includes the interaction between log $KF_{i,s,t}$ and an indicator variable equal to one if a district is in the top quintile by 1831 population. Column (3) includes the interaction between log $KF_{i,s,t}$ and an indicator variable equal to one if a district's 1851 labor in sector s is in the top quintile (for that sector). Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Controlling for district i's local rail status in column (7) only marginally reduces the coefficient. The coefficient on the local rail station term decreases from .466 in column (6) to .299 in column (7) when including log knowledge flows, indicating that 35.8% of the rail station effect on within sector patenting can be explained by knowledge flows.

Knowledge flows produce disproportionately higher returns to patenting in districts which occupy the top quintile by 1831 population. I find this result by interacting log $KF_{i,s,t}$ with an indicator for districts in the top quintile. Given that cities are mostly excluded from the main treatment definition, these districts can be considered to be "large towns." It is unsurprising that districts with the highest population receive the largest gains. Having a larger number of individuals increases the chances that someone in the district is a potential inventor who can benefit from the increased knowledge flows. I find further evidence of this when interacting log $KF_{i,s,t}$ with the top quintile by 1851 sector s labor. The positive result shows that it is necessary to have a large number of people with expertise in the given field to produce returns to innovation. In other words, knowledge flows require a large knowledge base in the relevant sector to foster innovation. Districts in the top quintile by population are likelier to have a large enough knowledge base as they simply have more people. Additionally, the labor composition of these districts is over-indexed in the sectors in which they produce a patent, further pointing towards the benefits of a large, relevant knowledge base.

It is plausible that sectors with relatively new technology are disproportionately affected by an increase in face-to-face communication as the knowledge in these sectors has not yet fully disseminated. For example, someone seeking to work in AI is likely to learn the most from people who are working at the cutting edge of the field rather than from a textbook that might be out of date within five years of publishing. The same is unlikely to be true for an older field such as plumbing. I find that sectors which grew the most by stock of patenting between 1823-1841 and 1842-1861 experienced a disproportionately large effect from knowledge flows. The effect is present in sectors above the median by patent growth and more pronounced in sectors

Table 10: Effect of Knowledge Flows on Fast Growing and Knowledge Sharing Sectors

	(1)	(2)	(3)	(4)	(5)	(6)
Log Knowledge Flows	0.370***	0.394***	0.418***	0.130**	0.130**	0.177**
	(0.059)	(0.054)	(0.064)	(0.062)	(0.059)	(0.075)
Log Knowledge Flows X Growth (Median)	0.078*			0.081*		
	(0.043)			(0.043)		
Log Knowledge Flows X Growth (Quintile)		0.109***			0.183***	
		(0.040)			(0.051)	
Log Knowledge Flows X Knowledge Sharing			0.176*			0.172*
			(0.100)			(0.095)
Observations	203573	203573	203573	188209	188209	188209
Districts	167	167	167	167	167	167
Years	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861
Definition	Main	Main	Main	Main	Main	Main
Sector FE				Y	Y	Y

Notes: Results of Poisson estimation of equation (7) where the dependent variable is the log of sector s patents in district i and year t and the main independent variable is log $KF_{i,s,t}$. Columns (1) and (4) include the interaction of $KF_{i,s,t}$ with sectors above the median by patent growth and columns (2) and (5) the sectors in the top quintile by patent growth. Columns (3) and (6) include the interaction between $KF_{i,s,t}$ and the metallurgy and dyes sectors. Columns (4)-(6) include sector fixed effects. As $KF_{i,s,t}$ is calculated using patents from 1823-1839, I restrict the data to 1840-1861 in all columns. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

in the top quintile.

I also find that sectors which have a culture of sharing ideas are disproportionately affected. The iron industry in England relied on "collective invention" during this period - the practice of competitors openly sharing secrets with each other (Allen 1983). The dyes industry relied on secrecy to prevent knowledge extraction rather than patenting (Moser 2013). In both cases, an increase in face-to-face communication allows outsiders to travel to a sector hub, learn trade secrets, and return home with insider knowledge. The interaction between log $KF_{i,s,t}$ and an indicator equal to one if sector s is either metallurgy or dyes produces a positive coefficient indicating that these sectors are indeed more likely to see a higher rate of patenting than other sectors due to an increase in knowledge flows.

As a robustness check, I recreate the main results in Table 8 using a five year rolling aggregate of patents in the definition of $KF_{i,s,t}$, rather than the static stock of patents between 1823-1839, such that $KF_{i,s,1845}$ uses all sector s patents between 1840-1844. While this ignores potential endogeneity concerns in that the rolling stock of patents which comprises $KF_{i,s,t}$ is produced in districts which are increasingly treated by rail stations, it is a more realistic view of knowledge flows. An inventor in 1860 is more likely to improve upon an invention from the previous few years than upon an invention from 1830. The results are similar to the main results.

As in the case with market access, systematically removing every district in one of the ten Census Divisions does not meaningfully change the results, indicating a lack of evidence that knowledge flows have

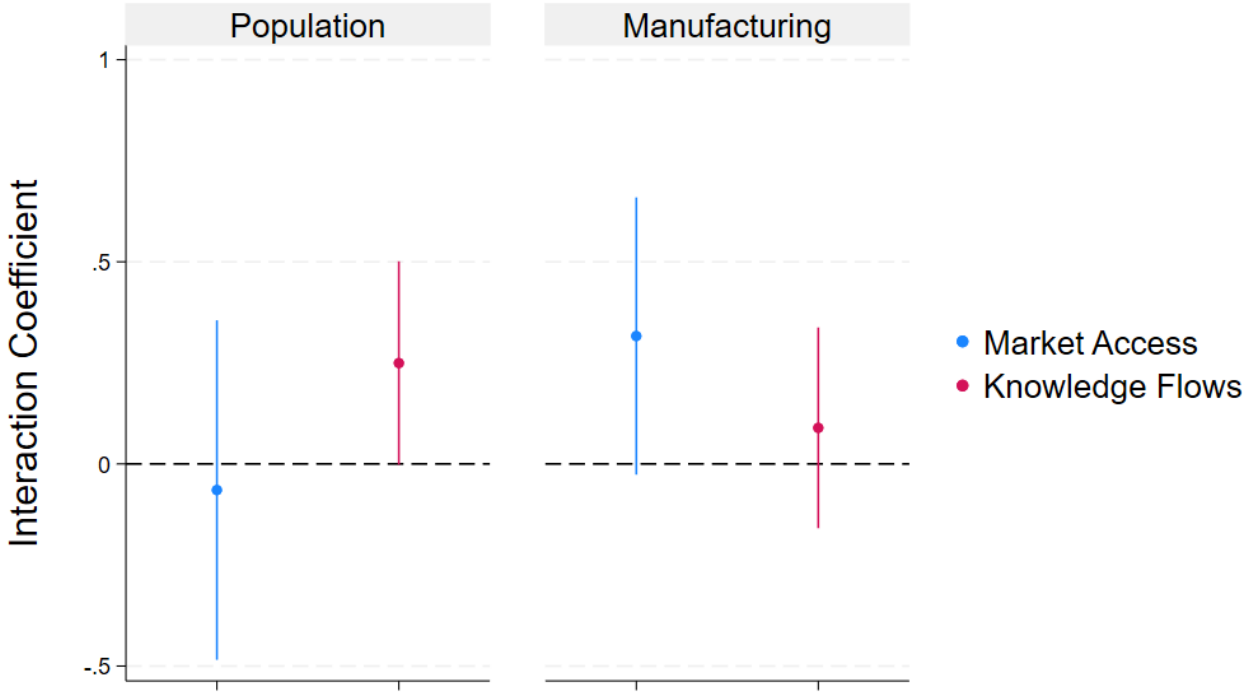


Figure 10: The left graph shows the coefficient of the interaction between an indicator equal to one for districts in the top half by 1831 population and either $MA_{i,t}$ (in blue) or $KF_{i,s,t}$ (in red). The right graph shows the same interaction using the top half by the 1851 manufacturing share of labor in a district. Halves are calculating using districts which comprise the main treatment definition.

a geographic component.

5.2.2 Disentangling of Market Access and Knowledge Flows

The market access and knowledge flows channels appear distinct. However, serial correlation is very high due to variation across time being driven entirely by changes to the rail network which creates difficulties in disentangling the variables. I present three pieces of evidence that these mechanisms measure distinct effects.

First, the market access and knowledge flows channels' largest effects go to different districts. In Sections 5.15 and 5.2 I showed that an increase in market access disproportionately affects patenting in districts with high manufacturing employment shares and knowledge flows disproportionately affects large towns. The reverse does not hold; the interaction term between $MA_{i,t}$ and a dummy variable equal to one for districts in the top half by 1831 population is essentially zero. The same is true for the interaction between $KF_{i,s,t}$ and a dummy for districts in the top half by the share of labor employed in manufacturing. This finding implies both that the channels affect different types of districts and that the channels' largest effects do not overlap.

Second, the correlation between the 1839-1861 log change in KF_i and log change in MA_i is negative and statistically insignificant. A lack of a relationship between these variables (when removing the serial time component) suggests that they are measuring different effects. The difference between the variables stems from the different weights that they place on the cost to travel between districts: other districts' GDP for market access and patents for knowledge flows. Population and patents are only somewhat correlated in this time period as half of the districts have zero patents between 1823-1838, though they clearly have a positive population. The lack of correlation between KF_i and MA_i could therefore perhaps be mechanical; given

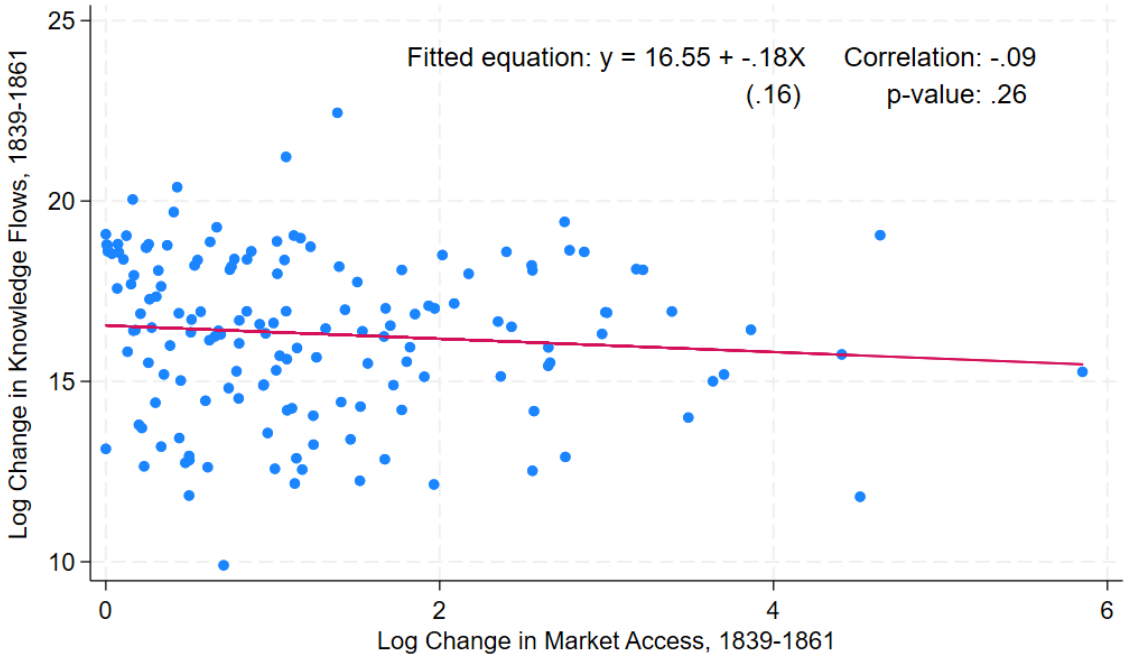


Figure 11: Districts' 1839-1861 log change in MA_i plotted against the 1839-1861 log change in KF_i. Only districts treated by 1861 and which are included in the main treatment definition are included.

Table 11: Effect of Market Access and Knowledge Flows on within Sector Patenting

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log Market Access	0.401***	0.340***	0.300***			0.271***	0.160***
	(0.144)	(0.074)	(0.066)			(0.077)	(0.028)
Log Market Access X Top KF Quintile		0.127	0.140				
		(0.254)	(0.216)				
Log Knowledge Flows				0.883***	0.872***	0.826***	0.887***
				(0.114)	(0.069)	(0.091)	(0.025)
Log Knowledge Flows X Top MA Quintile					0.013		
					(0.152)		
Observations	20856	20856	20856	20856	20856	20856	59268
Districts	158	158	158	158	158	158	449
Years	1851-1861	1851-1861	1851-1861	1851-1861	1851-1861	1851-1861	1851-1861
Definition	Main	Main	Main	Main	Main	Main	All
Pseudo R2	0.238	0.263	0.272	0.317	0.319	0.326	0.643

Notes: Poisson estimation where the dependent variable is the log of sector $\overline{s}$ patents per capita and the main independent variables are log $MA_{i,\overline{s},t}$ and log $KF_{i,\overline{s},t}$. Columns (2) and (3) include the interaction between log $MA_{i,\overline{s},t}$ and districts in the top quintile by 1851 and 1840 $KF_{i,\overline{s},t}$ respectively. Column (5) includes the interaction between log $KF_{i,\overline{s},t}$ districts in the top quintile by 1851 $MA_{i,\overline{s},t}$. Column (6) includes both log $MA_{i,\overline{s},t}$ and log $KF_{i,\overline{s},t}$ as regressors. Columns (7) includes all districts and all other columns include only districts which comprise the main treatment definition. The time period is restricted to 1851-1861 due to the use of labor data from the 1851 Census in the definition of $MA_{i,\overline{s},t}$. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

a positive number of patents across all districts the correlation might be positive. However, it is equally possible that in the first half of the 19th century, many towns had zero inventors (potential or otherwise) and connecting with that town would not produce any positive knowledge flows.

Third, access to a large knowledge base is not necessary for a change in market access to increase patenting. Similarly, knowledge flows do not require a large market. Interacting $MA_{i,\overline{s},t}$ with districts in the top quintile by the 1851 starting point knowledge flows does not yield a statistically significant coefficient. The reverse is also true. This again points to a lack of a relationship between the two channels. A plausible explanation for this finding is that the districts in which market access causes an increase in innovation already have all the knowledge they require to innovate, but are missing the market size to engage in high fixed cost R&D spending. This explanation fits with the result that high fixed cost, low variable cost sectors benefit most from an increase in market access (Figure 7) and with the finding that districts with a high sector $\overline{s}$ labor share produce greater returns to patenting from an increase in market access (Table 5 column 3). Another plausible explanation is that by 1851 many places with isolated markets or no stock of patents had been connected to the rail network and had already reaped the benefits and engaged in catch-up growth, flattening the difference in market access and causing all districts to have "large enough" markets to support the commercialization of incoming knowledge flows. However, $KF_{i,\overline{s},t}$ can reasonably be measured as far back as 1839, prior to the first wave of large scale rail construction. Using the 1839 value to calculate the top quintile does not change this result, suggesting that these findings are not due to catch-up growth.

Market access and knowledge flows are strongly correlated within the panel as variation across time for

both variables is due solely to changes in the rail network. Including both $MA_{i,\overline{s},t}$ and $KF_{i,\overline{s},t}$ in the same regression produces statistically significant effects, but the high serial correlation means that the estimates are unlikely to be robust and should be taken with caution. Beyond the correlation issues, there are also data constraints; the coefficient estimates when regressing knowledge flows on patenting are much larger when using $KF_{i,\overline{s},t}$ instead of $KF_{i,s,t}$, yet, as market access cannot be measured at the sector s level (because it requires the input-output table which is at the sector $\overline{s}$ level), a comparison between how the two knowledge flows definitions compare with market access cannot be made.

5.2.3 Movement of Innovation

It is possible that railroads do not increase aggregate innovation, but rather move innovation from one location to another by displacing inventors.[34] I do not find evidence that these negative spillovers are driving the innovation effects described in Section 4.

I modify equation (1) such that $D_{i,t}$ is equal to one if at least one district within 50KM of district i has a station in year t and remove observations for years in which district i has a rail station. In essence, this regression shows whether there is an outflow of patenting from district i when its residents can relocate to a nearby district with a station before their district is treated. The estimated coefficient in column (1) of Table 12 is statistically indistinct from zero, indicating a lack of evidence for such spillovers.

In column (2) I estimate the direct effect of a district's station on its neighbors. I again estimate a modified form of equation (1) where the dependent variable is now the mean number of patents per capita of a district's neighbors.[35]

$$E(\overline{Patents_{j,t}}) = exp(\alpha_i + \lambda_t + \beta D_{i,t}) \tag{8}$$

The difference-in-difference estimator $D_{i,t}$ is equal to one for all years in which district i has a local rail station. Neighboring districts j are only included as part of the mean patents calculation until they are treated. Neighbors which are treated prior to district i are always excluded. The coefficient is close to zero and statistically insignificant. I interpret this to mean that there is essentially no effect on the patenting activity of the neighboring districts.

In column (3) I examine the movements of the inventors themselves. I remove all patents by inventors who move between districts. Movers are identified as inventors whose full name is linked to multiple districts.[36] The limitation of this methodology is that location data only exists when an individual patents. Movement prior to patenting and after the individual's final patent application is therefore not seen. There is also the potential that individuals with the same name are treated as movers because their names are linked to multiple districts. It is unlikely that there is a bias in removing these individuals as they should be randomly distributed across England. After removing all patent applications by "movers" I estimate equation (1). The results are still positive, though no longer statistically significant. This is potentially due to a large decrease in the number of observations; a third of the patent applications are removed from the dataset as they pertain to "movers" and the time frame is limited to the lower patenting period prior to the 1852 reform. The coefficient implies an elasticity of 46%, lower than the baseline elasticity of 64% using the 1823-1852 sample, but still roughly similar.

[34] Pogonyi et al. (2021), Faber (2014), and Baum-Snow et al. (2017) find displacement effects in different settings when looking at economic or population growth following transportation improvements.

[35] This is calculated as the mean patents in the neighboring districts with the inclusion of the neighbors' mean population as an exposure variable.

[36] I only have access to inventors' names for the 1823-1852 portion of the dataset, which restricts this analysis to these years.

