

Railways, population divergence, and structural change in 19th century England and Wales¹

Dan Bogart², Xuesheng You³, Eduard Alvarez⁴, Max Satchell⁵, and Leigh Shaw-Taylor⁶

September 25, 2020

Abstract

Railways transformed inland transportation during the nineteenth century. In this paper, we study how railways led to population change and divergence in an already urbanized economy, England and Wales. We make use of detailed data on railway lines, stations, and population change in more than 9000 spatial units. We also create a least cost path based on major 1801 towns and the length of the 1851 rail network to address endogeneity. Our instrumental variable estimates show that having railway station in a locality by 1851 led to significantly higher population growth from 1851 to 1891 and shifted the male occupational structure away from agriculture. Moreover, we estimate that having stations increased population growth more if localities had greater population density in 1801. Also, there were population losses for localities 5 to 15 km from stations, indicating a displacement effect. Overall, we find that railways reinforced the urban hierarchy of the early nineteenth century and contributed to further spatial divergence of the English and Welsh population. The resulting implications for national income and labor productivity are found to be large.

Keywords: Urbanization, railways, transport, reorganization, endowments

JEL Codes: N7, O1, R4

¹ Data for this paper was created thanks to grants from the Leverhulme Trust (RPG-2013-093), Transport and Urbanization c.1670-1911, from the NSF (SES-1260699), Modelling the Transport Revolution and the Industrial Revolution in England, from the ESRC (RES 000-23-0131), Male Occupational Change and Economic Growth in England 1750 to 1851, and the ESRC (RES-000-23-1579) the Occupational Structure of Nineteenth Century Britain. We thank Walker Hanlon, Gary Richardson, Petra Moser, Kara Dimitruk, Arthi Vellore, William Collins, Jeremy Attack, Stephan Heblich, James Fenske, and Elisabet Viladecans Marsal for comments on earlier drafts and seminar participants at Warwick, Bristol, UC Merced, UC Irvine, UC San Diego, NYU, Florida State, Trinity College Dublin, Queens Belfast, Los Andes, Vanderbilt, and EHA Meetings. We also thank Cranfield University for share soils data.

² Professor, Department of Economics, UC Irvine, dbogart@uci.edu

³ Research Associate, Faculty of History, University of Cambridge, xy242@cam.ac.uk

⁴ Senior Lecturer, Economics and Business, Universitat Oberta de Catalunya, ealvarezp@uoc.edu

⁵ Research Associate, Dept. of Geography, University of Cambridge, aems2@cam.ac.uk

⁶ Senior Lecturer, Faculty of History, University of Cambridge, lmws2@cam.ac.uk

1. Introduction

Britain's urbanization was exceptional during the nineteenth century. Between 1800 and 1900 its population percentage living in cities of 5000 or more increased from 19 to 67. In the whole of Europe, the urbanization rate increased far less from 11 to 30 between 1800 and 1900 and even in the United States urbanization rates increased less from 5 to 36 (Bairoch and Goertz 1986). Britain's urbanization process was remarkable in another respect. Between 1850 and 1900 its urban areas grew dramatically, but its rural areas had little growth. This was not true elsewhere in the world. Population growth was much more balanced between urban and rural in much of Europe for example.⁷

In this paper we study how the introduction of railways in the second quarter of the nineteenth century led to population change and spatial divergence in England and Wales up to 1891.⁸ Railways were a major innovation for inland transport and had the potential to create population divergence across space. England, including Wales, provides an excellent context to study such effects because as noted above it already had a high urbanization rate. When railways arrived in most European countries after 1825, they had low or moderate rates of urbanization. For example, Belgium was the second most developed economy in Europe and its urbanization rate was 33 in 1850. In fact, England and Wales is the only economy where one can study the effects of railways when the urbanization rate had already reached 40.

Initial urbanization matters because agglomeration effects then become more relevant in determining how transport improvements affect the spatial distribution of economic activity. Commercial and industrial firms would have had greater incentive to location near railway stations because the low cost, high speed network could be easily accessed. That implies that residential population should increase near railway stations because of greater employment opportunities. However, this process of re-location may not be uniform across the initial population distribution. Several theoretical models suggest that if agglomeration is strong

⁷ Consider that Britain's total population grew by 0.8% per year between 1850 and 1900, while in Europe total population increased by 0.7% per year, which is similar (Cameron 1993). However, British urbanization rates went from 40 to 67 between 1850 and 1900 while in Europe they increased from 16 to 30 (Bairoch and Goertz 1986).

⁸ Unfortunately, our population data do not include Scotland or Ireland, and we cannot study Britain or the UK.

enough, then as transport costs decrease from high to moderate levels, the most densely populated areas grow at the expense of the least densely populated areas (see Fujita, Krugman, and Venables 2001, Lafourcade and Thisse 2011). This would suggest that when more densely populated areas got a railway station, the positive effect on their population growth will be larger than when less densely populated areas got a railway station. Population change can beget other economic changes. As one example, land rents rise with greater population, and thus land-intensive sectors become less profitable. This implies the occupational structure should move away from agriculture near stations and into either manufacturing or services.

Building on these ideas, we estimate how being near a railway station in the mid-nineteenth century affected local population and changes in occupational structure over the following decades in England and Wales. Our analysis makes use of uniquely detailed and highly granular dataset. Our 9489 spatial ‘units’ are constructed from parishes and townships, the smallest places reported in the British Census.⁹ We observe populations in every decennial census year from 1801 to 1891 and male occupational shares in agriculture, secondary, and tertiary in 1851 and 1881. We also incorporate highly accurate GIS data on railway lines and stations in each census year, geographic characteristics, like coastline and coal, and pre-rail infrastructure networks like turnpike roads, ports, and inland waterways.