Table 12: Reallocation of Patenting Activity

	(1) Patents	(2) Patents	(3) Patents	(4) Moved to Rail
Nearby Rail	-0.122			
	(0.258)			
Rail		-0.079	0.375	
		(0.129)	(0.276)	
High Skilled				0.025
				(0.032)
Observations	3814	4901	3060	145703
Districts	127	137	102	402
Years	1823-1861	1823-1861	1823-1852	1851
Sample	Not Treated	Neighbors	Non Movers	Movers (ICEM)

Notes: Column (1) reports the Poisson regression estimates of a modified version of equation (1) where the dependent variable is patents per capita and the independent variable is an indicator equal to one if there is a district within 50KM of district i that has a local rail station in year t. A district is included up until it receives a local station. Column (2) reports the Poisson regression estimates of equation (8) where the dependent variable is the mean number of patents in the neighbors of district i in year t and the main independent variable is an indicator equal to one for years in which district i has a rail station. Neighboring districts are only included as part of the dependent variable until they receive a rail station; neighboring districts treated prior to district i are excluded. I control for the neighbors' mean population as an exposure variable. Column (3) reports the coefficient estimate of equation (1) using the subset of the patent data which includes names (1823-1852). Movers are identified by full name and inventors in the patent data who are matched to multiple Census districts are removed. Column (4) reports the logit model estimates of equation (9) where the dependent variable is an indicator equal to 1 if the individual has moved to a district which has a rail station. The main independent variable is an indicator equal to 1 if the individual is classed as a high skilled worker. I restrict the data to individuals from the 1851 Census whose birth districts are identified and are different from the districts in which they were born. Only individuals above 20 years of age are included. Additionally, only individuals who in 1851 reside in a district which is included in the main treatment definition are included. I control for individuals' age and include birth district fixed effects. Errors are clustered by district in columns (1)-(3) and by birth district in column (4). * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

I find no evidence that highly skilled workers are more likely to move to rail districts, indicating that inventors are no more likely than the rest of the population to move to rail districts. Individuals cannot reasonably be matched across censuses as they can only be linked by name. Additionally, individuals did not identify themselves as "inventors" in the 1851 Census. However, Schurer and Higgs (2023) provides the birth district of individuals in the census as well as the HISCO occupational category for employed individuals. The latter can be matched to a skill class using HISCLASS (Leeuwen and Maas 2011). Using only the universe of employed "movers" - individuals who moved from their district of birth - and who are above age 21 in 1851, I estimate a logit model wherein I regress an indicator equal to one if the individual is living in a district with a rail station in 1851 on an indicator for a high skilled occupation:

$$Station_{i,p,1851} = \alpha_j + \beta_1 Skilled_{p,1851} + \beta_2 Age_{p,1851} + \epsilon_{i,p,1851} \tag{9}$$

Station is equal to one if the district i that individual p is registered as living in for the 1851 Census has a station and zero otherwise, *Skilled* is equal to one if individual p is considered highly skilled in 1851 and zero otherwise, *Age* is a control variable for the individual's age as the decision to move to certain places is likely influenced by age, and α_j is a fixed effect term for the individual's birth district. Leeuwen and Maas (2011)

divide individuals across eleven categories of skilled labor. I categorize individuals in top two categories as highly skilled. These categories include entrepreneurs and managers, as well as engineers, occupations which 19th century inventors are likely to have been categorized as. The coefficient estimate in column (4) is essentially zero, indicating that the highly skilled were no more drawn to rail districts than lower skilled workers.

6 Additional Robustness Tests

The results presented presented above are robust to a variety of robustness checks. Outputs for these robustness exercises can be found in the Appendix.

The main results use an exposure variable to control for population changes. However, Poisson estimation cannot control for population in an exogenous manner by including it as part of the dependent variable as Poisson regression can only estimate a dependent variable that is a count. I replicate the main results in Table 2 with an OLS model estimating patents per 100,000 people. The implied elasticity of 83% is quite similar to the 77% implied by the main Poisson estimation, indicating that the Poisson estimation is unlikely to be misspecified by excluding population from the dependent variable. Given the sparsity of the data, I also apply an Inverse Hyperbolic Sine (IHS) transformation to the dependent variable. The results are positive and statistically significant, but the elasticity of 20% is much smaller than in the case without the transformation. Recent critiques of the IHS transformation (Mullahy and Norton 2022 and Chen and Roth 2023) argue that it should not be applied when there is a mass of zeros, as is the case in this data sample. In particular, one of the conclusions of Chen and Roth (2023) is to avoid interpreting the coefficient as a percent change, which likely explains the large difference between the coefficients and implied elasticities of the IHS model and the OLS and Poisson models.

As an additional test on the patent data sparsity, I estimate a Poisson model where the dependent variable is an indicator equal to one if the district has at least one patent in year t. The coefficient of .306 is somewhat lower than the baseline coefficient estimate of .571.[37] More than half of the difference can be attributed to the top four districts by maximum number of patents in a year. Given that the median number of patents in a district and year, conditional on patenting, is one, it can be expected that the extensive and intensive margin are similar.

I estimate the model using 2SLS with the treatment definition in which only the end point and rail hub districts are removed. The simulated lines are used to instrument the real lines. Treatment is assigned to districts through the 2SLS process, and the year of treatment is dictated by the opening of the district's first rail station. The implied elasticity of 93% is quite similar to the elasticity implied by the OLS model using the main treatment definition. The high Kleibergen-Paap Wald F statistic shows that the instrument is not weak.

One concern with regards to the empirical framework is serial correlation in the error term. To deal with this, the regressions thus far have all clustered errors by district, the unit of analysis. It is possible that there is serial correlation across a wider geographic area. Clustering at the spatial unit above district, the county level, however, yields a similar error term, indicating that this is not a concern.

The results are robust to a variety of definition and sample restrictions. Restricting the time component of the sample to 1823-1851 removes the years after the passage of the Patent Law Amendment Act when

[37] The coefficients are much more similar when restricting the data to 1823-1851, the period in which patenting was more expensive and in which there was significantly less patenting activity and therefore fewer districts with multiple patents in a single year.

patenting became significantly cheaper. This period exhibits a much lower patenting activity, making it prone to larger error terms. The coefficient is slightly smaller than the baseline coefficient but broadly similar indicating that the new law did not have an impact on the local rail station effect.[38]

As a district can occupy a large area, there is potential that assigning treatment according to whether a station is in a district is vulnerable to concerns over the precision of that definition. For example, a station can be near a district's border and serve multiple other districts, or even primarily serve a neighboring district depending on the population dispersion in the districts. I adopt an alternative treatment definition based on the 3KM radius circle of land around each station.[39] A district is considered treated if it has the largest percentage of land within the circle or if the district's land occupies a majority of the area of the circle. In this definition, a station can treat multiple districts when the former is not equal to the latter.[40] The results are very similar to the baseline results, indicating that the treatment definition is not sensitive to geographic concerns.

Rail construction decisions were likely to be more similar in a twenty-year window than across a forty or fifty year window. Removing districts treated after 1861 (the "Never Treated" districts) from the control group does little to change the results. The resulting control group is smaller, and so prone to larger errors, but is also a more similar comparison to the treated group as the vast majority of treated districts were treated between 1839-1861. This control group, comprised only of "Not Yet Treated" districts, and the treated group are therefore close to each other across time of treatment.

Parliamentary approval was required for most rail construction, leading to the possibility of political interference in the planning and construction process. Esteves and Geisler Mesevage (2021) find that Members of Parliament with connections to a rail project managed to get their projects approved more often than rail projects which did not have politically influential backers. The authors use the 1845 Board of Trade recommendations as a placebo test for their findings. The recommendations were issued only in the 1845 session by the Board's team of engineers who evaluated projects on their overall efficiency and feasibility. The Board of Trade engineers' recommendations were viewed as so impartial that rail firms used their political power to prevent the Board of Trade from issuing recommendations ever again (Casson 2009). Of the 68 rail projects that were evaluated by the board and built within the next fifteen years, 51 were approved by the board. A majority of the rejected rail lines were built in districts that were already treated and therefore have no effect on the difference in difference term in the model (which only considers first treatment).[41] The fact that so many of these projects were rejected offers some confirmation on the decision making process by the Board of Trade's engineers - building many lines in the same place is not necessarily efficient. I estimate equation (1) and include an interaction term between the rail treatment variable D and an indicator equal to one if the district was treated by a line approved by the Board of Trade. The coefficient of the interaction term is small and negative and with a large error term. This indicates that the districts treated by the Board of Trade approved rail lines did not exhibit different patenting effects. In the broader context, this implies that districts that were treated by rail lines that were plausibly free of political considerations had a similar treatment effect as districts treated by rail lines that were more exposed to potential political interference.

Rail construction in the 1840s saw two large spikes in 1840 and 1847, a period coinciding with the height of speculative interest in rail ("Rail Mania"). It is plausible that districts treated in these years are more likely

[38] Using the full 1823-1861 sample with an interaction between an indicator equal to one during the post Patent Law Amendment Act years and the local rail station variable yields a statistically insignificant coefficient, offering additional evidence that the effect of a rail station is not dependent on the change in patenting cost.

[39] Each circle encompasses a 28.27KM area.

[40] This can arise when a smaller district is fully within the circle, yet a larger district makes up most of the surrounding land.

[41] Only three districts are treated by the rejected lines.

to be exogenous - speculative booms are characterized by a decrease in rational behavior and functionality of markets. I estimate equation (1) and include an interaction term between the rail treatment variable D and an indicator equal to one if the district was treated in 1839, 1840, 1846, or 1847. The coefficient is positive, but statistically insignificant, indicating no meaningful difference between the districts treated during the peak Rail Mania years and districts treated in any other year.

I create a balanced panel dataset of individual inventors from 1823-1851 (the years in which I can match inventors by full name). The dependent variable is an indicator equal to one if inventor p applies for a patent in year t. Fixed effects for year and individual are included in the estimation equation as is a vector of population controls for district i. Treatment D is equal to one if the district in which the inventor lives in year t has a rail station.

$$Patents_{i,t} = \lambda_t + \gamma_p + \beta_1 * D_{i,t} + \beta_2 * Pop_{i,t} + \epsilon_{i,t} \tag{10}$$

Inventors are matched across time by full name. Addresses are only known for years in which an inventor patents. In all other years, inventors are assigned districts by the known district nearest in time. For example, if an inventor patents in 1830 in Bradford and 1840 in Clun, the inventor's district i is Bradford during the period 1823-1835 and Clun during 1836-1851. The model is estimated using an OLS estimation which clusters errors by inventor. The results imply a slightly lower patenting elasticity of 43% than the elasticity of 58% implied by the district level model for the same time period.

Rail innovation is potentially local to rail construction, and I therefore exclude all rail related patents from my main data sample to avoid this potential correlation. Adding back in the rail patents that were removed from the data only marginally changes the coefficient. It is also possible that industries which are adjacent to rail also exhibit spillover effects from rail construction. For example, steam engines were used to produce power for both trains and factories and it is possible that the same invention could be used in both sectors. While I remove steam engine related patents which directly mention rail technology, I generally retain other engine patents which do not. As an additional robustness check, I remove all patents which are classified as part of the engine and carriages sectors. The results are almost identical to the baseline estimates. Removing agricultural patents to account for compositional differences between the "Main" treatment definition and the full set of districts somewhat increases the coefficient estimates, but the results are broadly similar.

To assess whether the results are specific to patenting, and therefore potentially prone to some error in the patenting data, I assess the effect of a rail station on local population. I modify equation (1) by changing the dependent variable to local population:

$$log(Pop_{i,t}) = \alpha_i + \lambda_t + \beta D_{i,t} + \epsilon_{i,t} \tag{11}$$

I find that a local rail station increases local population by 2.9%, indicating that the rail station shifts other economic variables, rather than just having an effect on patenting.[42]

Though patents are widely used to measure innovation, they may be an imperfect proxy. For example, individuals may have patented low quality or useless inventions in the hope of reaping money through the legal system. To a large extent, this issue is accounted for by measuring the effect on high quality patents as I do in Table 2. Using data from the 1851 World's Fair (Moser 2012), I use exhibitions as an alternative

[42]I compare this result with the results found by Bogart et al. (2022) by dividing 2.9% by the mean number of treatment years (13.5) and I find a mean annual growth rate of .22%. While this number is smaller than the .87% rate that Bogart et al. (2022) find from 1851-1891, the authors use a different estimation technique, time period, and geographic treatment area. Given these differences, the difference in magnitude is not surprising.

proxy for innovation. As the data are only for the 1851 Fair, potential analysis is limited. However, I find a positive correlation between districts that have a rail station in either 1850 or 1851 and the number of exhibitions per capita that the district had at the Fair.

6.1 Postage Act

The introduction of the rail network coincided with the Postage Act of 1839 which made the cost of sending mail uniform, replacing the previous distance-based system. Hanlon et al. (2022) posit that the decreased cost of communication caused by the Postage Act led to an increase in innovation independent of a rail effect. Given the shared timing of these events, it is possible that my findings are misidentified and that the knowledge flows caused by the lower cost of mail are driving the results.

I parse the effects by exploiting the fact that a single patent office in London served all of England. Individuals who were not in London could mail their applications to the patent office. If the increase in patenting which occurs following a rail connection is caused by the decreased cost of posting the patent application to London, one would expect the mean distance between an inventor and London to increase as inventors further afield see a decrease in their patenting costs. However, regressing patents per capita against the triple interaction of a local rail station, the distance between an inventor and London, and an indicator equal to one for every year following the Postage Act yields a negative, weakly significant coefficient. This suggests that the elimination of distance-based mail fees is not a confounding factor as distance to London is not positively correlated with patenting activity.[43]

Following the 1852 patent reform, which decreased the overall cost of patenting, the number of patent applications greatly increased. This is in line with an inverse relationship between the cost of patenting and patenting. However, Figure 20 shows that there is no change in the mean distance between coauthors following the 1839 Postage Act until 1852, whereas an increase in distance is expected if the cost of communication across larger distances decreases and is relevant to patenting. This implies that the lowered cost of written communication from the Postage Act is not having a conflating effect.

The discrepancy between these results and the results that Hanlon et al. (2022) find could perhaps be due to the authors' focus on basic research rather than final inventions. Another possibility is that in-person communication - the type of knowledge flows that might be expected from a decrease in the cost of rail tickets - interacts with patenting activity in a distinct manner than mail correspondence.

7 Conclusion

I find that the English railroad network had a transformative impact on innovation. Exploiting the manner in which the railroad was built to study plausibly exogenously treated districts, I find large returns to local patenting: having a rail station causes a 77% increase in a district's patents per capita. By comparison, the railroad increases local population by less than 3%. The returns to patenting do not appear to be driven by sorting among inventors. Rather, using annual transportation network maps that I created to measure the cost of moving between districts along roads, canals, and railroads, I find that the construction of the rail network causes an increase in local market access, which in turn causes an increase in local patenting. 43% of the local rail station effect is explained by the increase in market access. I also find evidence that the railway causes an increase in patenting due to fostering knowledge flows and that this mechanism is

[43]The coefficient from regressing patenting against the interaction between having a local rail station and the distance to London is effectively zero.

distinct from the market access channel. The returns from knowledge flows disproportionately increase innovation in large towns which house "large enough" relevant knowledge bases and the returns from trade disproportionately increase innovation in locations with more manufacturing jobs. These findings fill a gap in the literature - previous research on transportation networks find a knowledge flows channel, but no papers explore the market access mechanism. My results indicate that these papers are potentially missing a large channel through which transportation networks affect innovation. This has important implications both for improving economists' understanding of innovation and for improving measurements of the gains from static trade.

Future work includes: obtaining data on local trade flows to more precisely measure the imports and exports of a district and how this changes with rail, disentangling the market access mechanism from a potential financial constraint channel, matching the data to 21st century data on local employment and patenting to see if an earlier rail connection produced effects that were so durable as to have permanently shifted the sectoral composition of a town, and obtaining data on Royal Society of Arts awards as a robustness measure in defining innovation.

References

Acemoglu, D. and Linn, J. (2004). Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry. *The Quarterly Journal of Economics*, 119(3):1049–1090.

Adsera, A. and Ray, D. (1998). History and coordination failure. *Journal of Economic Growth*, 3(3):267–76.

Aghion, P., Bergeaud, A., Lequien, M., and Melitz, M. J. (2024). The Heterogeneous Impact of Market Size on Innovation: Evidence from French Firm-Level Exports. *The Review of Economics and Statistics*, 106(3):608–626.

Aghion, P., Bergeaud, A., and Van Reenen, J. (2023). The impact of regulation on innovation. *American Economic Review*, 113(11):2894–2936.

Agrawal, A., Galasso, A., and Oettl, A. (2017). Roads and Innovation. *The Review of Economics and Statistics*, 99(3):417–434.

Allen, R. C. (1983). Collective invention. *Journal of Economic Behavior Organization*, 4(1):1–24.

Américo, P. (2022). The Industrialization Paths: Railroads and Structural Transformation in Brazil 1872-1950.

Andersson, D., Berger, T., and Prawitz, E. (2023). Making a Market: Infrastructure, Integration, and the Rise of Innovation. *The Review of Economics and Statistics*, 105(2):258–274.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., and Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118.

Atack, J., Bateman, F., Haines, M., and Margo, R. A. (2010). Did railroads induce or follow economic growth? urbanization and population growth in the american midwest, 1850-1860. *Social Science History*, 34(2):171–197.

Atack, J., Bateman, F., Haines, M., and Margo, R. A. (2011). Railroads and the Rise of the Factory: Evidence for the United States, 1850-1870. In P. Rhode, J. Rosenbloom, D. W., editor, *Economic Evolution and Revolution in Historical Time*. Stanford University Press.

Atkin, D. G., Chen, K., and Popov, A. (2022). The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley. *NBER Working Paper No. w30147*.

Autor, D., Dorn, D., Hanson, G. H., Pisano, G., and Shu, P. (2020). Foreign Competition and Domestic Innovation: Evidence from US Patents. *American Economic Review: Insights*, 2(3):357–74.

Aw, B. Y., Roberts, M. J., and Xu, D. Y. (2011). R&D Investment, Exporting, and Productivity Dynamics. *American Economic Review*, 101(4):1312–44.

Banerjee, A., Duflo, E., and Qian, N. (2020). On the Road: Access to Transportation Infrastructure and Economic Growth in China. *Journal of Development Economics*, 145:102442.

Baum-Snow, N. (2007). Did Highways Cause Suburbanization? *The Quarterly Journal of Economics*, 122(2):775–805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M., and Zhang, Q. (2017). Roads, Railroads, and Decentralization of Chinese Cities. *The Review of Economics and Statistics*, 99(3):435–448.

Baum-Snow, N., Henderson, J. V., Turner, M., Zhang, Q., and Brandt, L. (2020). Does Investment in National Highways Help or Hurt Hinterland City Growth? *Journal of Urban Economics*, 115(C):S0094119018300287.

Bloom, N., Draca, M., and Van Reenen, J. (2016). Trade Induced Technical Change? The Impact of Chinese Imports on Innovation, IT and Productivity. *The Review of Economic Studies*, 83(1):87–117.

Bogart, D. (2005). Turnpike trusts and the transportation revolution in 18th century england. *Explorations in Economic History*, 42(4):479–508.

Bogart, D., You, X., Alvarez-Palau, E. J., Satchell, M., and Shaw-Taylor, L. (2022). Railways, Divergence, and Structural Change in 19th Century England and Wales. *Journal of Urban Economics*, 128(C).

Bombardini, M., Li, B., and Wang, R. (2017). Import Competition and Innovation: Evidence from China. Technical report.

Bottasso, A., Conti, M., Robbiano, S., and Santagata, M. (2022). Roads to innovation: Evidence from Italy. *Journal of Regional Science*, 62(4):981–1005.

British Parliamentary Papers (1854a). Return in a Tabular Form of the Amount of the Esimates of the Annual Value of Real Property in each Parish, as Assessed to the Property Tax in April 1815.

British Parliamentary Papers (1854b). Return in a Tabular Form of the Amount of the Esimates of the Annual Value of Real Property in each Parish, as Assessed to the Property Tax in April 1843.

British Parliamentary Papers (1860). Return of gross Annual Value of Property assessed under Schedules (A.) and (B.) of Income Tax in England, Wales and Scotland; 1859-60, Return of net Amount of Profits assessed under Schedule (D.). 536.

Brynjolfsson, E., Hu, Y. J., and Smith, M. D. (2003). Consumer surplus in the digital economy: Estimating the value of increased product variety at online booksellers. *Management Science, Forthcoming*.

Buera, F. J. and Oberfield, E. (2020). The Global Diffusion of Ideas. *Econometrica*, 88(1):83–114.

Bustos, P. (2011). Trade Liberalization, Exports, and Technology Upgrading: Evidence on the Impact of MERCOSUR on Argentinian Firms. *American Economic Review*, 101(1):304–40.

Büchel, K. and Kyburz, S. (2020). Fast track to growth? railway access, population growth and local displacement in 19th century switzerland. *Journal of Economic Geography*, 20(1):155–195.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

Casson, M. (2009). *The World's First Railway System: Enterprise, Competition, and Regulation on the Railway Network in Victorian Britain*. Number 9780199213979 in OUP Catalogue. Oxford University Press.