Our baseline specification studies population change from 1851 to 1891 and uses an indicator for having a rail station within a unit’s boundary by 1851 as the main railway access variable. The aim is to estimate the effects of station connection to the main trunk lines open by the mid-nineteenth century. Endogeneity of station location is a major challenge in our analysis, especially as English and Welsh railways were built and owned by private companies pursuing profits (see Casson 2009). As a solution, we construct a use least cost path (LCP) based on the length of the 1851 network and locations of large 1801 towns, which serve as nodes in the LCP. The main instrumental variable or IV is an indicator for having the LCP in a unit. It is a strong predictor for having a station by 1851 and if we restrict the sample to exclude units near the nodes, having the LCP can also be excluded from the second stage analysis for population

⁹ Unfortunately, our population data do not include Scotland or Ireland, and thus we cannot study the whole UK.

growth from 1851 to 1891. The main idea is that the instrument selects “inconsequential units”, which attracted railway lines and stations only because of their favorable location along a route connecting large towns.

The IV estimates imply that having railway stations in a unit by 1851 caused its population to grow by an additional 0.87% per year from 1851 to 1891. The estimated effect is large considering that on average units lost 0.06% in population per year and the standard deviation in annual growth was 1.18%. We also estimate that having a station by 1851 led to a 0.121 *decrease* in the agricultural occupational share between 1851 and 1881 and to a 0.063 *increase* in the secondary occupational share. These effects equal -0.79 and 0.88 standard deviations for agricultural and secondary share changes.

Our main extension speaks to whether railways reinforced the urban hierarchy of the early nineteenth century and contributed to spatial divergence of the English and Welsh population. For this we estimate heterogeneous station access effects based on 1801 population density. We employ a similar identification strategy using the interaction of the LCP dummy and 1801 population variables as instruments. The estimates imply that having 1851 stations increased population growth between 1851 and 1891 more when units had greater log population density in 1801. Using different deciles of the 1801 population distribution, shows lower growth effects of stations if units were in the bottom 70th percentile. In fact, a zero effect of stations in the bottom 70th cannot be rejected, suggesting railways contributed to divergence.

Another extension estimates the size of population displacement effects using varying distance to 1851 stations. We find that being 5 to 15 km from an 1851 station led to lower population growth and increased the share of agricultural occupations compared to being more than 20 km from a station. Thus, there is also evidence for divergence at a local scale near stations.

Our estimates also speak to the economy-wide effects of railways. We predict the 1891 population of 9480 spatial units if none had stations by 1851. The counter-factual total 1891 population is found to be 22% smaller than the actual 1891 population. Moreover, the counter-factual implies that the 1891 population share of the top 5% of units would have been 0.575

rather than 0.687, which accounts for most of its actual change. Concerning male occupations, we estimate that agricultural occupations would have increased by 23% in 1891 if no units had stations by 1851. These effects have implications for labor productivity which we estimate to decline by more than 5% due to reduced economies of density and less structural transformation out of agriculture.

Our results contribute to a large literature on railways and the English and Welsh economy. There are many studies suggesting the importance of railways in affecting local populations.¹⁰ Among the quantitative studies there is agreement that getting railway stations was associated with increased population density.¹¹ However, the causal effects of getting stations have not been established. We address endogeneity by constructing a novel LCP and by analyzing occupational change and heterogeneous effects. Thus, our estimates speak more to how railways fostered population divergence in England and Wales. It also contributes to a more general understanding of spatial divergence in Britain, which despite its remarkable features, is relatively under-studied from a quantitative point of view.¹²

We also contribute to a large comparative historical literature aiming to quantify how proximity to railways affected population and economic change in different countries over the nineteenth and early twentieth centuries.¹³ Several will be discussed more fully below. We make several contributions here. First, with a few exceptions most studies use counties, districts, or cities as their spatial unit.¹⁴ We use small-scale spatial data, approximately at the village or town-level, to study effects of railways. Our study also introduces a richer set of geographic variables, like coal endowments, and a richer set of pre-railway infrastructures like roads and ports. Second, in constructing LCPs as instruments, most studies use straight lines to connect network nodes, however they are not accurate for small-scale spatial data. We use information on historical costs to create an LCP that incorporates slope. Third, several studies

¹⁰ See also Dyos and Aldcroft (1974), Gourvish (1986), Kellet (2012).

¹¹ See Gregory and Martí Henneberg (2010), Casson (2013), Casson et. al. (2013), Alvarez et. al. (2013)

¹² Hanlon (2020) is one of the few studies on nineteenth century city growth.

¹³ See Tang (2014), Hornung (2015), Berger and Enflo (2017), Atack, Bateman, Haines, and Margo (2010), Donaldson and Hornbeck (2016), Hodgson (2018), Jedwab, Kerby, and Moradi (2015), and Donaldson (2018).

¹⁴ An exception is Buchel and Kyburz (2020) who use finely grained spatial data in Switzerland.

analyze effects on firms and investment, but few examine effects on occupational change, one of the key transformations of the nineteenth century.¹⁵ We estimate railways effects on changes in male agricultural, secondary, and tertiary employment. Fourth, our context is unique as we study a highly urbanized economy prior to railways. As we argue it is more important to incorporate heterogenous effects based on initial population size in such settings.

There is another branch of the literature which analyzes the aggregate effects of railway effects through the added consumer surplus from lower freight rates, fares, and higher passenger speeds.¹⁶ Agglomeration is often missing from this framework, which is a limitation because increasing concentration in urban areas is perhaps an important channel by which railways contributed to greater economic growth. Our estimates illuminate this channel and show that is quantitatively significant.

Finally, our results contribute to a broader literature studying the effects of transport infrastructure and regional development.¹⁷ Most focus on local and regional outcomes in recent decades. Historical contexts complement this literature by demonstrating whether infrastructures lead to population gains as well as losses decades after they are built. The English historical context is particularly useful because it is closest to many current settings where infrastructure is built in developed economies with strong agglomeration forces.