Catalini, C., Fons-Rosen, C., and Gaulé, P. (2020). How Do Travel Costs Shape Collaboration? *Management Science*, 66(8):3340–3360.

Census of Great Britain (1852). Population Tables I. Number of Inhabitants in the years 1801, 1811, 1821, 1831, 1841 and 1851.

Chandra, A. and Thompson, E. (2000). Does Public Infrastructure Affect Economic Activity?: Evidence from the Rural Interstate Highway System. *Regional Science and Urban Economics*, 30(4):457–490.

Chen, J. and Roth, J. (2023). Logs with Zeros? Some Problems and Solutions. *The Quarterly Journal of Economics*, 139(2):891–936.

Cobb, M. (2015). *The Railways of Great Britain: A Historical Atlas*. Riley Dunn & Wilson Ltd.

Coluccia, D. M. and Dossi, G. (2023). Return Innovation: The Knowledge Spillovers of the British Migration to the United States, 1870-1940. *Working Paper*.

Damijan, J., Kostevc, C., and Polanec, S. (2010). From innovation to exporting or vice versa? *The World Economy*, 33(3):374–398.

Davis, D. R. and Dingel, J. I. (2019). A Spatial Knowledge Economy. *American Economic Review*, 109(1):153–170.

Donaldson, D. (2018). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, 108(4-5):899–934.

Donaldson, D. and Hornbeck, R. (2016). Railroads and American Economic Growth: A "Market Access"

Approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Duranton, G. and Puga, D. (2004). Chapter 48 - Micro-Foundations of Urban Agglomeration Economies. In Henderson, J. V. and Thisse, J.-F., editors, *Cities and Geography*, volume 4 of *Handbook of Regional and Urban Economics*, pages 2063–2117. Elsevier.

Esteves, R. and Geisler Mesevage, G. (2021). Private Benefits, Public Vices: Railways and Logrolling in the Nineteenth-Century British Parliament. *The Journal of Economic History*, 81(4):975–1014.

Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *The Review of Economic Studies*, 81(3):1046–1070.

Fogel, R. W. (1964). *Railroads and American Economic Growth: Essays in Econometric History*. Johns Hopkins University Press.

Furman, J. L., Porter, M. E., and Stern, S. (2002). The determinants of national innovative capacity. *Research Policy*, 31(6):899–933.

Garcia-López, M.-A., Holl, A., and Viladecans-Marsal, E. (2015). Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics*, 85(C):52–67.

Ghani, E., Goswami, A. G., and Kerr, W. R. (2016). Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing. *The Economic Journal*, 126(591):317–357.

Gibbons, S., Lyytikäinen, T., Overman, H. G., and Sanchis-Guarner, R. (2019). New Road Infrastructure: The Effects on Firms. *Journal of Urban Economics*, 110:35–50.

Glaeser, E. L. and Gottlieb, J. D. (2009). The wealth of cities: Agglomeration economies and spatial equilibrium in the united states. *Journal of Economic Literature*, 47(4):983–1028.

Gregory, I. and Marti, J. (2010). The railways, urbanization, and local demography in england and wales, 1825-1911. *Social Science History*, 34:199–228.

Hanlon, W. W., Heblich, S., Monte, F., and Schmitz, M. B. (2022). A penny for your thoughts. Working Paper 30076, National Bureau of Economic Research.

Hanlon, W. W. and Miscio, A. (2017). Agglomeration: A long-run panel data approach. *Journal of Urban Economics*, 99:1–14.

Henneberg, J., Satchell, M., You, X., Shaw-Taylor, L., and Wrigley, E. (2018a). 1851 England, Wales and Scotland Railway Stations [Data Collection]. *UK Data Service. SN: 852994*.

Henneberg, J., Satchell, M., You, X., Shaw-Taylor, L., and Wrigley, E. (2018b). 1861 England, Wales and Scotland Rail Lines [Data Collection]. *UK Data Service. SN: 852992*.

Henneberg, J., Satchell, M., You, X., Shaw-Taylor, L., Wrigley, E., and Cobb, M. (2018c). 1861 England, Wales and Scotland Railway Stations [Data Collection]. *UK Data Service. SN: 852995*.

Holl, A. (2016). Highways and productivity in manufacturing firms. *Journal of Urban Economics*, 93(C):131–151.

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

Horrell, S., Humphries, J., and Weale, M. (1994). An Input-Output Table for 1841. *The Economic History Review*, 47(3):545–566.

Jaffe, A. B., Trajtenberg, M., and Henderson, R. (1993). Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations. *The Quarterly Journal of Economics*, 108(3):577–598.

Jaworski, T., Kitchens, C., and Nigai, S. (2023). Highways and globalization. *International Economic Review*, 64(4):1615–1648.

Juhász, R. and Steinwender, C. (2023). Industrial Policy and the Great Divergence. *NBER Working Paper No. w31736*.

Kline, P. and Moretti, E. (2014). Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. *The Quarterly Journal of Economics*, 129(1):275–331.

Leeuwen, M. V. and Maas, I. (2011). *Hisclass: A Historical International Social Class Scheme*. Leuven University Press.

Lileeva, A. and Trefler, D. (2010). Improved access to foreign markets raises plant-level productivity...for some plants. *The Quarterly Journal of Economics*, 125(3):1051–1099.

Melitz, M. J. and Redding, S. J. (2014). Chapter 1 - heterogeneous firms and trade. In Gopinath, G., Helpman, E., and Rogoff, K., editors, *Handbook of International Economics*, volume 4 of *Handbook of International Economics*, pages 1–54. Elsevier.

Melitz, M. J. and Redding, S. J. (2023). *Trade and Innovation*. Harvard University Press.

Michaels, G. (2008). The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System. *The Review of Economics and Statistics*, 90(4):683–701.

Moser, P. (2012). Innovation without Patents: Evidence from World's Fairs. *The Journal of Law Economics*, 55(1):43–74.

Moser, P. (2013). Patents and innovation: Evidence from economic history. *Journal of Economic Perspectives*, 27(1):23–44.

Mullahy, J. and Norton, E. C. (2022). Why transform y? a critical assessment of dependent-variable transformations in regression models for skewed and sometimes-zero outcomes. Working Paper 30735, National Bureau of Economic Research.

Nasmyth, J. (1897). *James Nasmyth, Engineer: An Autobiography*. John Murray.

Natural England (2023). National Historic Landscape Characterisation 250m Grid (England).

Nuvolari, A., Tartari, V., and Tranchero, M. (2021). Patterns of innovation during the industrial revolution: A reappraisal using a composite indicator of patent quality. *Explorations in Economic History*, 82:101419.

Oxford Dictionary of National Biography (2024). Oxford University Press.

Pauly, S. and Stipanicic, F. (2024). The creation and diffusion of knowledge: Evidence from the Jet Age.

Perlman, E. (2017). Dense enough to be brilliant: Patents, Urbanization, and Transportation in Nineteenth Century America. *Working Paper*.

Pogonyi, C. G., Graham, D. J., and Carbo, J. M. (2021). Metros, agglomeration and displacement. Evidence from London. *Regional Science and Urban Economics*, 90(C).

Purdy, F. (1860). The statistics of the english poor rate before and since the passing of the poor law amendment act. *Journal of the Statistical Society of London*, pages 286–329.

Railway Commission (1848). *Return of the Passenger and Goods Traffic on each Railway in Great Britain and Ireland for the Two Years ending 30th July 1846 and 30th June 1847*. William Clowes and Sons, Stamford Street.

Redding, S. and Turner, M. (2015). Transportation costs and the spatial organization of economic activity. volume 5, chapter Chapter 20, pages 1339–1398. Elsevier.

Redding, S. and Venables, A. (2004). Economic geography and international inequality. *Journal of International Economics*, 62(1):53–82.

Rosevear, A., Satchell, M., Bogart, D., Shaw Taylor, L., Aidt, T., and Leon, G. (2017). Turnpike roads of

england and wales, 1667-1892.

Satchell, A., Kitson, P., Newton, G., Shaw-Taylor, L., and Wrigley, E. (2023). 1851 England and Wales census parishes, townships and places. [Data Collection]. *Colchester, Essex: UK Data Archive*.

Satchell, M. and Shaw-Taylor, L. (2018). 1851 england and wales navigable waterways . [data collection]. *UK Data Service. SN: 852998*.

Satchell, M., Wrigley, E., Shaw-Taylor, L., You, X., and Henneberg, J. (2018). 1851 England, Wales and Scotland Rail Lines [Data Collection]. *UK Data Service. SN: 852991*.

Schmookler, J. (1966). *Invention and Economic Growth*. Harvard University Press.

Schurer, K. and Higgs, E. (2023). Integrated Census Microdata (I-CeM), 1851-1911. [data collection].

Shaw-Taylor, L. and You, X. (2018). *The Online Historical Atlas of Transport, Urbanization and Economic Development in England and Wales c.1680-191*, chapter The development of the railway network in Britain 1825-1911.

Sokoloff, K. L. (1988). Inventive activity in early industrial america: Evidence from patent records, 1790-1846. *The Journal of Economic History*, 48(4):813–850.

Storper, M. and Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, 4(4):351–370.

Tang, J. P. (2014). Railroad expansion and industrialization: Evidence from meiji japan. *The Journal of Economic History*, 74(3):863–886.

The National Archives (2023). Intellectual property: Patents of invention. *https://www.nationalarchives.gov.uk/help-with-your-research/research-guides/patents-of-invention*.

UK Parliament (2024). Railways in Early Nineteenth Century Britain - UK Parliament.

Wolmar, C. (2008). *Fire and Steam: A New History of the Railways in Britain*. Atlantic Books.

Woodcroft, B. (1854). *Titles of Patents of Invention, Chronologically Arranged: From March 2, 1617 (14 James I.) to October 1, 1852 (16 Victoriae)*. Titles of Patents of Invention: Chronologically Arranged from March 2, 1617 (14 James I.) to October 1, 1852 (16 Victoriæ). G.E. Eyre & W. Spottiswoode.

Yamasaki, J. (2017). Railroads, Technology Adoption, and Modern Economic Development: Evidence from Japan. ISER Discussion Paper 1000, Institute of Social and Economic Research, Osaka University.

You, X., Bogart, D., Alvarez-Palau, E., Satchell, M., and Shaw-Taylor, L. (2021). Transport Development and Urban Population Change in the Age of Steam: A Market Access Approach.

Appendix

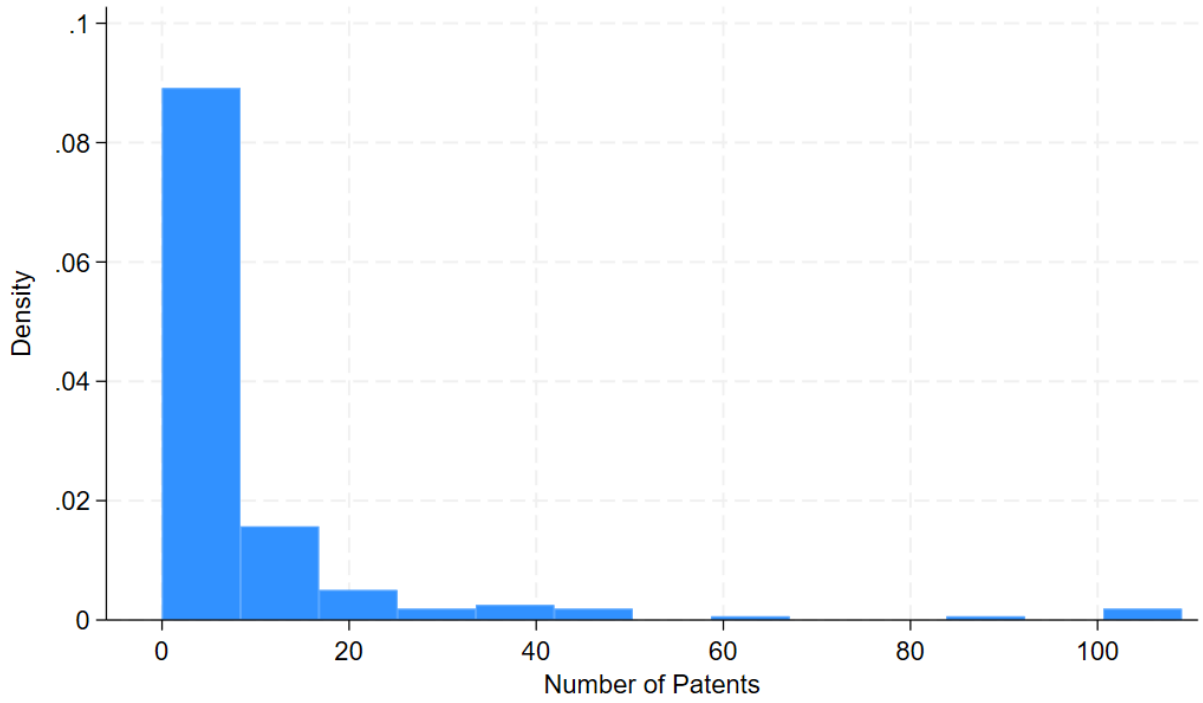


Figure 12: Histogram of patents produced in a district aggregated across 1823-1861. Only districts in the main treatment definition are included.

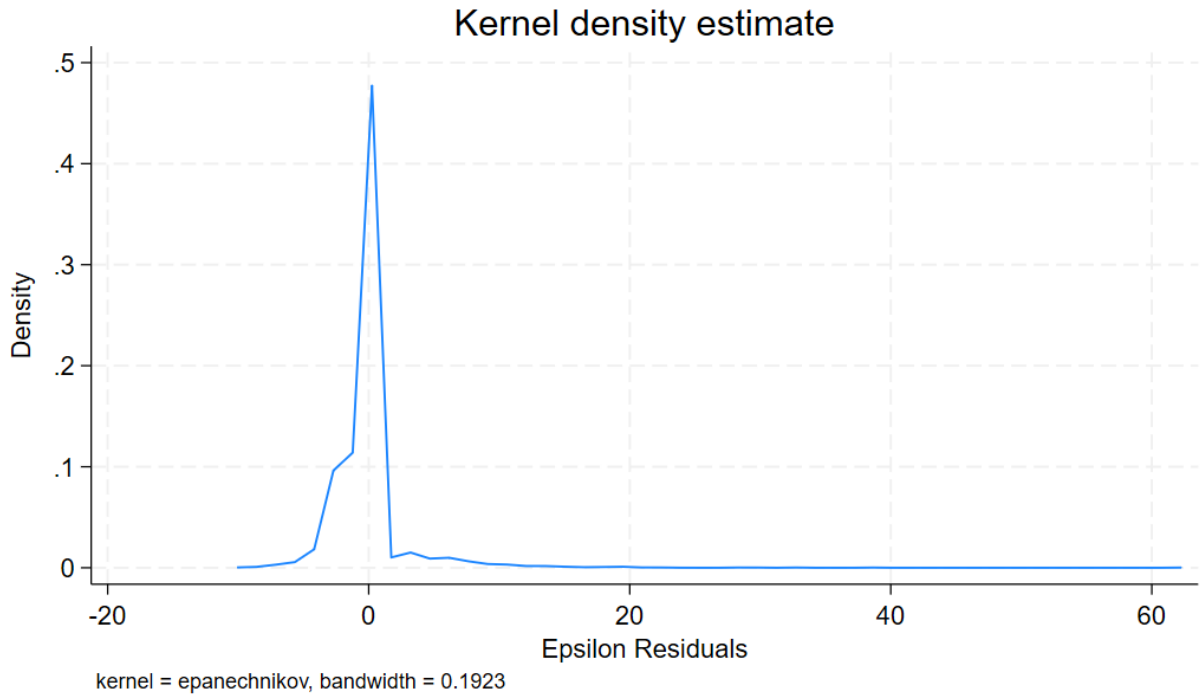


Figure 13: Kernal density graph of epsilon residuals from the OLS estimation of equation (1) where the dependent variable is patents per capita. Only districts in the main treatment definition are included.

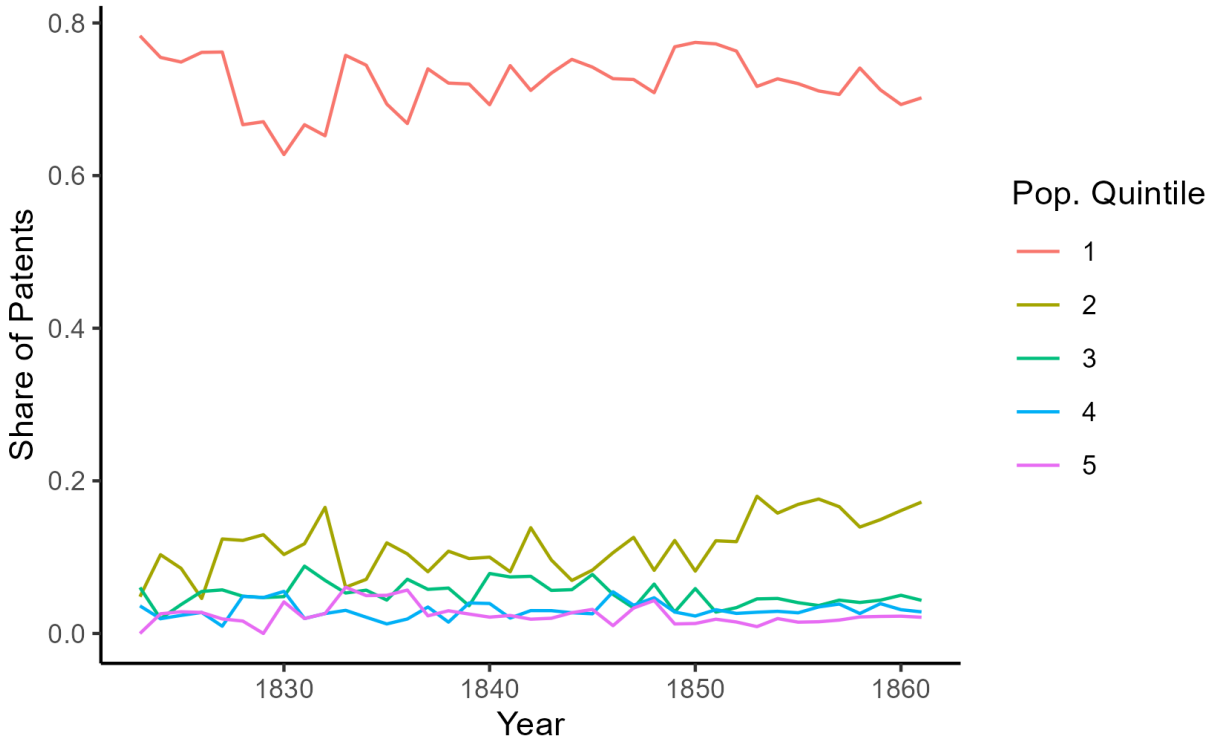


Figure 14: Share of patenting across time by the 1821 population deciles. The most populous quintile is "1" and the least populous is "5". Includes all districts.