The paper is organized as follows. Section 2 provides background. Section 3, 4 and 5 introduce data and methods. Section 6 describes baseline results and sections 7 and 8 examine heterogeneity by 1801 population and displacement. Section 9 focuses on counterfactuals.

2. Background on urbanization and railways

¹⁵ Hornung (2015) studies number and size of firms, Attack, Haines, and Margo (2008) study factories, Tang (2014) studies firm capitalization. Berger (2019) is one of the few to studies occupations occupational change.

¹⁶ See Hawke (1974), Foreman-Peck (1991), and Leunig (2006) for examples. See Heblich, Redding, and Sturm (2018) for a more detailed model of railways effects on London.

¹⁷ A survey is provided by Duranton and Puga (2014). Also see Baum-Snow (2007), Duranton and Turner (2012), Faber (2014), Jedwab et. al. (2015), Garcia-Lopez et. al. (2015), Storeygard (2016), Ghani (2016), Holl (2016), Baum-Snow et. al. (2017), and Gibbons et. al. (2019).

As noted in the introduction, England and Wales was highly urbanized in the nineteenth century compared to other economies. Decadal trends at the top of the urban hierarchy are illustrated in table 1. The share of the population in cities of 20,000 or more grew substantially from 1801 to 1891. The share increased because the total population grew more in cities of 20,000 or more than outside them (see columns 2 and 3). The decades of the greatest divergence in their growth were the 1820s, 1840s, 1870s, and 1880s. Another perspective comes from definitions of urban and rural in the census available starting in 1851.¹⁸ From the 1860s a growing urban population was combined with a shrinking rural population.

Table 1: Decadal trends for the distribution of population in England and Wales, 1801-1891

	(1)	(2)	(3)	(4)	(5)
	population share in cities with 20,000 or more	Growth rate over previous decade, population in cities of 20,000 or more	Growth rate over previous decade, population <i>not</i> in cities of 20,000 or more	Growth over previous decade in <i>urban</i> population as defined by census	Growth over previous decade in <i>rural</i> population as defined by census
year					
1801	0.169				
1811	0.181	22.12	12.71		
1821	0.207	35.33	14.25		
1831	0.249	39.27	9.66		
1841	0.290	32.93	8.39		
1851	0.349	35.94	3.16		
1861	0.382	22.39	6.31	23.15	1.88
1871	0.420	24.47	6.21	28.10	-4.77
1881	0.479	30.49	2.68	25.60	-3.84
1891	0.537	24.98	-0.61	18.48	-2.77

Source: Distribution of Population in England and Wales,
<https://www.le.ac.uk/esh/teach/ug/modules/eh3107/popdist.pdf>

The spatial patterns suggest population growth was very concentrated in England and Wales. Many high growth areas were in the north, near the large industrial centers of Manchester, Liverpool, and Leeds (see Law 1967). By 1901 a densely populated region had formed around them. The other high growth area was near London, industrial Birmingham, and

¹⁸ The census from 1851 to 1871 includes municipal boroughs, towns of improvement acts, and towns of some 2,000 or more inhabitants, without any organization other than the parish vestry" (Census of 1871, Introd., p.xxxi). In 1881 and 1891 it consisted of the urban sanitary districts.

the mining center Cardiff in Wales. However, outside of these ‘hotspots,’ there were few rapidly growing areas in Wales, the south, and east of England. Many villages and small towns had close to zero population growth after 1851.

Differences in net migration were the primary reasons for varying patterns of population growth. To illustrate, Shaw-Taylor and Wrigley (2014) document that between 1801 and 1851 industrial counties grew by 38% more than the average English county, but around 1841 the rate of natural increase (birth rate minus death rate) in industrial counties was only 6% higher than the English average. The implication is that the rate net migration must have been higher in the industrial counties. The primacy of migration was even stronger for counties near London where population growth was above average, yet its rate of natural increase was below average. These calculations support a large literature showing that English and Welsh population was very mobile.¹⁹ Many individuals migrated within regions, but some migrated great distances between regions, especially if a migrant’s birthplace was agricultural. The literature also suggests that better employment opportunities was a key reason for migrating.

Concerning employment there was evolution in occupations that was related to increasing urbanization. The percent of males with agricultural and other primary occupations decreased from 32.4% in 1851 to 25.6% in 1871. The percent in secondary increased from 44.7% in 1851 to 46.3% in 1871. The remaining category, tertiary increased more rising from 22.8% to 28% from 1851 to 1871 (Shaw-Taylor and Wrigley 2014). The increasing occupational share in secondary and tertiary was part of a longer-term process starting with the early stages of the industrial revolution (footnote to explain further).

2.1 Development of railways

England and Wales had the second most dense rail network in the world by 1900 (Belgium was first) and it became a steam driven economy consuming large amounts of coal (Crafts and Mills 2004, Wrigley 2010). But it is important to recognize that England and Wales had a well-developed transport network before railways. It had many good roads suitable for coaches and

¹⁹ See Boyer and Hatton (1997), Long (2005), Schurer and Day (2019).

large wagons and a large inland waterway network for barges. There were also hundreds of ports with wet docks and lighthouses, supporting a thriving coastal shipping trade based on sailing vessels.²⁰

The pre-railway network was created and financed through local and private initiative. Government's role was mainly to approve or reject proposals and regulate user-fees. Railways were developed using this system. Local business groups would introduce a bill in parliament that specified where the proposed railway would go and called for the creation of a joint stock company that would finance its construction and continue with its operation. If approved, local business groups would form a railway company, collect subscription money from investors, and start the construction process. Notably the railway companies did not receive any significant financing from the government. The planning process was largely driven by expectations of private profit, not government funding priorities (Casson 2009).

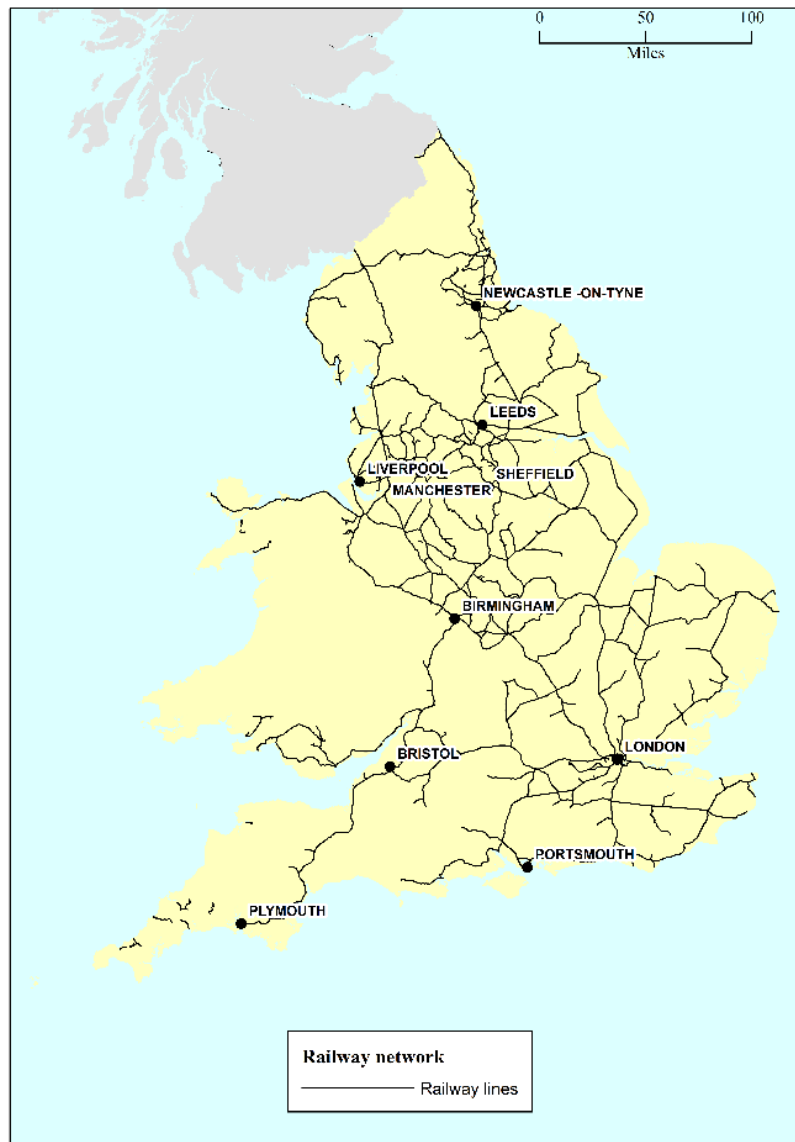
The construction of railways by private companies spanned several decades in England and Wales. The first steam powered passenger rail service opened in 1825 between the towns of Stockton and Darlington. Then in 1830, the Liverpool and Manchester railway opened, followed by several other railways in the mid-1830s. At this early stage, railway companies were mainly interested in connecting the largest urban centers, because they had the most pre-existing passenger and high-value freight services. By 1841, 9 of the 10 largest cities in England and Wales had railway connections, whereas few small and medium towns had railways by 1841.

The 'railway mania' of the mid-1840s saw the biggest expansion of the network. Between 1845 and 1847, 330 Railway Acts were passed to establish new railway companies or extend company networks. The raising of nearly 170 million pounds of capital was more than twice as much as the state spent on the military (Odlyzko, 2010 p.4). According the literature, the mania was partly driven by the early railway company's strategy to maintain their position serving the large cities and by politicians wanting railway stations in their constituencies.²¹

²⁰ For a summary see Bogart (2014).

²¹ For the literature on the railway mania see Casson (2009), Odlyzko (2010), Campbell and Turner (2012, 2015)