Table 13: Patents Applicants by Sector

Sector	All Patents	Main Analysis Patents	Sh. of Nat., All	Sh. of Nat., Main	Sh. of Nat., All w/o Ag.	Sh. of Nat., Main w/o Ag.
barrels	217	12	0.9	0.72	0.95	0.84
bottle manufacture	18	1	0.07	0.06	0.08	0.07
carding	92	5	0.38	0.3	0.4	0.35
carpets	24	6	0.1	0.36	0.11	0.42
carriages	334	43	1.39	2.57	1.46	3.02
ceramic glass	409	31	1.7	1.85	1.79	2.17
chemistry	743	28	3.09	1.67	3.25	1.96
clocks	137	9	0.57	0.54	0.6	0.63
clothing	1958	111	8.15	6.63	8.58	7.78
construction	558	36	2.32	2.15	2.44	2.52
distilling	170	16	0.71	0.96	0.74	1.12
dyes	177	2	0.74	0.12	0.78	0.14
energy	653	35	2.72	2.09	2.86	2.45
engines	2275	153	9.47	9.14	9.97	10.73
fertilizer	88	7	0.37	0.42	0.39	0.49
fire safety	354	18	1.47	1.08	1.55	1.26
firearms	657	45	2.74	2.69	2.88	3.16
fluids water	131	14	0.55	0.84	0.57	0.98
food	391	41	1.63	2.45	1.71	2.88
furniture	452	29	1.88	1.73	1.98	2.03
gas liquid_meter	85	0	0.35	0	0.37	0
gas manufacture	18	0	0.07	0	0.08	0
implements goods_other	791	45	3.29	2.69	3.47	3.16
india rubber	88	9	0.37	0.54	0.39	0.63
infrastructure	137	12	0.57	0.72	0.6	0.84
light	352	11	1.47	0.66	1.54	0.77
locks	199	5	0.83	0.3	0.87	0.35
machinery	916	54	3.81	3.23	4.01	3.79
measuring navigation	310	15	1.29	0.9	1.36	1.05
medical	208	12	0.87	0.72	0.91	0.84
metallurgy	1627	91	6.78	5.44	7.13	6.38
mining	107	4	0.45	0.24	0.47	0.28
music	225	13	0.94	0.78	0.99	0.91
oils	301	10	1.25	0.6	1.32	0.7
paper printing	880	72	3.66	4.3	3.86	5.05
photography	60	6	0.25	0.36	0.26	0.42

plumbing pumps	466	19	1.94	1.14	2.04	1.33
preserving	51	4	0.21	0.24	0.22	0.28
rail	1130	76	4.71	4.54	4.95	5.33
refrigerators	36	0	0.15	0	0.16	0
rope	45	1	0.19	0.06	0.2	0.07
sea	21	1	0.09	0.06	0.09	0.07
shipbuilding	915	52	3.81	3.11	4.01	3.65
shoes	182	9	0.76	0.54	0.8	0.63
spinning	936	54	3.9	3.23	4.1	3.79
sugar	67	5	0.28	0.3	0.29	0.35
tanning leather	72	3	0.3	0.18	0.32	0.21
telegraphs	266	13	1.11	0.78	1.17	0.91
textiles	1052	58	4.38	3.46	4.61	4.07
timber	154	8	0.64	0.48	0.67	0.56
weaving	1148	115	4.78	6.87	5.03	8.06
weighing	20	5	0.08	0.3	0.09	0.35
wire	11	0	0.05	0	0.05	0
writing implements	113	2	0.47	0.12	0.5	0.14
agriculture	1185	248	4.94	14.81	NA	NA
Median	221	12.5	.92	.75	.95	.84

Note: Summary of patent applicants by their patents' sectors. Columns 1, 3, and 5 include patents from all districts. Columns 2, 4, and 6 only include patents in the districts included in the main treatment definition (and the never treated districts). Columns 1 and 2 report the total number of patents between 1823-1861 in each sector. Columns 3 and 4 report the sector's share of national patents. Columns 5 and 6 report the sector's share of national patents excluding agriculture from the calculation.

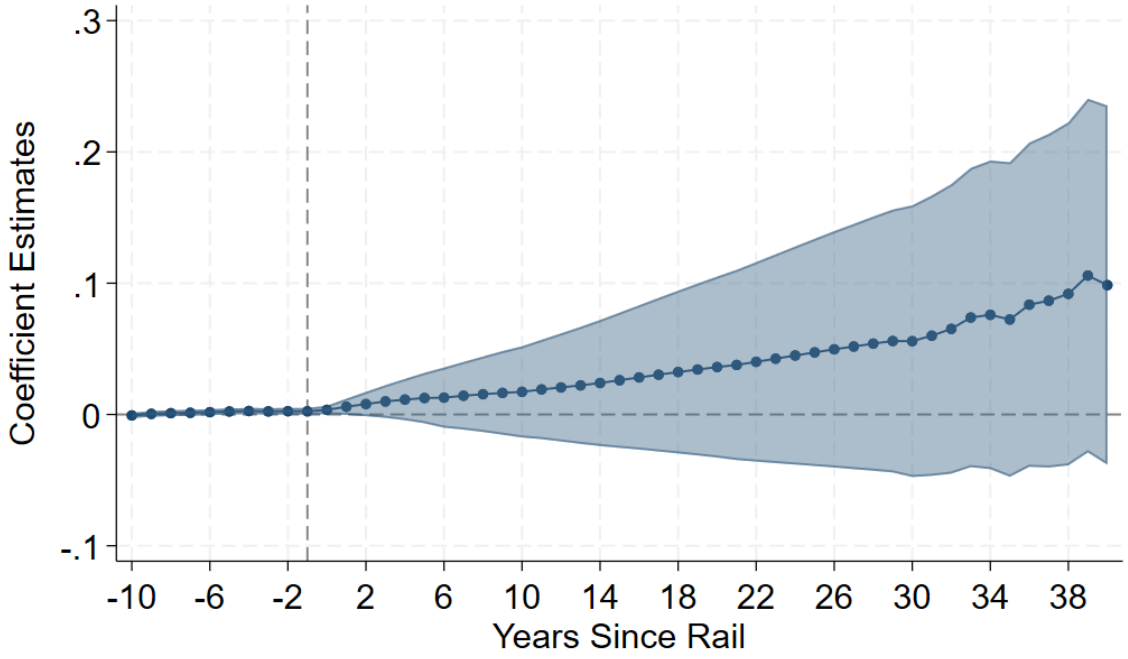


Figure 15: Event study plot with imputed population as the dependent variable and 1823-1891 as the time period. Districts are treated in the year in which they receive a rail station. Because of limitations to the rail data, districts which are not yet treated by 1861 are considered as "never treated" districts. I use the difference-in-difference estimator proposed by Callaway and Sant'Anna (2021) to correctly estimate the "never treated" control group.

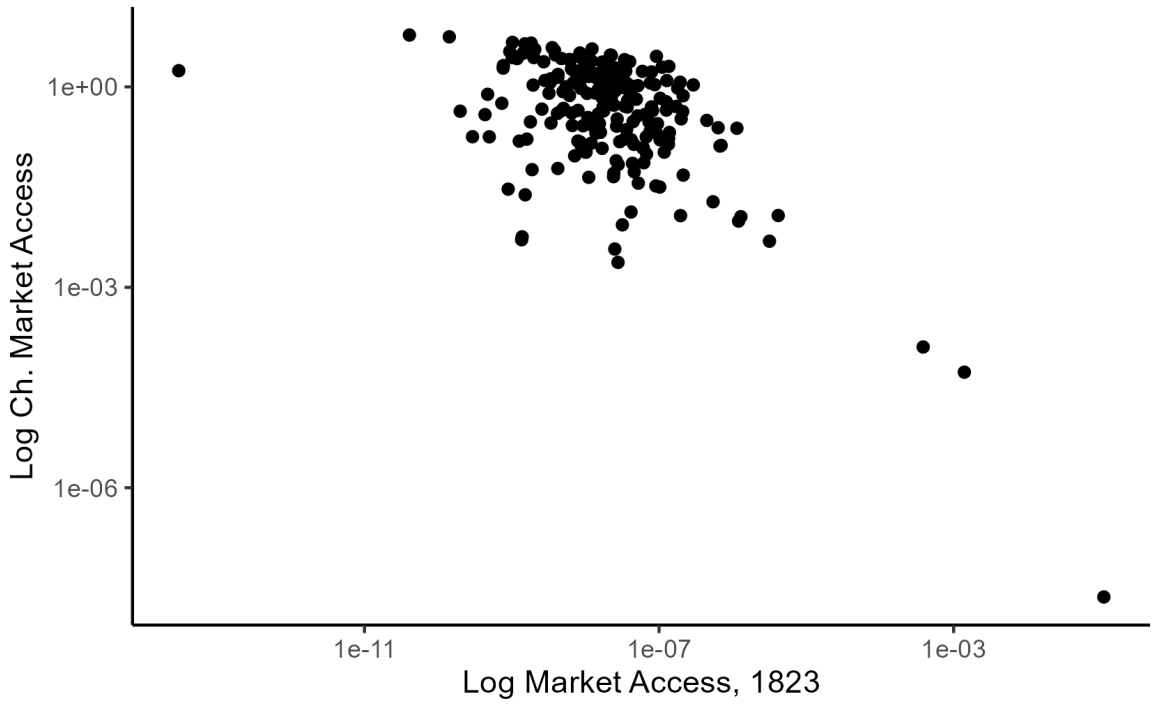


Figure 16: Scatter plot of the log increase in districts' market access between 1823-1861 by districts' log market access in 1823. Only districts in the main treatment definition are included.

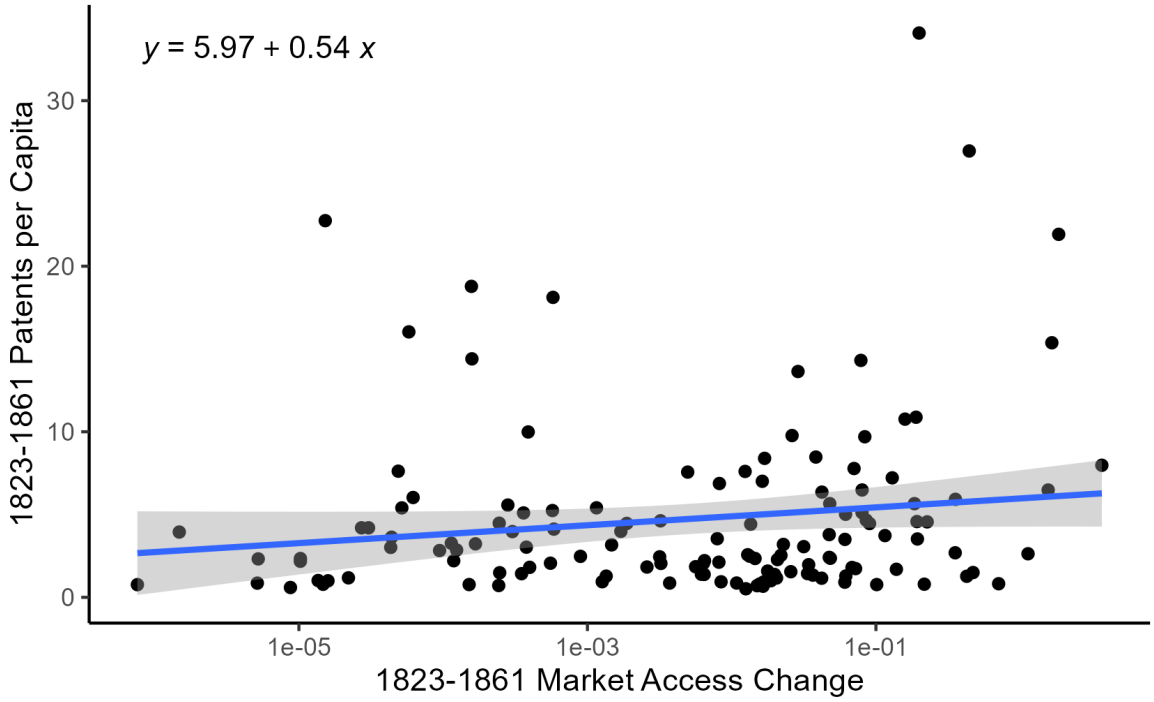


Figure 17: Scatter plot of the log increase in districts' market access between 1823-1861 by districts' log change in patents between 1823-1861. The stock of patents between 1824-1828 and 1857-1861 is used to measure patents at the beginning and end of the period. I use an IHS transformation to get the log of patents. Only districts in the main treatment definition are included. A fitted polynomial is drawn through the points dictated by the equation in the top left corner.

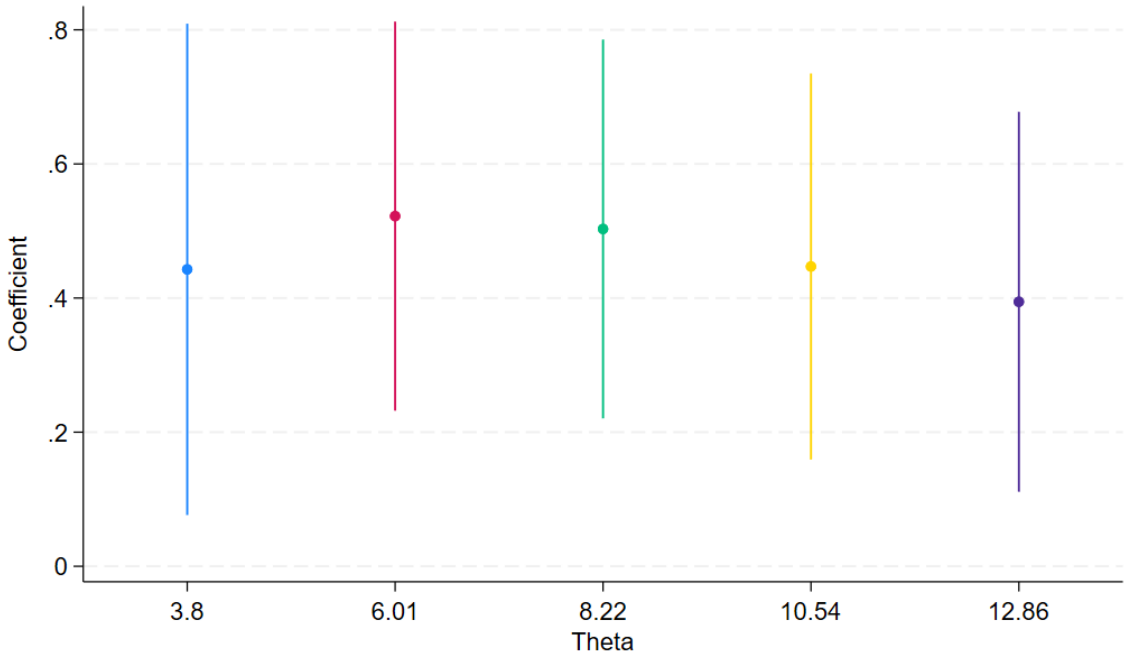


Figure 18: Estimation of equation (3) using different values of theta. Errors are clustered by district. 3.8 and 12.86 are the upper and lower bounds suggested by Donaldson and Hornbeck (2016) and 8.22 is the main value the authors use following a structural estimation of the parameter. 6.01 and 10.54 are chosen as midpoints between the boundary points and 8.22.

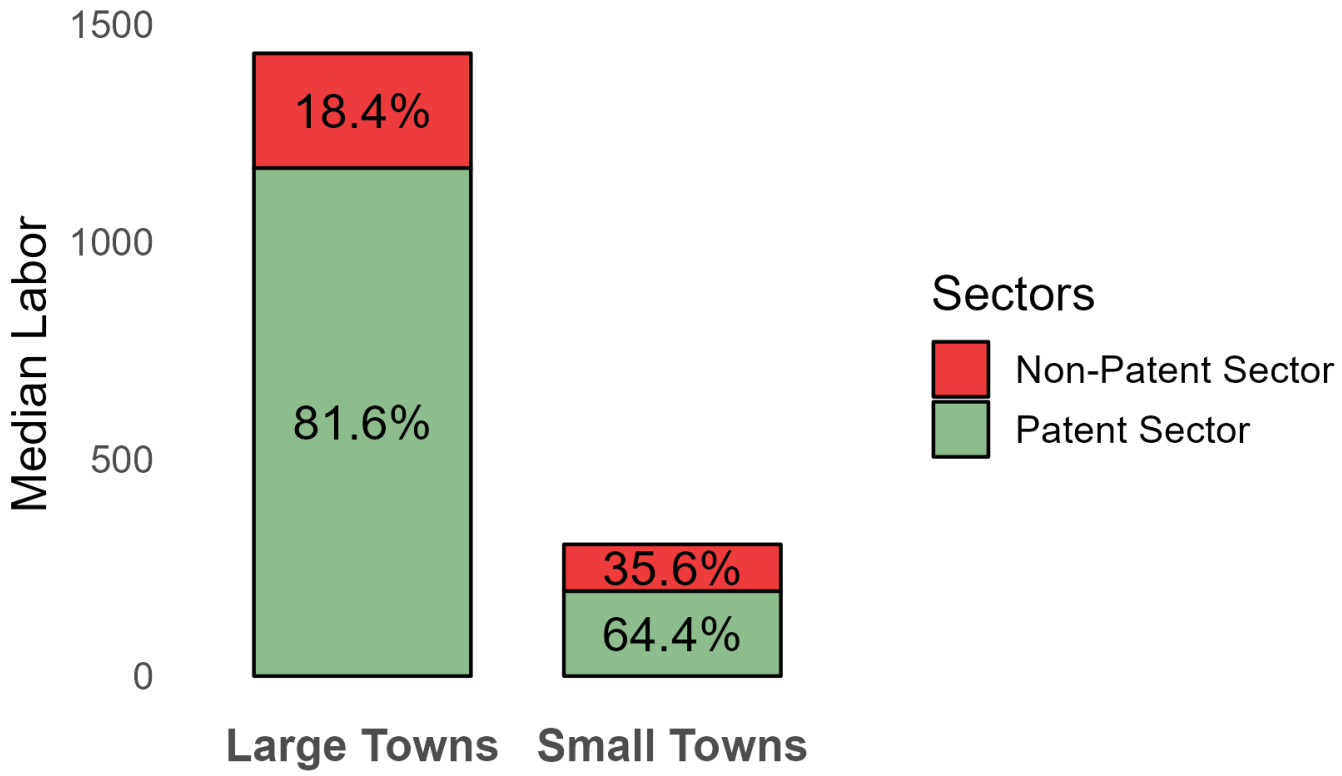


Figure 19: Median labor across districts and sectors divided by population size. The left bar included districts in the top quintile by 1831 population and and the right bar includes all other districts. The green bars include only labor within sectors that the district produces at least one patent between 1840-1861 and the red bars include the labor in all other sectors. Only districts in the main treatment sample are included and only sectors in the top half by number of total number of national patents across 1823-1861 are included.

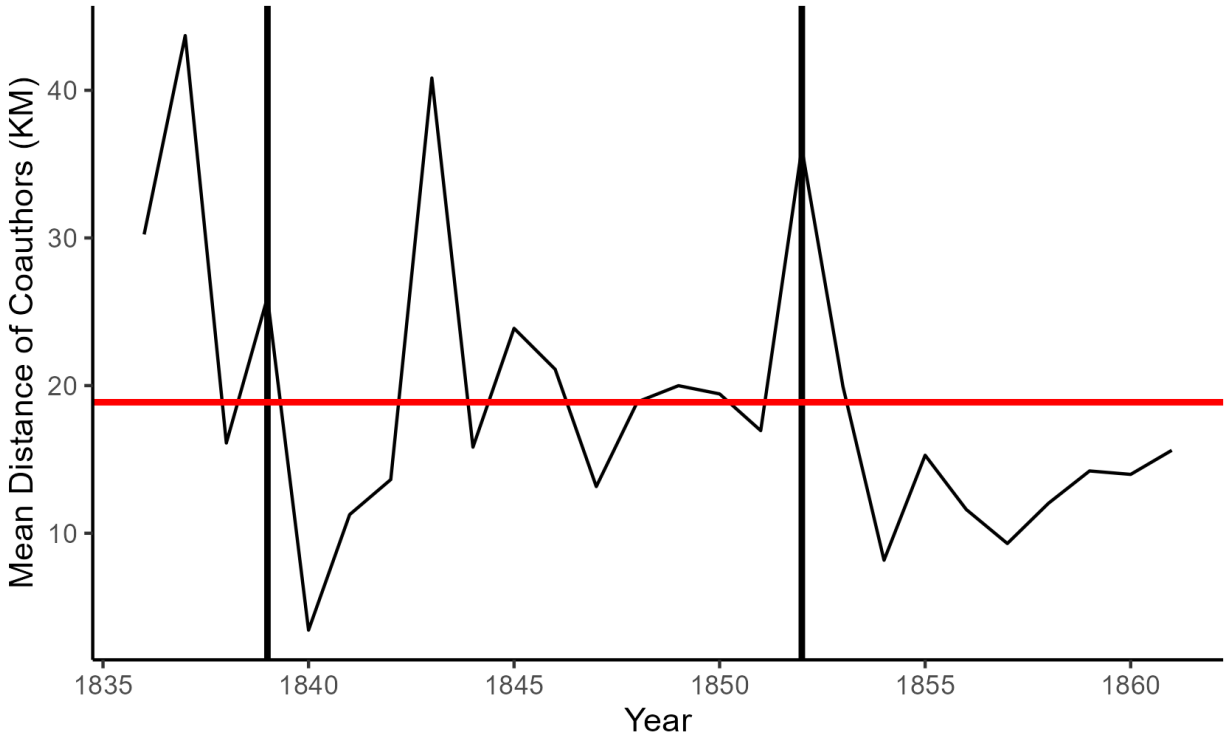


Figure 20: Mean distance between patent coauthors in years with at least 20 patents in which multiple authors are listed as applicants. The vertical line in 1838 represents the Postage Act of 1838 which decreased mail based communication costs and the vertical line in 1852 represents the Patent Law Amendment Act of 1852 which decreased the cost of patenting. The red horizontal line is the mean distance between coauthors across all years included in the graph.