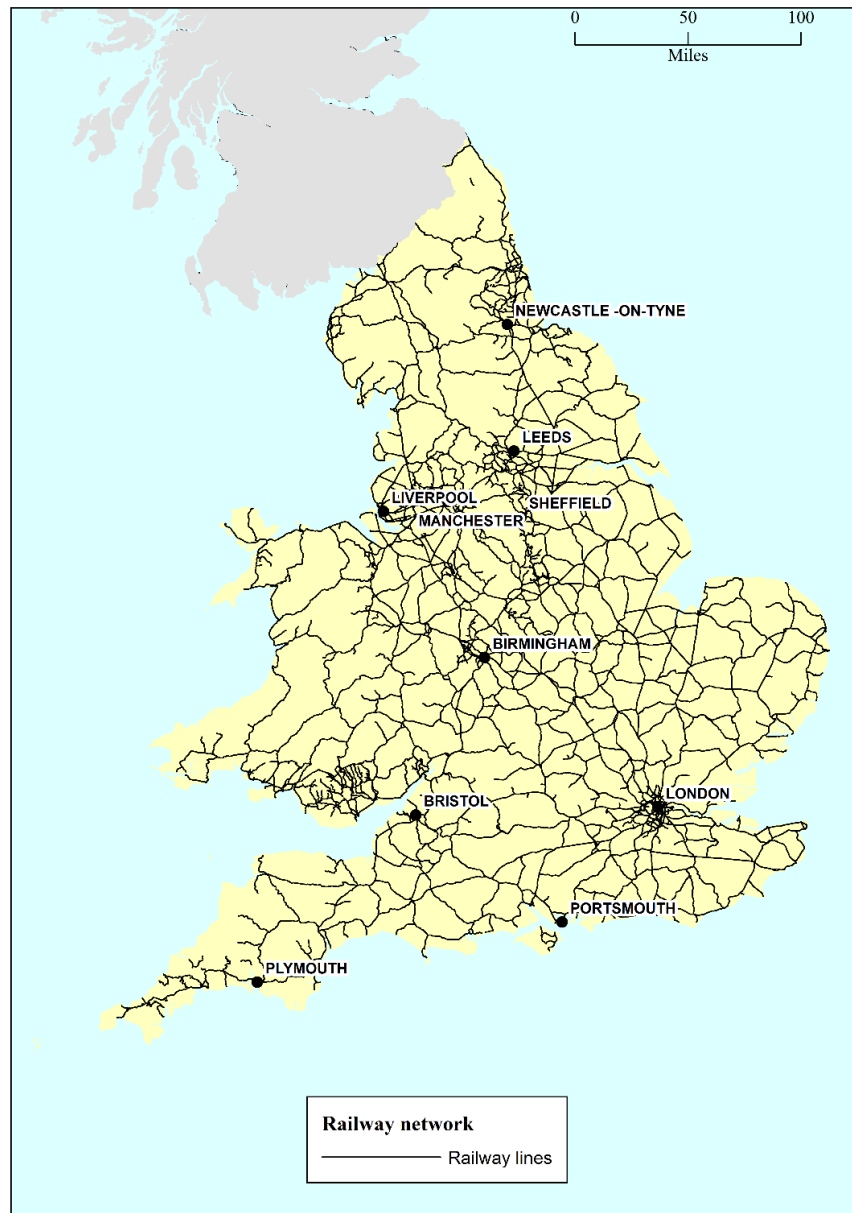
Figure 1. The railway network in 1851



Source: Shaw-Taylor and You.

The significance of the railway mania can be seen in the growth of track mileage. Between 1839 and 1844, railway km grew from 1,560 to 3,456 and between 1845 and 1851 it grew to 10,082 km. The rail network 1851 is shown in figure 1. By this time regional networks had formed around the large towns in addition to connections via the trunk lines. Yet there were still some regions that were under-served, most notably Wales and the southwest.

Figure 2. The railway network in 1881



Source: Shaw-Taylor and You.

The rail network further expanded after 1851 and was nearly 25,000 km in 1881, or twice its size three decades earlier. A map of the network 1881 is shown in figure 2. Railway lines were now in every region of England and Wales. Within these regions there were some towns and rural areas that were better served than others, but none was very far from a railway. By 1881 the network continued to be owned and operated by companies. A process of consolidation

around a few large companies had begun but it would not become significant until after 1900 (Crafts, Leunig, and Mulatu 2008).

Railways came to dominate the internal transport market because they were far superior in both speed and cost. Railways displaced stagecoaches almost immediately when stations opened between destinations. Passenger miles increased at annual rate of 20% and 10% in the 1840s and 1850s. The annual growth rate of passenger miles fell to 5% or less by the 1860s reflecting a rate of increase more in line with GDP growth (Hawke 1970, p. 50).²² In freight, canals offered some competition to railways as barges charged similar freight rates, but they were much slower and less reliable (Maw 2013). Railways poached most of the existing canal traffic after the 1850s. Railways also managed to divert some coastal shipping to its inland network. One revealing statistic is that railways accounted for only 10% of the coal imported into London in 1851. The rest came by sea. But in 1870 railways accounted for 55% of the coal imported to London (Hawke 1970). Improvements to steamships would lead to some reversal to sea, but railways remained an important shipper of coal in London and most towns (Armstrong 2009).

In our analysis one crucial issue relates to the routing of lines and placement of stations. The first and main consideration for lines built in the 1830s and 40s was to connect large cities by the most direct and flat route in order to save construction costs (Simmons 1986, pp. 169-171). Land acquisition costs were another consideration. When railway companies approached urban areas, they often avoided built areas. Sometimes they built lines through slums and working-class neighborhoods because there was less political opposition (Kellet 2012, p. 306, 335). When placing stations along the line, railway companies considered the economic potential in the surrounding area. The review process for railway bills shows they collected information on existing traffic levels and the populations of nearby towns (Casson 2009, Odlyzko 2010). But it was expected that individuals would travel to stations, perhaps as much as 20 or 30 miles. For this reason, companies did not build stations directly in towns, and

²² Another revealing statistic is that there were 0.65 railway journey per head of population in 1841; 20 in 1881 and 32 in 1911. See Mitchell, *British Historical Statistics*, pp. 545-7.

instead placed them at road junctions or near coaching inns, optimal for collecting traffic from the town and its hinterland.

3. Data

Our population data come from British censuses, available every decade starting in 1801. Individuals are counted at the smallest place where they lived, usually the parish or township. The census population counts have been digitized for all ‘census years’ from 1801 to 1911. Male occupational shares for agriculture, secondary production, tertiary, mining, and an unspecified category are also digitized at the smallest census place from 1851 to 1911.²³ Currently, the best occupational data for analysis is in 1851 and 1881.

To address boundary changes, we have created 9764 consistent spatial units between 1801 and 1891 and linked them with census population and occupation data.²⁴ See appendix A.3 for details and maps. A few have missing variables and so our sample size is 9489. Henceforth we call them ‘units’ and each will contain a central point, which we use to calculate various distance variables. The units are quite small in area averaging 15 sq. km. Note that the 9489 units in our data are contained within 55 counties, which were an important administrative unit of local government. The exception are units associated with metropolitan London, which we treat as four ‘counties,’ including south, west, east, and central London.

We also associate each unit with a ‘center’ using GIS. The center corresponds to a town marketplace, if the unit had a town within its boundary at some point between 1600 and 1850.²⁵ If not the centroid is used, which arguably makes sense for a rural unit without a marketplace. Regardless, little error is introduced by using the town market or centroid since our units are only 15 square km on average.

²³ The digitization of the population data is described in Wrigley (2011). Later data comes from the Integrated Census Microdata (I-CeM), 1851-1911. See Schurer and Higgs (2014).

²⁴ We create 9489 consistent units mapping population from 1801 to 1891 and male occupations in 1817, 1851, and 1881. We thank Gill Newton, of the Cambridge Group for History of Population and Social Structure, who developed the Python code.

²⁵ Satchell, Potter, Shaw-Taylor, Bogart (2017) provide a dataset on 1746 towns and their centers. 746 of our units have at least one town in them. If there is a single town, we choose its center. If there are multiple, the town center with the largest 1801 population is used.

Our railway data includes GIS shapefiles for railway lines and stations in every census year starting in 1831. The rail networks and stations are created using highly accurate historical maps.²⁶ From this we create two measures for access to railway stations: (1) an indicator if there was an open station within the boundaries of the unit in a particular year and (2) the distance from the center of each unit to its nearest station in a particular year.²⁷

In addition to railways we create a rich set of variables on ‘first-nature geography.’ These include an indicator for being on exposed coalfields, an indicator being on the coast, average elevation in the unit, average slope in the unit, and standard deviation of slope within the unit, average rainfall, average temperature, wheat suitability, latitude, longitude, and the share of land in 10 different soil types.²⁸ Coastal is identified using an intersection of the seacoast with unit boundaries. The elevation and slope variables are calculated in GIS (see appendix 2 for details). Annual rainfall and temperature (both averaged from 1961 to 1990) and wheat suitability come from FAO.²⁹ Of special significance to the English and Welsh economy, Satchell and Shaw-Taylor (2013) identify those areas with exposed coal bearing strata (i.e. not overlain by younger rocks). Exposed coalfields were more easily exploited compared to concealed coal (see appendix A.3 for details).

Variables for second-nature geography are also incorporated in our analysis. They include distance to one of the ten largest cities in 1801³⁰, log population density in 1801,

²⁶ Martí-Henneberg, Satchell, You, Shaw-Taylor, and Wrigley (2017) created the GIS of England, Wales and Scotland railway stations 1807-1994. It is derived from a railway atlas by Cobb (2005).

²⁷ Note it was rare for stations to close in the nineteenth century (Simmons 1986, p. 325). But it did happen, which means a few units get more distant from stations.

²⁸ Soils data (c) Cranfield University (NSRI) 2017 used with permission. The 10 soil categories are based on Avery (1980) and Clayden and Hollis (1985). They include (1) Raw gley, (2) Lithomorph, (3) Pelosols, (4) Brown, (5) Podzolic, (6) Surface-water gley, (7), Ground-water gley, (8) Man made, (9) peat soils, and (10) other. See http://www.landis.org.uk/downloads/classification.cfm#Clayden_and_Hollis. Brown soil is the most common and serves as the comparison group in the regression analysis.

²⁹ See the Global Agro-Ecological Zones data at <http://www.fao.org/nr/gaez/about-data-portal/agricultural-suitability-and-potential-yields/en/>. We selected low input and rain fed for wheat suitability.

³⁰ The ten largest cities are London, Manchester, Birmingham, Liverpool, Leeds, Bristol, Newcastle, Plymouth, Portsmouth, and Sheffield (near Nottingham)

distance to turnpike roads in 1800, distance to inland waterways in 1800, and distance to ports in 1780. The last four are calculated using detailed pre-rail infrastructure data.³¹

Summary statistics for several variables are shown in table 2. This data reveals several important facts about population and economic change across space. First, despite the total English and Welsh population increasing from 17.9 million to 29 million between 1851 and 1891, the average difference in log 1891 and 1851 population was negative. This is consistent with the share of the population in the top 5% of units increasing from 0.56 in 1851 to 0.69 in 1891.³² In terms of occupational change between 1851 and 1881, the average unit saw a decrease in its male agricultural share and an increase in its male tertiary share. This matches with the national trend to lower shares in agriculture and higher shares in tertiary. However, the average unit had a decrease in the secondary share between 1881 and 1851. This is despite the national trend to slightly higher secondary shares.

Table 2: Summary statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Population and occupation variables					
Diff. Ln. 1831 and 1801 population	9489	0.268	0.247	-1.800	3.126
Diff. Ln. 1891 and 1851 population	9489	-0.023	0.468	-3.388	4.599
Ln pop density 1851	9489	4.242	1.367	0.808	11.625
Diff. 1881 and 1851 male agriculture share	9,488	-0.067	0.153	-0.820	0.928
Diff. 1881 and 1851 male secondary share	9,489	-0.007	0.072	-0.707	0.639
Diff. 1881 and 1851 male tertiary share	9,489	0.045	0.092	-0.700	0.806
Rail variables					
At least one Station in unit by 1851	9489	0.107	0.309	0	1
At least one Station in unit by 1891	9489	0.276	0.447	0	1
Has LCP in unit	9489	0.229	0.421	0	1

³¹ Rosevear, Satchell, Bogart, Shaw Taylor, Aidt, and Leon (2017) created a GIS of turnpike roads, Satchell, Shaw-Taylor, and Wrigley (2017) created a GIS of inland waterways, and Alvarez, Dunn, Bogart, Satchell, Shaw-Taylor (2017) created a GIS of ports.

³² Analysis also shows an inverted U-shaped relationship between the difference in log 1891 and 1851 population and the log of 1801 population density. A plot in the appendix shows that 1891 and 1851 difference increases with 1801 density up until the 90 to 95th percentile and then it begins to decline.