Maps

Patents by District, 1823-1835

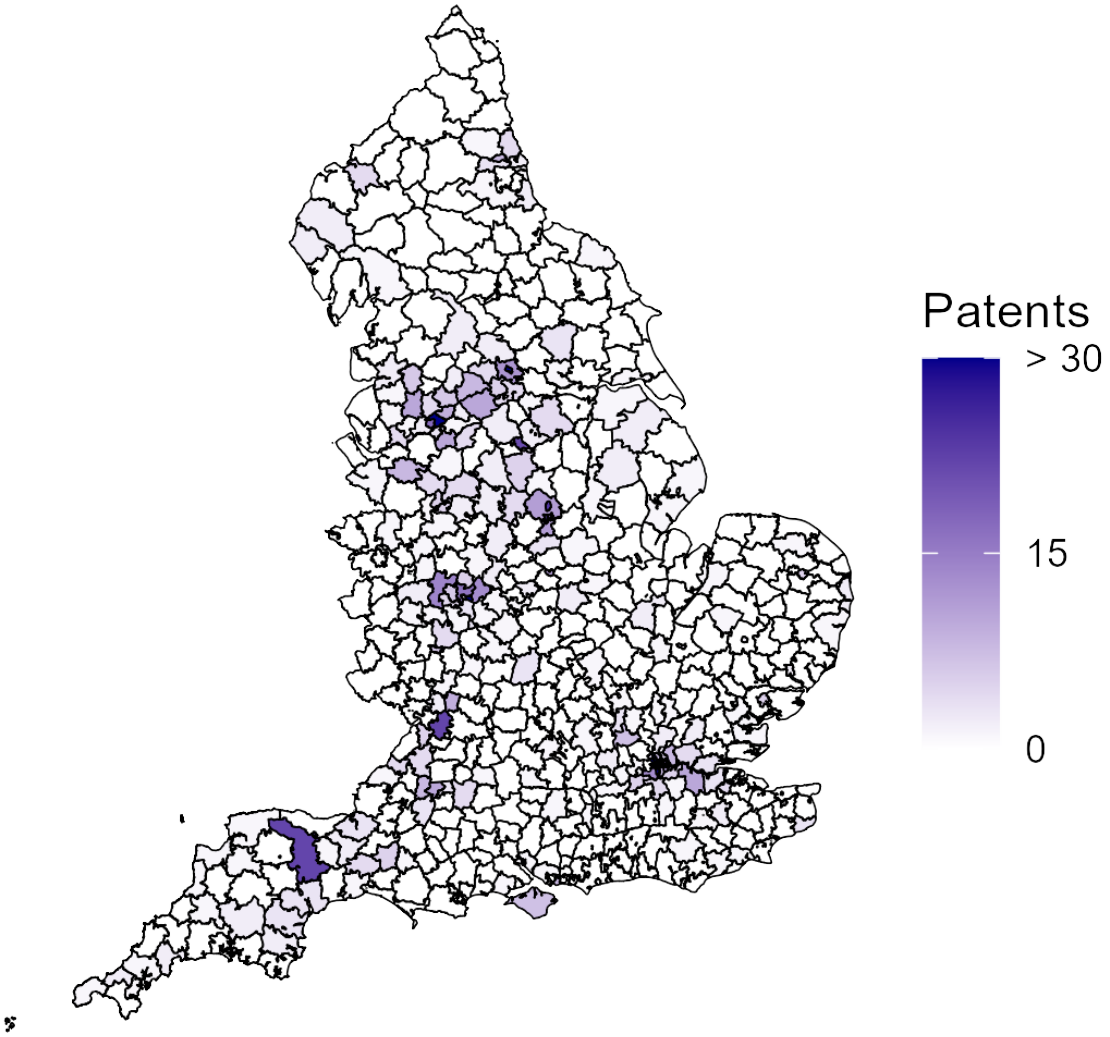


Figure 21: Heat map of all patents between 1823-1835 by district.

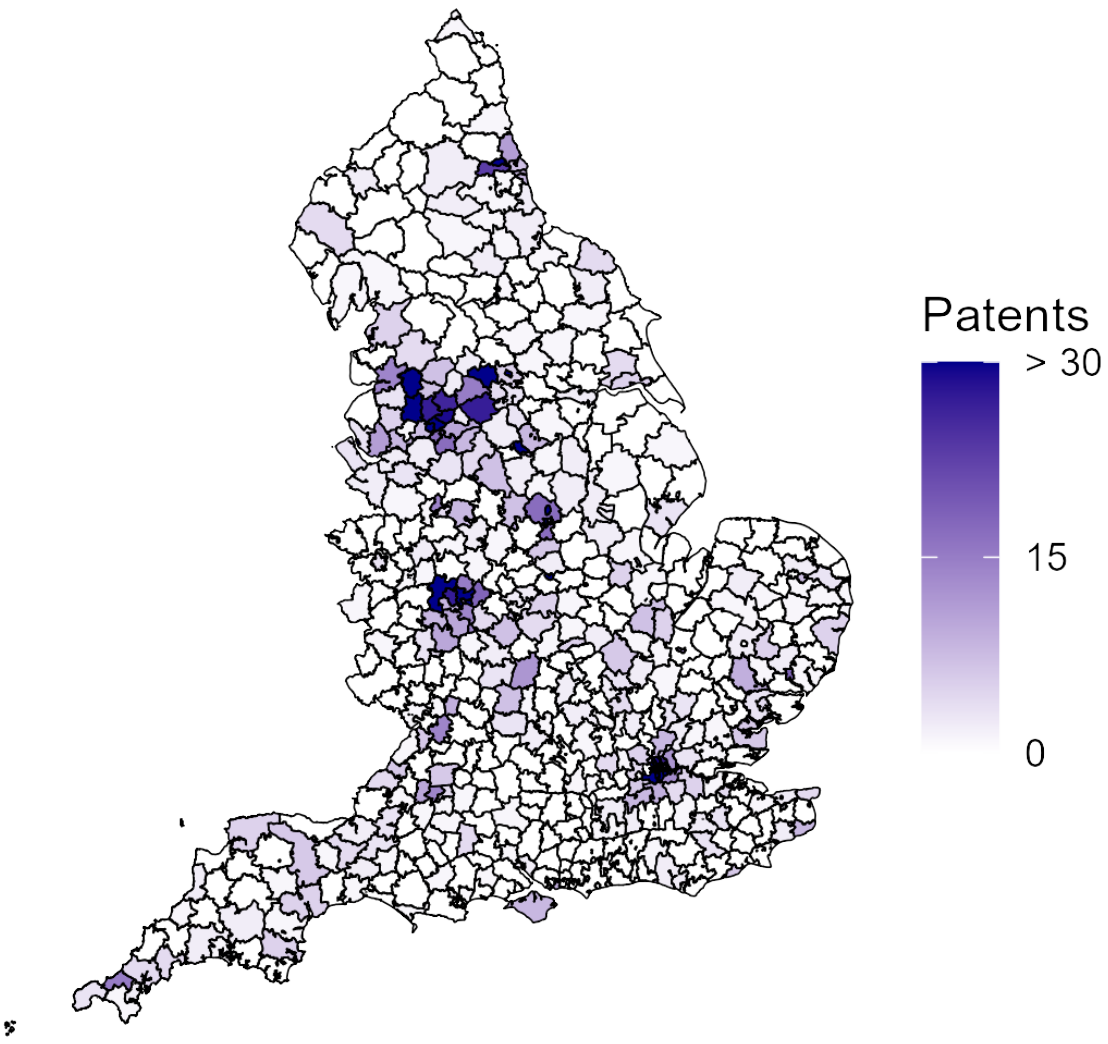


Figure 22: Heat map of all patents between 1836-1848 by district.

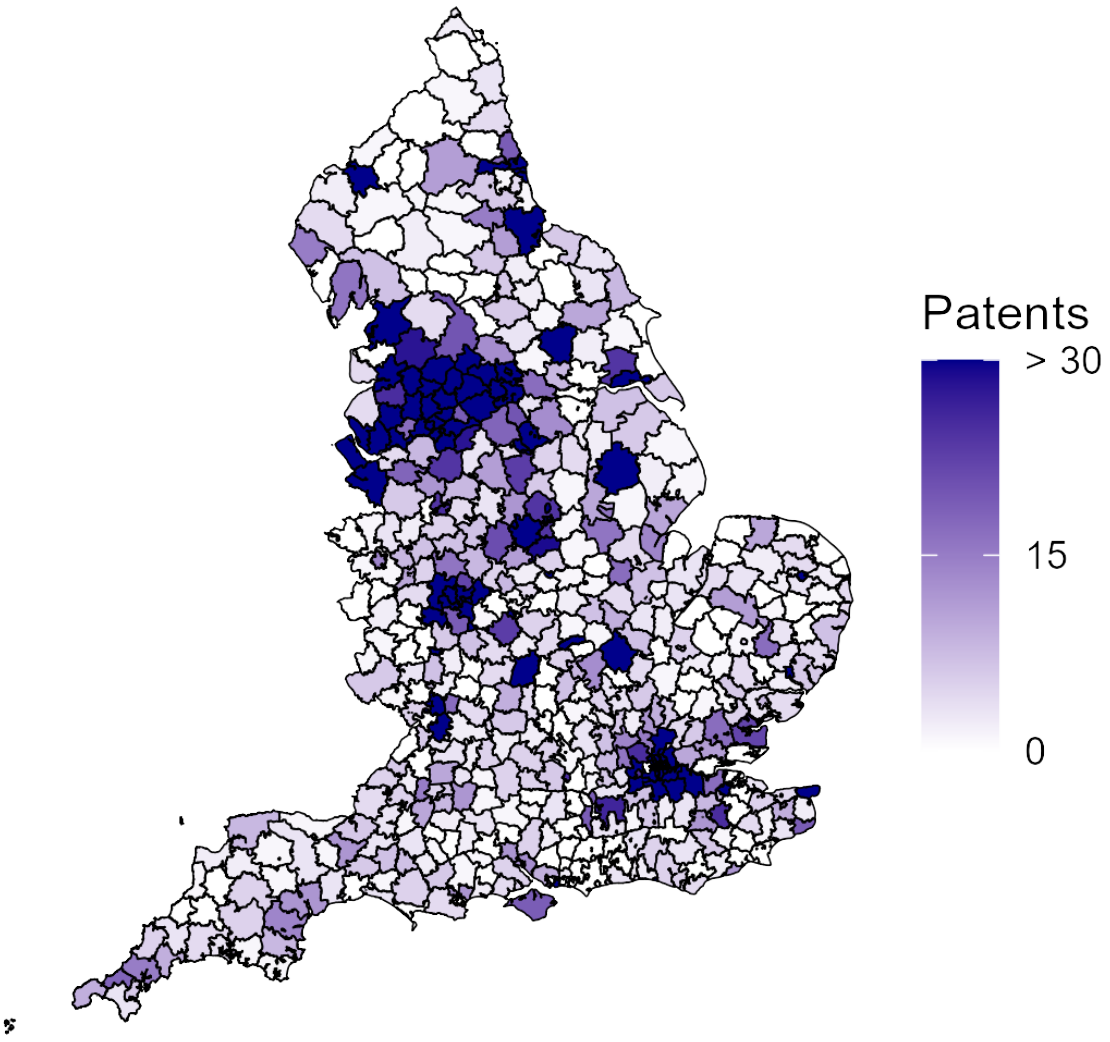


Figure 23: Heat map of all patents between 1849-1861 by district.

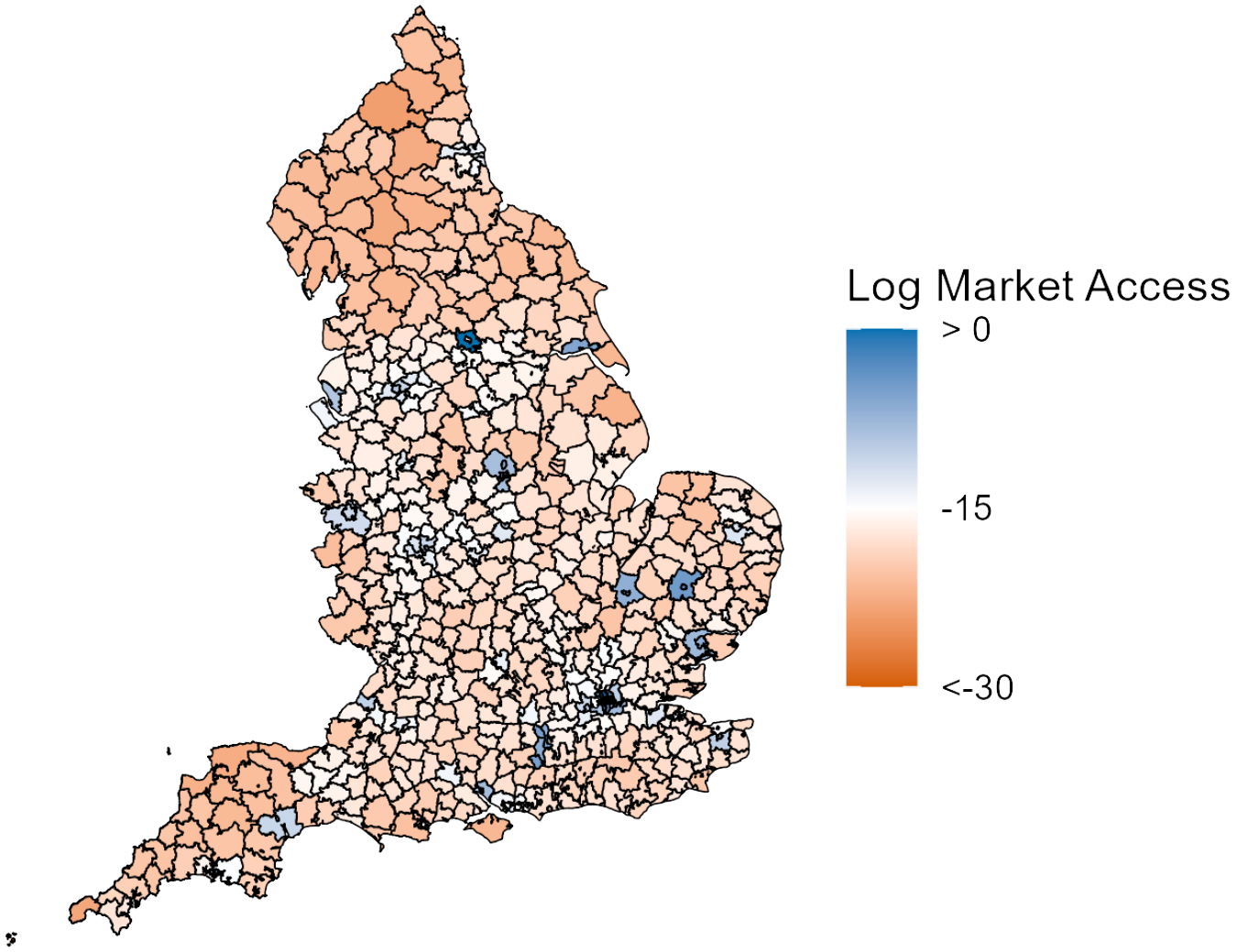


Figure 24: Log market access value for every district in England in 1823 by district.

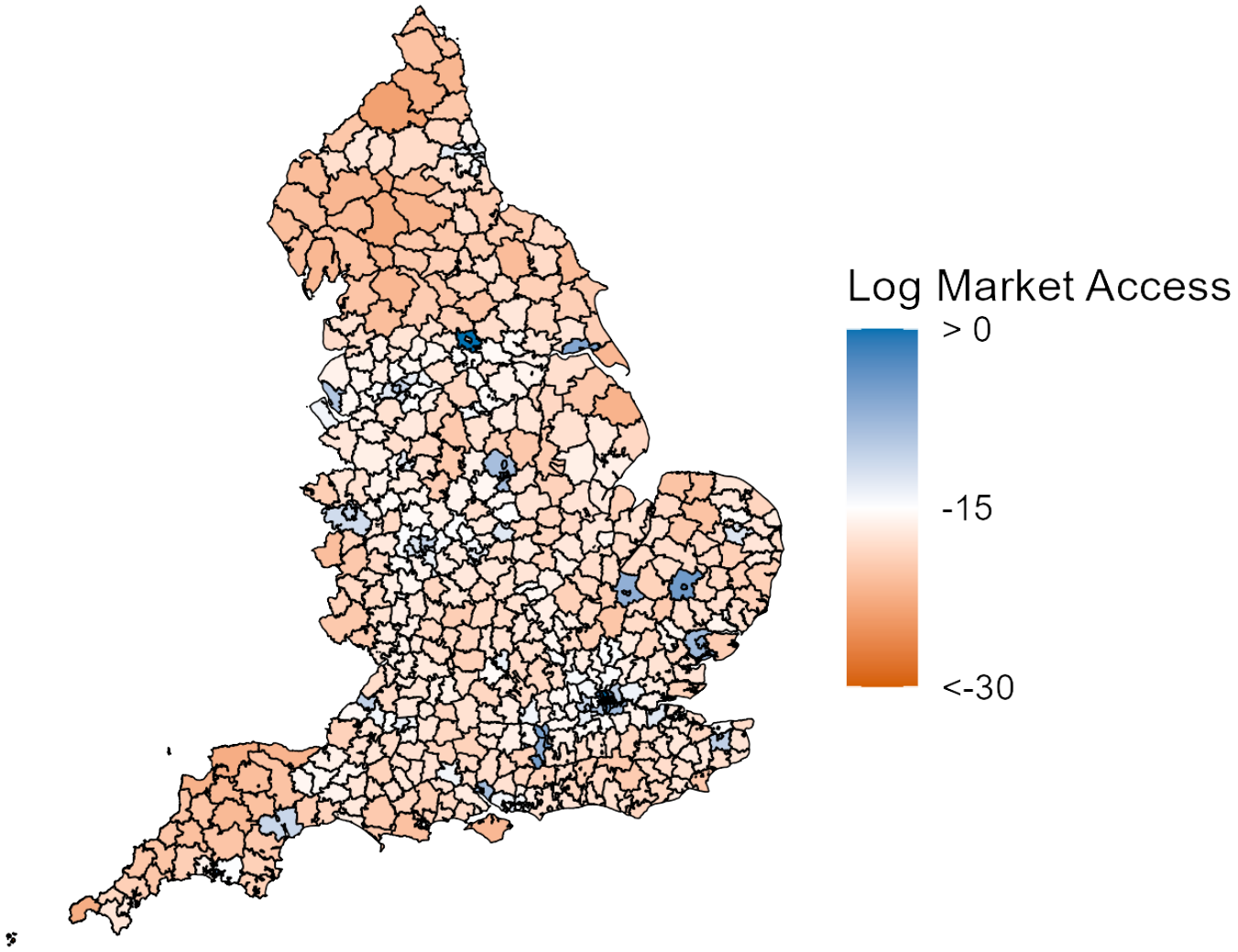


Figure 25: Log market access value for every district in England in 1841 by district.