Has stage coaching inn by 1802	9489	0.079	0.269	0	1
Has LCP & stage coaching inn by 1802	9489	0.031	0.174	0	1
<hr/>					
First-nature controls					
Indicator exposed coal	9489	0.080	0.271	0	1
Indicator coastal unit	9489	0.147	0.355	0	1
Elevation	9489	89.72	74.02	-1.243	524.3
Average elevation slope within unit	9489	4.767	3.615	0.484	37.42
SD elevation slope within unit	9489	3.432	2.717	0	23.17
Rainfall in millimeters	9484	755.7	191.7	555	1424
Temperature index	9484	8.958	0.658	5.5	10
Wheat suitability (low input level rain-fed)	9484	2188.1	273.25	272	2503
Latitude	9484	259871	115236	13522	652900
Longitude	9484	443389	112073	136232	654954
Land area in sq. km.	9484	15.63	22.18	0.003	499.8
Perc. of land with Raw gley soil	9489	0.084	1.327	0	76.49
Perc. of land with Lithomorphic soil	9489	8.615	19.83	0	100
Perc. of land with Pelosols soil	9489	8.203	20.63	0	100
Perc. of land with Podzolic soil	9489	4.624	14.32	0	99.56
Perc. of land with Surface-water gley soil	9489	24.63	29.46	0	100
Perc. of land with Ground-water gley soil	9489	10.187	20.11	0	100
Perc. of land with Man made soil	9489	0.363	3.262	0	94.99
Perc. of land with Peat soil	9489	1.187	5.279	0	91.44
Perc. of other soil	9489	0.535	1.966	0	65.15
<hr/>					
Second nature controls					
Ln 1801 population per sq. km	9489	3.877	1.310	0.483	11.43
Distance to inland waterway in 1800 in km	9489	8.121	7.063	0.006	48.67
Distance to turnpike road in 1800 in km	9489	2.431	3.185	0.00	27.95
Distance to port in 1780 in km	9489	33.39	22.33	0.078	99.71
Distance to top 10 city in 1801 in km	9487	68.29	38.69	0	184.14

Sources: see text.

The summary data also inform how station access differed across space and time. In 1851 10.7% of units had at least one station open and by 1891 27.6% of units had at least one

station open. Consistent with these figures, distance to stations fell. In 1851 the median unit was 6.9 km from a station, while the average distance was 10.45 km. In 1881 the median unit was 3.3 km from a station and strikingly the average was only 3.8 km.

As a preview it is useful to remark on how railway access in 1851 is correlated with population growth from 1851 to 1891. A difference in means test shows that the difference in log 1891 and 1851 population is 0.433 higher for units with at least one 1851 rail station versus all other units (p-value 0.00). If the sample is restricted to units below the median for 1801 population density, then the difference in log 1891 and 1851 population is 0.266 higher for units with at least one 1851 rail station versus all other units (p-value 0.00). Thus, the difference in population growth is smaller if units are below the median. Several other correlations are worth mentioning. The 8% of units that had exposed coal had 0.442 higher difference in log 1891 and 1851 pop. compared to other units. This provides one illustration of how endowments matter. With respect to the second nature variables, distance to 1800 turnpike, distance to 1800 waterway, distance to 1800 port, and distance to top 10 1801 city are all negatively and significantly related to the difference in log 1891 and 1851 pop.

4. Methodology

There are two common econometric models for estimating the effects of infrastructure on population. This section explains how they are used in our analysis. The first model analyzes the effect of changes in infrastructure on simultaneous changes in population, hereafter changes-on-changes. As explained by Duranton and Puga (2014), the changes-on-changes model is akin to assuming that populations are in equilibrium. It estimates the change in equilibrium population implied by the change in infrastructure. The second model analyzes infrastructure levels and their effects on changes in population going forward, hereafter changes-on-levels. It is akin to assuming an adjustment process where every year the population comes closer to the equilibrium. It estimates population change over a specific time period implied by the base infrastructure level.

The historical railways literature has employed both econometric models. We think the adjustment process is more reasonable in studying change over several decades and so we give more attention to changes-on-levels. A common specification is a cross-section like equation (1)

$$y_{it+k} - y_{it} = \lambda * y_{it} + \beta * railaccess_{it} + x_i + \varepsilon_{it} \quad (1)$$

where y_{it} is locality i 's log population in t and $y_{it+k} - y_{it}$ is the difference in log population $t+k$ and log population t . An alternative dependent variable is annual population growth from t to $t+k$ but the difference in log population is similar. The vector x_{it} includes controls like resources, infrastructures built before railways, and prior development measures besides initial population captured by y_{it} on the right-hand side. The main variable $railaccess_{it}$ is usually defined as having a railway station within locality i 's boundaries in the year t . The idea is that rail transport services were so much cheaper or faster that many industrial and commercial firms had to be very near stations to be competitive. Stations gave them better access to consumers in other markets and helped reach low-cost suppliers. On the workers side, many had to live very near stations because of jobs, and because commuting costs were too prohibitive to live elsewhere. As a result of the positive net-migration near stations, the expectation is that having a railway station in a locality should cause its population to grow more than in localities without railway access all else equal.

Regarding the timing, it is common in this literature to estimate effects of 'first-wave' rail construction on population change over the next 20 to 50 years. For example, Hornung (2015) studies the effect of Prussian railway stations built by 1848 and estimates that they increased city population growth by 2.1% per year from 1849 to 1871. Büchel, and Kyburz (2020) study the effect of Swiss railway stations built by 1864 and estimates they increased municipal population growth by 0.6% per year from 1850 to 1900. There is a potential concern that the effects of subsequent railway building, say after 1848 or 1864, affected growth. To address this problem a control for later railway access is often included. However, even if not, it is still informative to estimate the persistence effects of getting first-wave access.

A related version of change-on-levels examines persistence effects of first-wave access over several decades using panel data. Specification (2) is one version

$$y_{it} = \alpha_i + \delta_t + \beta_t \text{railaccess}_i + \tau_t x_i + \varepsilon_{it} \quad (2)$$

where y_{it} is locality i 's log population in t , α_i is a locality fixed effect, and δ_t is a year fixed effect omitting the first year of the panel, which serves as the base year. The coefficients β_t is the year t log difference in population for locality i with rail access relative to its base year and relative to units without rail access. In the Swedish context, Berger and Enflo (2017) use this specification to estimate that having first-wave railway lines in towns by 1870 led to a 50 log point increase in town population between 1850 and 1900.