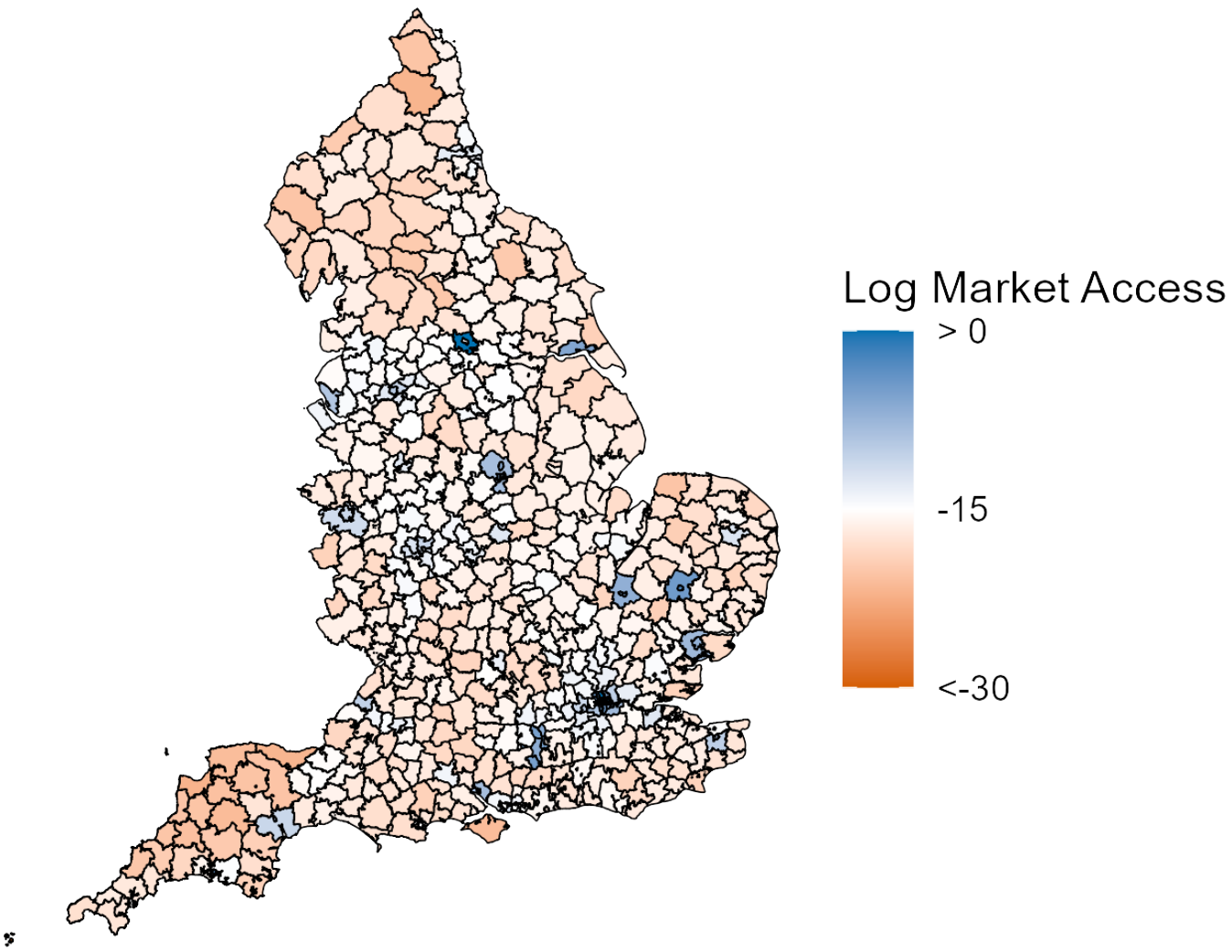


Figure 26: Log market access value for every district in England in 1861 by district.

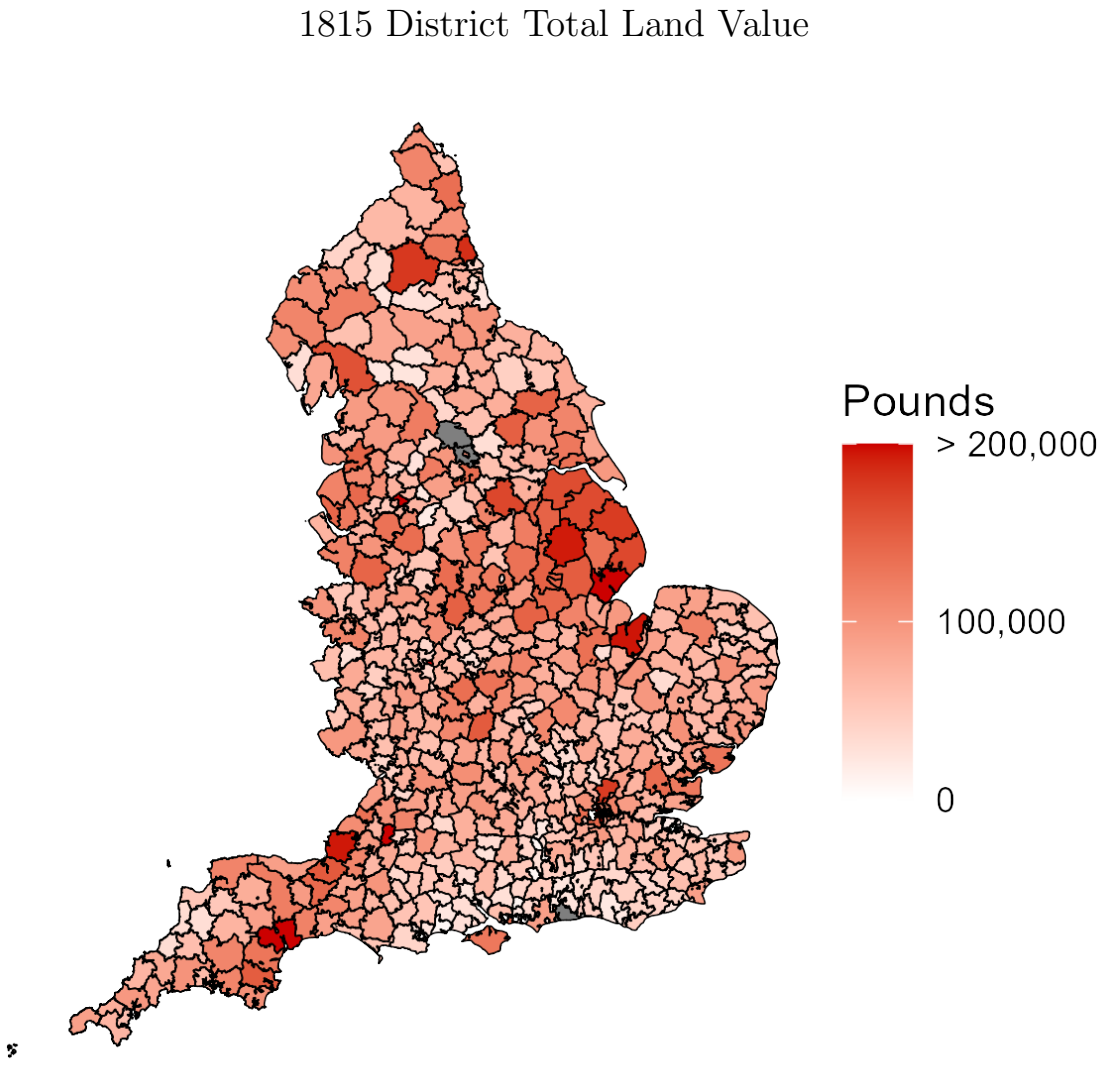


Figure 27: Districts by 1815 total land value. These values are used in the main definition of market access (Equation (2)) as a proxy for local GDP. Districts in gray (Hunslet, Kensington Town, Otley, Westminster, and Worthing) have no data. 16 districts have values above 200,000 pounds and three have values above 500,000 pounds.

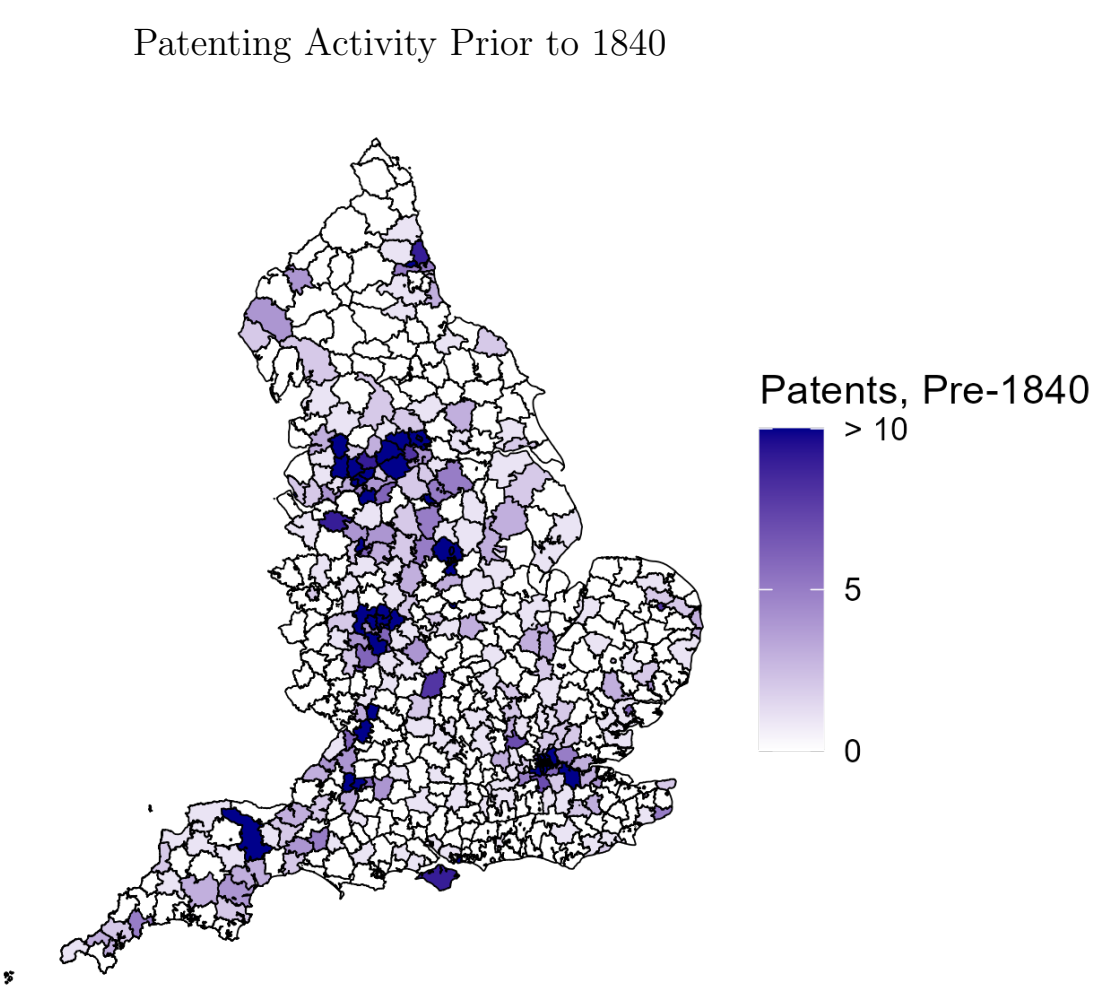


Figure 28: Maps of England with districts shaded by number of cumulative patents between 1823-1839.

Tables for Robustness Tests

Main Results

Table 14: Econometric Robustness Tests

	(1)	(2)	(3)	(4)	(5)
Rail Access	0.902***	0.202***	0.306**	0.814***	0.571***
	(0.253)	(0.053)	(0.137)	(0.200)	(0.124)
Observations	6864	6864	6864	10374	6864
Districts	176	176	176	266	169
Years	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861
Mean of Dep. Var.	1.085	0.336	0.128	0.874	0.247
Test	OLS	IHS	Indicator	SLS	County Cluster
Kleibergen-Paap Wald rk F				4802.059	

Notes: Robustness exercises for the results presented in Table 2. The main independent variable in all columns is an indicator equal to one if district i has a rail station in year t. Column (1) reports the results of an OLS estimation in which the dependent variable is patents per 100,000 people. Column (2) uses the IHS transformation of patents per capita as the dependent variable. Column (3) estimates a Poisson model and uses an indicator equal to one if there is at least one patent in district i in year t as the dependent variable. Column (4) reports the results of a SLS regression using the "Accidentally Connected" treatment definition which excludes rail end points and hubs. The simulated lines are used as the instrument to estimate the treatment status of each district. The treatment year is the actual year of treatment. Column (5) estimates a Poisson regression where the errors are clustered by county, rather than by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 15: Definition and Sample Robustness Tests

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Rail	0.571***	0.459**	0.491**	0.496***	0.611***	0.494***	0.019*
	(0.139)	(0.208)	(0.215)	(0.139)	(0.154)	(0.150)	(0.011)
Rail X Board of Trade					-0.133		
					(0.346)		
Rail X Mania						0.420	
						(0.350)	
Observations	6864	3451	6864	5226	6864	6864	7553
Clusters	176	119	176	134	176	176	286
Years	1823-1861	1823-1851	1823-1861	1823-1861	1823-1861	1823-1861	1823-1851
Mean of Dep. Var.	0.247	0.123	0.247	0.284	0.247	0.247	0.044
Test	Main	Pre 1852	Alt. Geo.	Alt. Never	BoT	Mania	Indiv.
Year FE	Y	Y	Y	Y	Y	Y	Y
District FE	Y	Y	Y	Y	Y	Y	
Individual FE							Y

Notes: This table shows estimation results with varying samples and definitions. In columns (1)-(5) equation (1) is estimated with a Poisson model where the dependent variable is patents per capita and the main independent variable is an indicator equal to one if a district has a station in year t. Column (1) reports the baseline Poisson results reported in Table 2. Column (2) uses a sample that spans 1823-1851, excluding years after the Patent Law Amendment Act. Column (3) utilizes an alternate geographic boundary to define treatment of a district. A 3KM radius circle is drawn around each station and treatment is assigned to districts that constitute the majority of the circle or districts with the highest percentage of their land within the circle. Column (4) excludes districts that are treated after 1861 from the control group. Column (5) includes an interaction term between the rail indicator variable and an indicator equal to one if a district is treated by a rail line approved by the 1845 Board of Trade. Column (6) includes an interaction term between the rail indicator variable and an indicator equal to one if the district was treated during 1839-1840 or 1846-1847, the years at the height of Rail Mania. Columns (7) reports OLS estimates of equation (10) which uses a panel of individual inventors, rather than districts. The dependent variable is an indicator equal to one if the inventor issues a patent application in year t, zero otherwise. Individual and year fixed effects are included as independent variables in both models as is a vector of population controls. Errors are clustered by individual in column (7) and clustered by district in all other columns. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 16: Sector Robustness Tests

	(1)	(2)	(3)	(4)
Rail Access	0.571***	0.574***	0.687***	0.567***
	(0.139)	(0.152)	(0.152)	(0.137)
Observations	6864	6630	6045	6864
Districts	176	170	155	176
Years	1823-1861	1823-1861	1823-1861	1823-1861
Sector	All	No Rail Adjacent	No Agriculture	Rail Included

Notes: Poisson estimation results of equation (1) where the dependent variable is patents per capita and the main independent variable is an indicator equal to one if a district has a station in year t. Column (1) reports the baseline results reported in Table 2. Column (2) removes patents in sectors that are adjacent to rail (engines and carriages) and column (3) removes agricultural patents. Column (4) retains patents from all sectors including rail, which are otherwise excluded from all other analyses. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 17: Effect of Rail Network Access on Population Growth

	(1)	(2)	(3)	(4)
	Population	Population	Log Population	Log Population
Rail Access	609.051*	609.051*	0.029**	0.029**
	(362.542)	(336.691)	(0.013)	(0.014)
Observations	6864	6864	6864	6864
Districts	176	176	176	176
Years	1823-1861	1823-1861	1823-1861	1823-1861
Mean of Dep. Var.	18071.541	18071.541	9.699	9.699
Cluster	District	County	District	County

Notes: OLS estimation based on equation (1) where the dependent variable is population in columns (1) and (2) and log population in columns (3) and (4). The main independent variable is an indicator equal to one if district i has a rail station in year t. Errors are clustered by district in columns (1) and (3) and by county in columns (2) and (4). * p < 0.1, ** p < 0.05, *** p < 0.01

Table 18: Relationship between Railroad Stations and World's Fair Exhibitions

	(1)	(2)	(3)	(4)
Rail Station	0.738*	0.781*	0.680**	0.735**
	(0.401)	(0.401)	(0.302)	(0.302)
Observations	214	214	214	214
Mean of Dep. Var.	1.350	1.350	0.318	0.318
Pseudo R-Squared	0.033	0.037	0.019	0.023
Station Year	1851	1850	1851	1850
Model	Poisson	Poisson	Logit	Logit

Notes: Poisson and logit model estimates where the dependent variable is the number of exhibitions per capita for a district in the 1851 World's Fair in columns (1) and (2) and a dummy variable equal to one if a district has at least one award in column (3) and (4). In columns (1) and (3) a district is considered to have a rail station if it has one in 1851 and in columns (2) and (4) if it has one in 1850. One observation is equal to one district; the number of observations differs from other analyses as I include districts from the Main sample that had zero patents in all years. Robust errors are reported. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 19: Effect of Postage Cost Reduction

	(1)	(2)	(3)	(4)
Rail	0.490***	0.344	-0.558	-5.248*
	(0.123)	(0.689)	(0.878)	(2.910)
Rail X Distance to London		0.031		0.889
		(0.147)		(0.551)
Rail X Post 1839			1.059	5.884**
			(0.892)	(2.881)
Rail X Distance to London X Post 1839				-0.918*
				(0.548)
Observations	8658	8658	8658	8658
Districts	222	222	222	222
Years	1823-1861	1823-1861	1823-1861	1823-1861

Poisson estimation of equation (1) where the dependent variable is the number of patents per capita where population is controlled for as an exposure variable. The main independent variables are "Rail," which is an indicator equal to one for years in which district i has a rail station, "Distance to London," which is the distance between district i and London, and "Post 1839," which is an indicator equal to one if the year is after 1839. The treatment definition used in all columns is the "No Ends or Rail Hubs" definition due to lack of variance in the main definition to handle the interaction effects. Errors are clustered by district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Market Access

Table 20: Effect of Market Access on Innovation - Navigable Waterways and Market Size

	(1)	(2)	(3)	(4)	(5)
Log Market Access	0.145**	0.181**	0.251***	0.330***	0.511***
	(0.066)	(0.074)	(0.073)	(0.081)	(0.130)
Observations	18642	16419	8892	8190	6864
Districts	478	421	228	210	176
Years	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861
Definition	All	Top Income	Waterways	Both	Main

Notes: Poisson estimation of (3) where the dependent variable is log patents per capita and the main independent variable is log market access. Column (1) includes all districts; column (2) excludes districts in the top decile by 1815 income level; column (3) excludes districts with immediate access to the main navigable waterway (defined as the waterway passing within a district's geographic boundaries); column (4) excludes all districts excluded in columns (2) and (3); column (5) uses the main treatment definition. Errors are clustered by district in all columns. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 21: Effect of Market Access on Innovation - Various Robustness Exercises

	(1)	(2)	(3)	(4)	(5)
Log Market Access	0.503***	0.443**	0.394***	0.466***	0.815***
	(0.144)	(0.187)	(0.145)	(0.128)	(0.256)
Observations	6864	6864	6864	6864	528
Districts	176	176	176	176	176
Years	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861
GDP def.	Land Value	Land Value	Land Value	Pop.	Land Value
Theta	8.22	3.8	12.86	8.22	8.22

Notes: Various robustness exercises on the market access mechanism. Columns (1) - (3) report Poisson estimates of equation (3) where the dependent variable is log patents per capita and the main independent variable is log market access. Column (1) reports the baseline estimate with a value of θ equal to 8.22 and columns (2) and (3) use values of θ on either end of the scale proposed by Donaldson and Hornbeck (2016). Column (4) uses districts' 1821 population as the proxy for GDP in the definition for market access. Column (5) estimates a stacked OLS regression model where each period is ten years. The dependent variable is the change in patents per capita between t and t+1 and the main independent variable is the log change in market access between t-1 and t. Errors are clustered by district in all columns. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 22: Effect of Market Access on Innovation - Geographic Variation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log Market Access	0.600***	0.441***	0.521***	0.459***	0.501***	0.618***	0.544***	0.480***	0.524***	0.419***
	(0.149)	(0.118)	(0.136)	(0.128)	(0.131)	(0.151)	(0.137)	(0.140)	(0.140)	(0.129)
Observations	5814	5928	6396	6786	6786	5694	5967	5460	6279	6513
Districts	153	152	164	174	174	146	153	140	161	167
Years	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861	1823-1861

Notes: Poisson estimation results based on equation (3) where the dependent variable is the log of patents in district i and year t and the main independent variables is log market access. Each column removes all districts contained in one of the ten Census Divisions. Results can be compared with the baseline result in Table 4, which is a coefficient of .511. Errors are clustered by district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Knowledge Flows

Table 23: Effect of Knowledge Flows on Patenting - High Patent Count Sectors

	(1)	(2)	(3)	(4)
Log Knowledge Flows	0.336***	0.240***	0.146***	0.333***
	(0.073)	(0.090)	(0.031)	(0.074)
Rail				0.181
				(0.193)
Observations	96278	96278	272090	96278
Districts	161	161	455	161
Years	1839-1861	1839-1861	1839-1861	1839-1861
Definition	Main	Main	All	Main
Sector FE		Y		

Notes: This table replicates Table 8, but restricts sectors to those in the top half by patent count in the 1823-1839 period. Each column shows the results of the Poisson estimation of equation (7) where the dependent variable is the log of sector s patents in district i and year t and the main independent variable is log $KF_{i,s,t}$. Column (2) includes sector fixed effects. Columns (3) includes all districts while all other columns include only districts which comprise the main treatment definition. Column (4) includes the rail treatment variable $D_{i,t}$ as a control. As $KF_{i,s,t}$ is calculated using patents from 1823-1839, I restrict the data to 1840-1861. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 24: Effect of Knowledge Flows on Patenting - Aggregated Sectors

	(1)	(2)	(3)	(4)
Log Knowledge Flows	0.820***	0.576***	0.904***	0.821***
	(0.097)	(0.220)	(0.028)	(0.098)
Rail				-0.020
				(0.227)
Observations	49933	49933	139035	49933
Districts	167	167	465	167
Years	1839-1861	1839-1861	1839-1861	1839-1861
Definition	Main	Main	All	Main
Sector FE		Y		

Notes: Results of Poisson estimation of equation (7) where the dependent variable is the log of sector $\bar{s}$ patents in district i and year t and the main independent variable is log $KF_{i,\bar{s},t}$. Column (2) includes sector fixed effects, column (3) includes districts from all of England, and column (4) includes the rail treatment variable $D_{i,t}$ as a control. As $KF_{i,\bar{s},t}$ is calculated using patents from 1823-1839, I restrict the data to 1840-1861 in all columns. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 25: Effect of Knowledge Flows on Patenting - Geographic Variation

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log Knowledge Flows	0.200***	0.213***	0.226***	0.192***	0.169***	0.200***	0.199***	0.200***	0.206***	0.191***
	(0.030)	(0.037)	(0.037)	(0.031)	(0.032)	(0.031)	(0.031)	(0.030)	(0.032)	(0.030)
Observations	493218	514206	454740	500214	504878	513040	465234	487388	469898	502546
Districts	423	441	390	429	433	440	399	418	403	431
Years	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861

Notes: Results of Poisson estimation of equation (7) where the dependent variable is the log of sector s patents in district i and year t and the main independent variable is log $KF_{i,s,t}$. Each column removes all districts in one of the ten Census Divisions of England. As $KF_{i,s,t}$ is calculated using patents from 1823-1839, I restrict the data to 1840-1861. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Table 26: Effect of Knowledge Flows on Patenting - Updating Years

	(1)	(2)	(3)	(4)
Log Knowledge Flows	0.534***	0.237***	0.265***	0.535***
	(0.067)	(0.069)	(0.039)	(0.067)
Rail				-0.082
				(0.236)
Observations	191565	175890	537532	191565
Districts	165	165	463	165
Years	1839-1861	1839-1861	1839-1861	1839-1861
Definition	Main	Main	All	Main
Sector FE		Y		

Notes: Results of the Poisson estimation of equation (7) where the dependent variable is the log of sector s patents in district i and year t and the main independent variable is log $KF_{i,s,t}$. Columns (2) includes sector fixed effects. Columns (3) includes all districts while all other columns include only districts which comprise the main treatment definition. Column (4) includes the rail treatment variable $D_{i,t}$ as a control. Unlike the main definition of $KF_{i,s,t}$, $KF_{i,s,t}$ is calculated here using a rolling stock of sector s patents from the five years leading up to year t. I restrict the data to 1840-1861 to allow comparisons with the main results in Table 8. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Comparison of Market Access and Knowledge Flows

Table 27: Effect of Market Access and Knowledge Flows on Patenting

	(1)	(2)	(3)	(4)	(5)	(6)
Log Market Access	0.481***		0.885***	0.482***		-0.164
	(0.119)		(0.177)	(0.118)		(0.162)
Log Knowledge Flows		0.124	-0.274***		0.427***	0.433***
		(0.082)	(0.090)		(0.062)	(0.064)
Observations	3841	3841	3841	203573	203573	203573
Districts	167	167	167	167	167	167
Years	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861	1839-1861
Definition	Main	Main	Main	Main	Main	Main
Pseudo R2	0.457	0.454	0.459	0.195	0.251	0.252

Notes: Poisson coefficient estimates based on equation (3) where the dependent variable is the log of patents per capita in columns (1) - (3) and column (7) and the log of sector s patents per capita in columns (4) - (6). The main independent variables are log $MA_{i,t}$ and log $KF_{i,t}$ in columns (1) - (3) and log $MA_{i,t}$ and log $KF_{i,s,t}$ in columns (4) - (6). Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

Synthetic Difference in Difference Graphs

I provide here graphs in the three year periods for which I run the Synthetic Difference in Difference (SDID) estimation. The vertical line shows the treatment date. The dotted line is the control and the solid line is the treated. SDID estimation weights the control group to force parallel trends in the pre-treatment period. These graphs shows that this process is largely successful using the three year periods discussed in Section 4.

Figure 29: SDID estimation trend graph where treatment year is between 1832-1834.

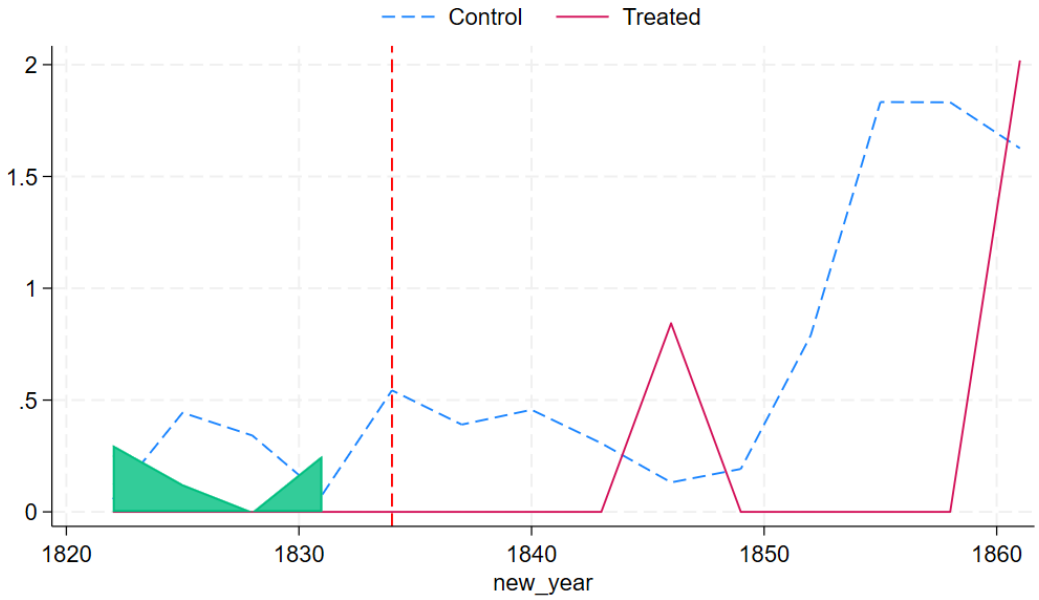


Figure 30: SDID estimation trend graph where treatment year is between 1835-1837.

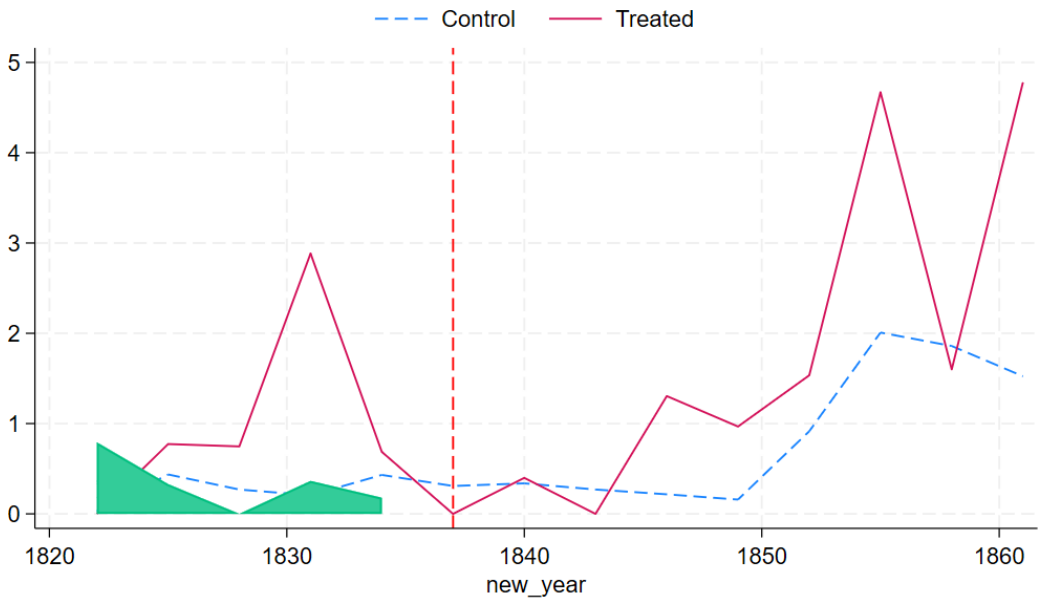


Figure 31: SDID estimation trend graph where treatment year is between 1838-1840.

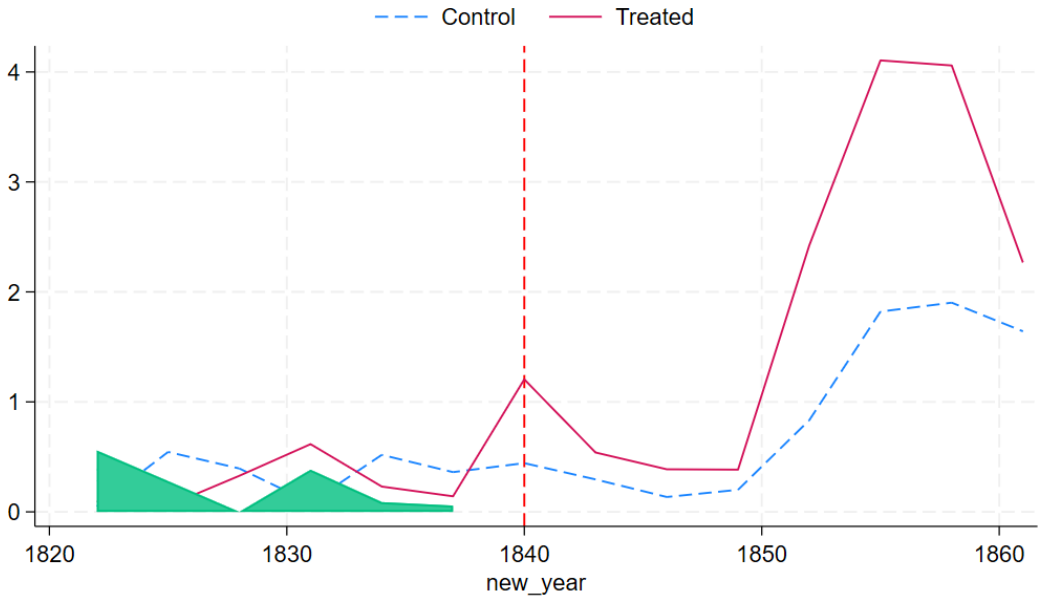


Figure 32: SDID estimation trend graph where treatment year is between 1841-1843.

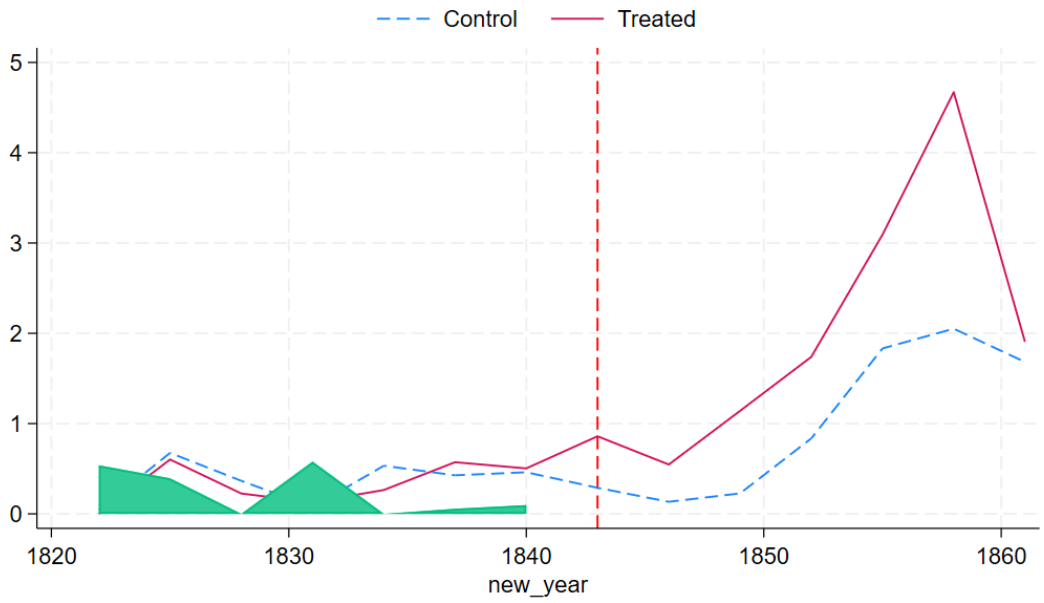


Figure 33: SDID estimation trend graph where treatment year is between 1844-1846.

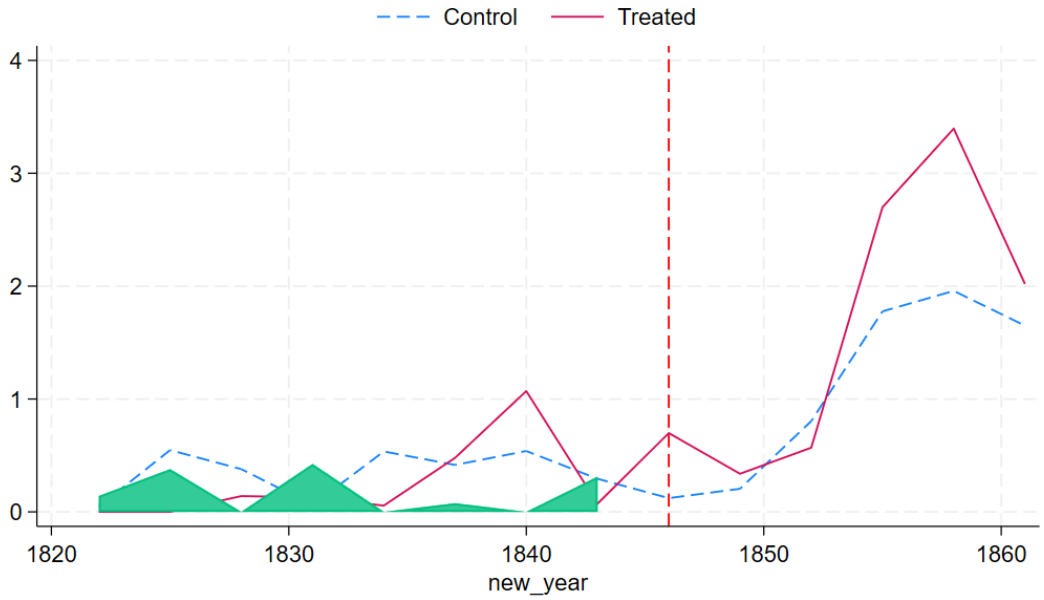


Figure 34: SDID estimation trend graph where treatment year is between 1847-1849.

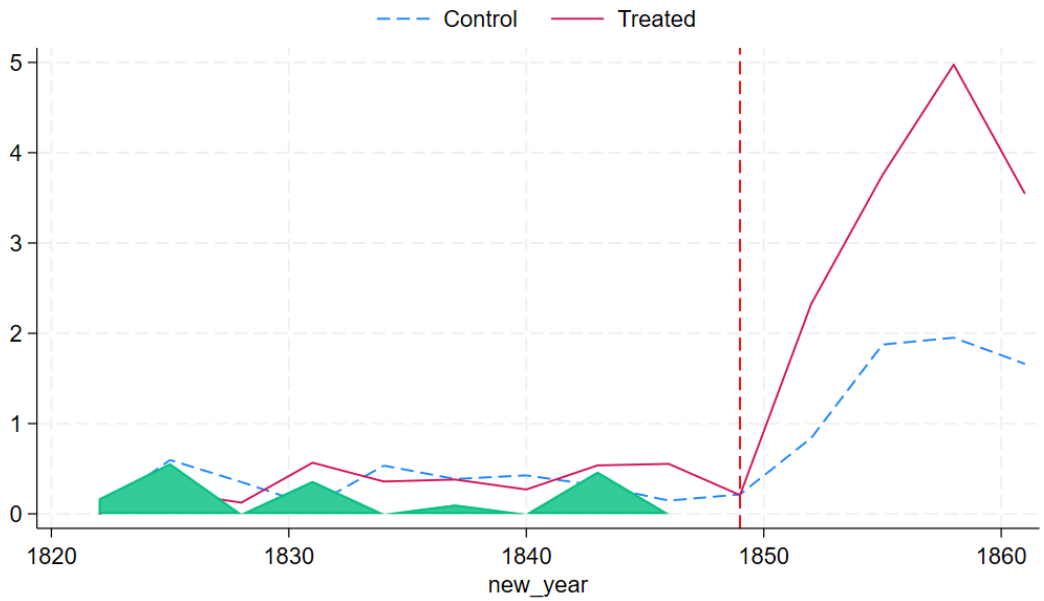


Figure 35: SDID estimation trend graph where treatment year is between 1850-1852.

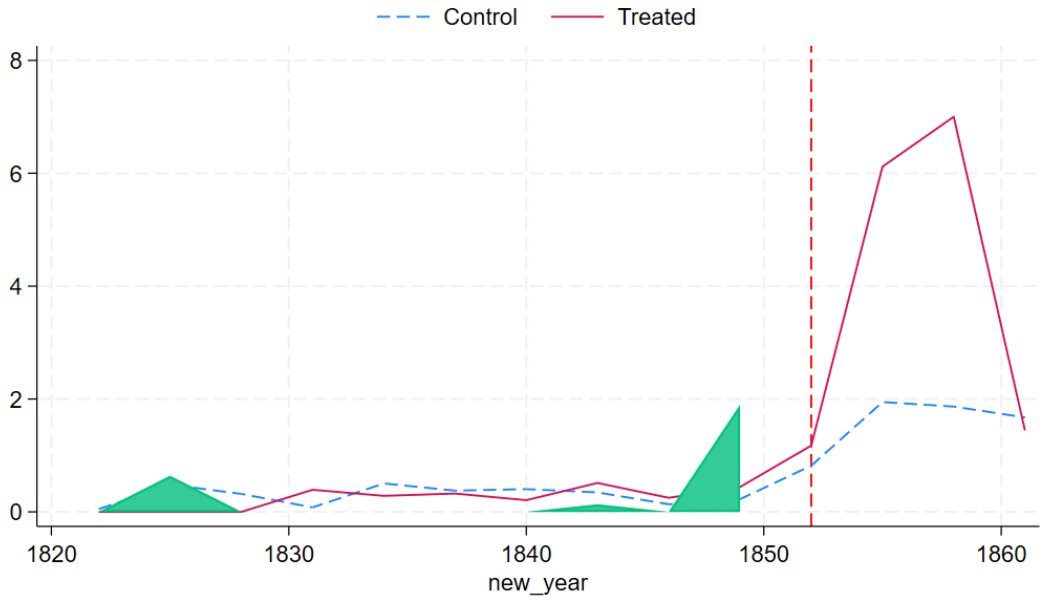


Figure 36: SDID estimation trend graph where treatment year is between 1853-1855.

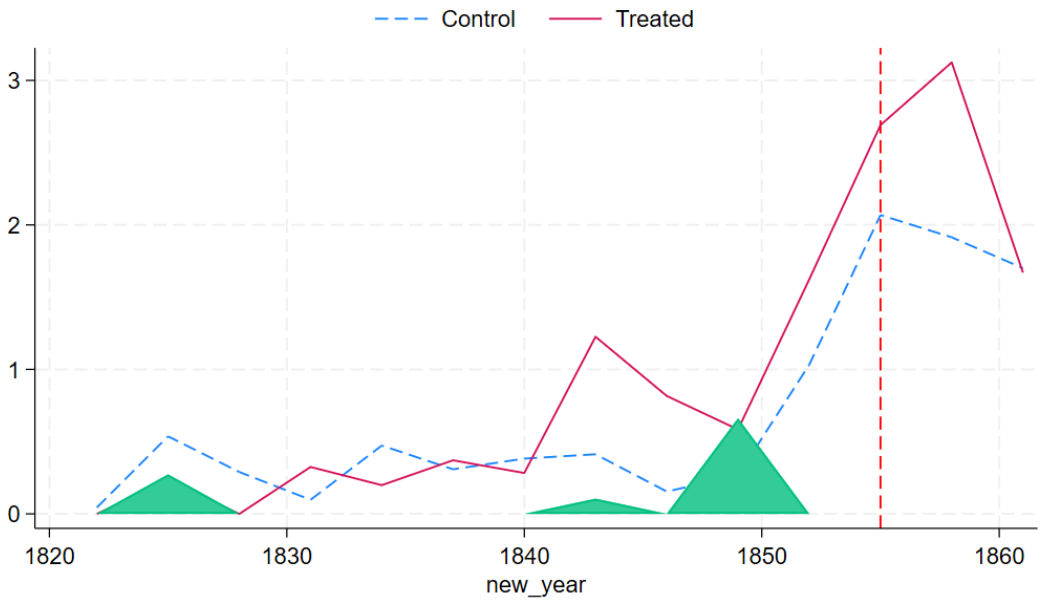


Figure 37: SDID estimation trend graph where treatment year is between 1855-1858.

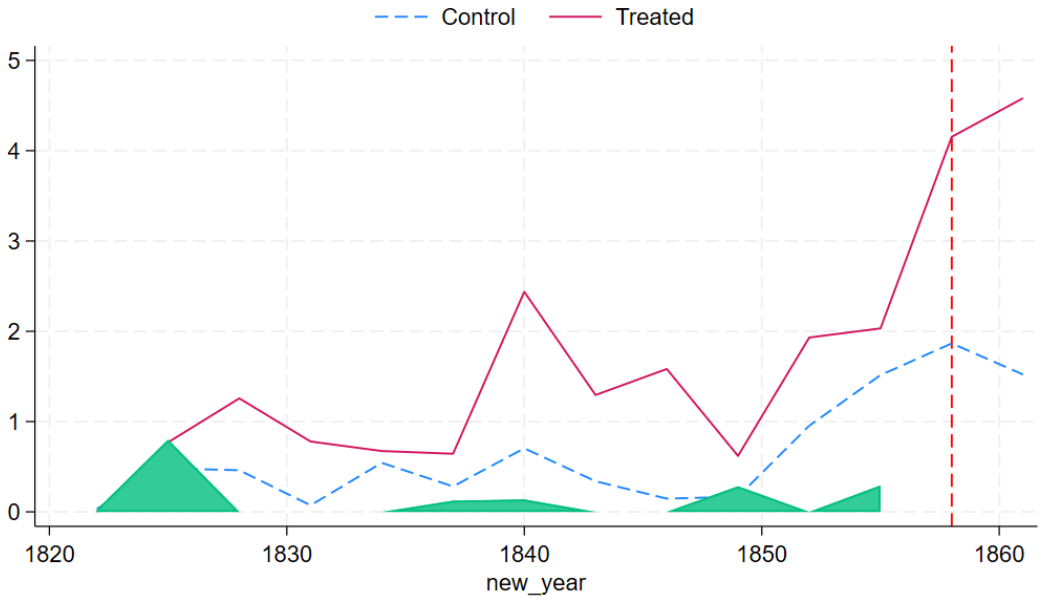
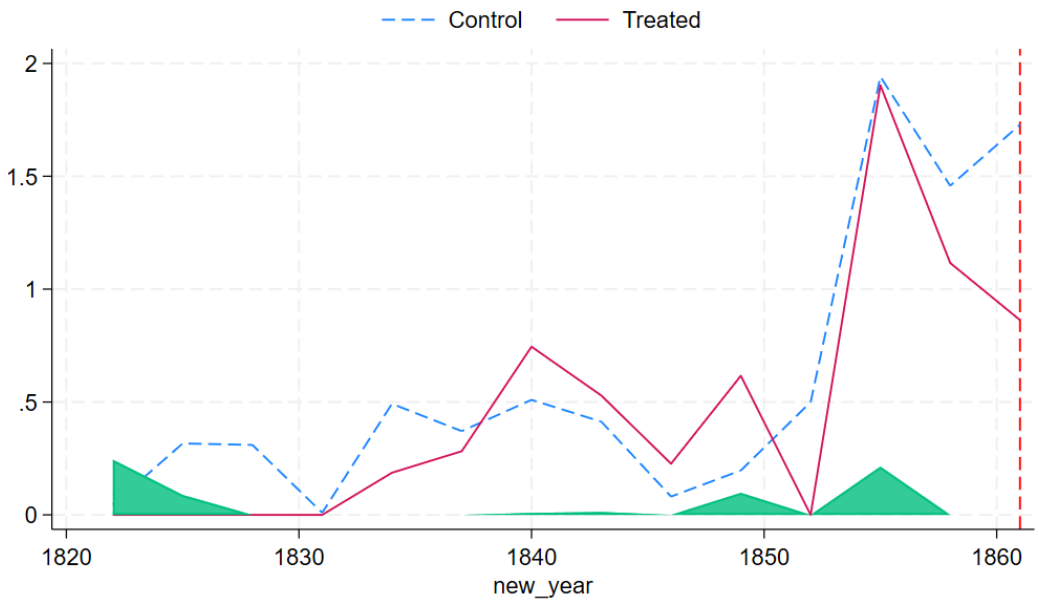


Figure 38: SDID estimation trend graph where treatment year is between 1858-1861.



Model

The nature of railroads makes it difficult to evaluate national effects on output or productivity. The empirical results provide evidence of a positive effect in districts that received a rail station. However, this could come at the expense of other districts. If higher productivity workers in a district that did not receive a rail station move to a district that did, the positive productivity effects would cancel out at the national level. Section 5.4 provides evidence that this is not the case in this setting. I follow Kline and Moretti (2014) and Agrawal et al. (2017) in creating a model that provides additional evidence that disentangles a population agglomeration effect and a patenting effect on labor productivity and I show evidence that the agglomeration channel does not appear to cancel out at the national level.

I model districts as small, open economies. Firms take the prices of capital K, labor L, and output Y as given. Worker utility is modeled as a function of wages w and amenities M:

$$ln(u_{it}) = ln(w_{it}) + ln(M_{it}) \tag{12}$$

Output Y is produced according to a Cobb-Douglas production function with a fixed factor F and productivity A:

$$Y_{it} = A_{it}K_{it}^{\alpha}F_{t}^{\beta}L_{it}^{1-\alpha-\beta} \tag{13}$$

I normalize the price of Y to 1 and denote w as the wage and r as the cost of capital to obtain an inverse labor demand curve:

$$ln(w_{it}) = ln(1-\alpha-\beta) + \frac{\alpha}{1-\alpha}ln(\alpha) - \frac{\beta}{1-\alpha}ln(L_{it}) + \frac{\alpha}{1-\alpha}ln(F_i) - \frac{\alpha}{1-\alpha}ln(r_t) + \frac{1}{1-\alpha}ln(A_{it}) \tag{14}$$

I model productivity A as a function of labor, L, and patents, P:

$$ln(A_{it}) = \sigma ln(L_{it}) + \xi ln(P_{it}) \tag{15}$$

σ and ξ can respectively be interpreted as the strength of agglomeration forces and patenting on productivity. I find evidence in Section 6 that growth in both L and P is dependent on having a rail station and I use a reduced form specification to model this relationship:

$$L = aR^{\mu} \tag{16}$$

$$P = bR^{\theta} \tag{17}$$

Combining equations (14) - (17) and taking the derivative with respect to a rail station R yields the following effect of rail on local wages:

$$\frac{dln(w)}{dln(R)} = -\frac{\beta\mu}{1-\alpha} + \frac{\xi\theta}{1-\alpha} + \frac{\mu\sigma}{1-\alpha} \tag{18}$$

Thus far, this model replicates Agrawal et al. (2017) and I follow the authors' interpretation and calibration. The negative term is the competitive effect that a rail station has on wages; rail attracts labor and a larger labor force lowers wages. The positive term is the productivity effect through patenting and

agglomeration forces; higher labor productivity increases wages. I set $\xi = .11$, $\beta = .47$ and $\alpha = .68$. The regression estimates in Section 4 and Section 6 yield patenting and labor elasticities and I set $\theta = .77$ and $\mu = .03$. Substituting these values into equation (18) yields:

$$\frac{dln(w)}{dln(R)} = -.04 + .27 + .09\sigma \tag{19}$$

Given the large discrepancies in the labor and patenting effects I find compared with Agrawal et al. (2017), the predicted effect is expected to be quite different. Indeed, with a value of $\sigma = 0$, Agrawal et al. (2017) find an elasticity of wages of -.14. However, the patenting effect in this setting is so great that even without agglomeration forces the effect on labor productivity would be positive and equal to .23. Setting $\sigma = .72$, as Agrawal et al. (2017) do, yields a total effect of .25, larger than the authors' findings of .20. This is in line with the larger patenting effect I find in this setting. The total productivity effect is .42, exactly the same as Agrawal et al. (2017) but for different reasons. The authors find that the patenting channel makes up 19% of the total productivity effect. I find that the patenting channel accounts for 64% of the productivity effect. Mechanically, this large discrepancy is due to the difference in estimated values.

The model thus far emulates the Kline and Moretti (2014) style model that Agrawal et al. (2017) use in order to highlight the differences between Agrawal et al. (2017) and this paper. I now continue to follow Kline and Moretti (2014) to model national level effects of the rail network.

I now model A as a function of population density and keep the rail station variable invariant across time.

$$ln(A_{it}) = \sigma_i ln(V_{it-1}) + \delta_t D_i + \eta_i + \epsilon_{it} \tag{20}$$

V_{it-1} is the population density of district i. It includes a time lag of one period because, as noted by Kline and Moretti (2014) and Adsera and Ray (1998), this ensures that the model yields deterministic predictions which exclude implausible scenarios. D_i is a dummy for whether a district has a rail station and δ_t is the direct effect of that access on local productivity in year t. I include district fixed effects η_i to capture time invariant effects in each district. The error term ν_{it} represents the idiosyncratic component of district productivity.

Rewriting equation (14) in terms of labor and taking first differences yields:

$$ln(L_{it}) - ln(L_{it-1}) = -\frac{1-\alpha}{\beta}(ln(w_{it}) - ln(w_{it-1})) + \frac{\delta_t - \delta_{t-1}}{\beta}R_i + \frac{\sigma_i}{\beta}(ln(V_{it-1}) - ln(V_{it-2})) + \frac{1}{\beta}\nu_{it} \tag{21}$$

The coefficient $\frac{\delta_t - \delta_{t-1}}{\beta}$ gives the change in the effects of having a rail station and the coefficient $\frac{\sigma_i}{\beta}$ gives the indirect effects of rail through a labor agglomeration channel.

One issue in estimating equation (21) is isolating the agglomeration effect from the local rail station effect. As productivity increases from receiving a station, labor demand will increase causing individuals to move to the effected district and increasing population density. An agglomeration effect would then cause another increase in productivity. This dual effect necessitates a method to isolate the agglomeration effect. I follow Kline and Moretti (2014) by introducing further lags in the population density variable P_{it}. Given a long enough time frame, the productivity shocks ν_{it} should be independent. I define a time period as five years and use additional lags which measure the change in population density between fifteen and twenty years before 1861 as the instrument:

Table 28: Model Structural Estimates

	(1) OLS	(2) 2SLS
Bottom Half	0.916***	0.906***
	(0.257)	(0.259)
Top Half	1.007***	1.126***
	(0.132)	(0.136)
Rail Access	0.006	0.004
	(0.008)	(0.008)
Year Effects	Yes	Yes
Observations	93	93
F Statistic	39.472	43.406
Kleibergen-Paap F		14.039
Equal Slopes	0.751	0.436
All Equal 0	0.000	0.000

Results of regressions based on estimation of equation (21). Column 1 shows the OLS estimates and column 2 shows the 2SLS estimates using the lagged population density variable as the instrument. P-values for the null hypotheses of equal slopes of the bottom and top half density splines and that both spline coefficients are equal to zero are reported at the bottom of the table. Regressions are weighted by 1821 population. Errors are clustered by district. * p < 0.1, ** p < 0.05, *** p < 0.01

$$Z_{it} = ln(V_{it-3}) - ln(V_{it-4}) \tag{22}$$

For the coefficient to contain an agglomeration effect, the district must already be treated at the time of the lag so that a rail station has an effect on the lagged population density. I therefore restrict the data to districts that were either treated prior by the end of 1841 or never treated as of 1861. I use the 1851 and 1861 population data provided by Schurer and Higgs (2023) to construct annual population data between those years and the original census data from 1841 to 1851 to construct the 1841-1851 annual population data. Given that the data are imputed, sequential years from the same decade would exhibit serial correlation. I limit the regression to a single time period (where t = 0 is 1861) to avoid this issue at the cost of additional observations.

Wage data are constructed from the 1843 and 1860 district level land value data and annualized. Following Kline and Moretti (2014), the coefficient on the wage change variable, $\frac{1-\alpha}{\beta}$, is assigned a value of 1.5.

Kline and Moretti (2014) create splines of the instrument to assess whether the agglomeration coefficient is constant across densities. The authors shows that with constant agglomeration elasticities, national level agglomeration effects cancel out. I similarly spline the instrument in estimating equation (21). The results in Table 28 indicate that the coefficient for the instrument representing the top half of districts by 1841 district level population density is 20% higher than the coefficient for the bottom half. While this seems to indicate that distribution of population density elasticities is not constant, a null hypothesis of equal slopes cannot be rejected.

The rail station coefficient is essentially zero. This is not surprising as equation 21 is a first difference effect and only "always treated" and "never treated" districts are included. The coefficient is therefore estimating whether receiving a rail station at least fifteen years ago had a direct effect on labor demand between two consecutive periods.

Table 29: Population Growth by Population Density

	(1) Bottom Half	(2) Bottom Half	(3) Top Half	(4) Top Half
Rail Access	0.037***		0.044***	
	(0.008)		(0.011)	
Pre-Treatment		0.028***		0.021***
		(0.007)		(0.008)
Observations	3915	3391	3915	3251
Districts	135	135	135	135
R-Squared	0.685	0.666	0.697	0.708

OLS estimation of equation (11) where the dependent variable is the log of the imputed population. I use the same Analytic Sample used in calculating the results in Table 1, but divide the districts into halves based on 1831 population density. Columns (1) and (2) include the districts in the bottom half of the data sample and columns (3) and (4) include districts in the top half. In columns (1) and (3) the main independent variable is an indicator equal to one in all periods where the district has a rail station. In columns (2) and (4) it is an indicator equal to one starting exactly ten years prior to having a rail station until 1851, so that it includes pre-treatment population growth. Comparing columns (2) and (4) indicates that pre-treatment growth was higher in the less dense districts and comparing across (1) and (3) indicates that post-treatment growth was higher in the denser districts. Errors are clustered by district. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Taking the results as evidence that lower density locations had lower agglomeration elasticities, I now examine the potential ramifications at the national level. Per Kline and Moretti (2014), endogenous reallocation of workers from a lower elasticity district to a higher elasticity district raises aggregate output. I return to the Analytic Sample used in calculating the results in Table 2 to maximize the observation count, divide the districts into halves based on 1831 population density and estimate equation (1), where the dependent variable is the imputed population. Districts in the top half by population density grow faster after treatment by approximately .7 percentage points than districts in the bottom half, indicating that individuals from low density districts are relocating to high density districts after the treatment. Crucially, this does not appear to be the case prior to treatment - the low density districts are actually growing faster in the years leading up to the local rail station event, indicating causality of the treatment. Taken together with the results presented in Table 28, this supports the hypothesis that agglomeration forces from railroad access have a positive effect on national GDP because railroads moved individuals from low density locations (which have lower agglomeration elasticities) to high density districts (which have higher elasticities). Though the results of the model indicate that agglomeration forces are small in this setting, this is still a notable result as it provides evidence that transportation infrastructure can cause a national increase in productivity solely through an agglomeration effect.