Endogeneity of rail access is a major issue in the literature and several instruments have been proposed. The most common is to an indicator for having least cost path (LCP) pass through a locality. The LCP connects pre-selected cities likely to serve as endpoints for a rail line. The first applications used straight lines to connect endpoints (e.g. Atack, Bateman, Haines and Margo 2010), but subsequent studies use slope and geographic impediments to create the LCP (e.g. Berger 2019). The key idea is that some localities became close to railway lines simply because they were on the route designed to connect larger towns at a low capital cost.³³ The analysis usually drops all localities within a certain distance of the end-points as they are positively selected for railway development. It is also common to test for pre-trends, specifically whether having the LCP affects population growth in the early nineteenth century conditioning on other factors. A rejection of the LCP effect supports the plausibility of the exclusion restriction.³⁴

Another common test concerns population displacement effects. The hypothesis is that beyond some distance to station threshold, population growth is lower in localities closer to stations than in localities very far from stations. It builds on the theory that there are advantages to being either very close to centers of economic activity or being far away (Fujita, Krugman, and Venables 2001). Locations in-between face too much competition from the center and cannot sustain their economic firms and businesses. The size and range of the

³³ Redding and Turner (2014) call this the 'inconsequential places' approach. See Chandra and Thompson (2000), Michaels (2008), Faber (2014), and Lipscombe et. al. (2013) for early applications.

³⁴ An alternative approach is to matching localities based on pre-railway characteristics. Matching is less common perhaps because it supposes that conditional on observables being similar, the assignment of railway access was random.

displacement zone has been estimated with a modified version of equation (1) using rail distance-bins like 0 to 2 km, 2 to 4 km, up to some cutoff. For example, Büchel, and Kyburz (2020) use this approach to show that Swiss municipalities between 2 and 10 km from railway lines built by 1864 had between 0.12% and 0.28% lower annual population growth between 1850 and 1900 than the comparison group beyond 12 km distance. By comparison units between 0 to 2 km distance had 0.19% higher annual growth than the comparison group.

We incorporate all these methodologies and estimate the effects of 1851 railway station access on population growth from 1851 to 1891. The main trunk lines were open by approximately 1850 and many had stations spread across these lines. Therefore, having a station in 1851 provides a measure of access to a network connecting most of the large cities and towns of the early nineteenth century.

Our baseline specification (3) is a cross section regression of population growth

$$\Delta \ln pop_{ij,1891-1851} = \beta * Station1851_i + \gamma \cdot x_i + c_j + \varepsilon_{ij} \quad (3).$$

where the dependent variable $\Delta \ln pop_{ij,1891-1851}$ is the difference in log 1891 and 1851 population for unit i in county j . As explained in the next section, we observe population in every decade from 1801 to 1891. The main variable, $station1851_i$, is a dummy that equals 1 if unit i has at least one open station within its boundary by 1851 and zero otherwise. The control group is all units without station access by 1851. We will also include variables for later station access as robustness checks. The vector x_i always includes the natural log of population density in 1851, 1841, and 1831 to capture effects of base year population levels and prior trends in population growth. In some specifications, x_i includes our rich set of ‘first nature’ and ‘second nature’ characteristics. The variables c_j are 59 county fixed effects, which capture common growth patterns at higher administrative units of local government

The instrument for unit 1851 station access in equation (3) is going to be an indicator for having the LCP pass through the unit boundary. The LCP is created using historical construction cost information combined with elevation data (section 6 explains in detail). We also introduce

a second instrument, having coaching inns by 1801. It is designed to capture where along the route of the LCP stations are most likely.

One of our key extensions is to estimate heterogeneous effects of station access. Most of the historical railway literature fails to find significant heterogeneous effects. However, we think early nineteenth century population density is likely to matter in the English and Welsh context. Drawing on theory of agglomeration economies, we postulate that when railways stations arrived in low 1801 density units they brought increased competition from high density units who were more productive. The greater competition could result in employment losses and offset some of the positive net-migration effects from getting stations. The expectation is that station effects on population growth will be small or close to zero for low 1801 density units, and much larger for high 1801 density units. The heterogeneous effects are also important in assessing railways role in population divergence. If there is a positive interaction between $Station1851_i$ and variables for greater 1801 population density, then this would imply railways contributed to divergence because it increased population more in the most dense units.

We also study occupational change using the difference in 1881 and 1851 male agricultural, secondary, or tertiary employment shares as the dependent variables in equation (3).³⁵ The idea is that land-intensive economic sectors, like agriculture, should become less profitable relative to labor-intensive sectors as population grows. Therefore, the occupational share in agriculture is expected to decline in locations with station access. If so, the occupational share must necessarily rise in other sectors, but there is not a clear prediction on whether it will rise in secondary, tertiary, or both. If being near railway stations was especially important for firms purchasing inputs and selling their goods to customers, then having a station should increase secondary shares. If timely access to information was especially important in the production of services, then having a station should increase tertiary shares.

The last methodological point relates to our use of changes-on-changes specifications. Most previous studies analyze outcome changes across two points in time. One example is Berger (2019) who studies Swedish trunk railroad lines built between 1850 and 1900. Berger

³⁵ Berger (2019) is the only paper that we know of which analyzes effects of rail access on occupational change.

estimates that being within 5 km of a Swedish state trunk railroad lines by 1900 led to a 0.066 increase in the share of industrial employment between 1850 and 1900. This is a changes-on-change specification because it uses connection by 1900 not an earlier date. We think the most informative changes-on-changes specification in our setting analyzes the longer term, as in equation (5)

$$\Delta \ln pop_{ij,1891-1821} = \beta * \Delta station_{i,1891-1821} + \pi x_i + \varepsilon_{it} \quad (5)$$

where the dependent variable $\Delta \ln pop_{ij,1891-1821}$ is the difference in log 1891 and 1821 population. On the right hand side $\Delta station_{i,1891-1821}$ is the difference in railway station access in 1891 versus year 1821. Since rail station access is zero in 1821 for all units, $\Delta station_{i,1891-1821}$ is simply a dummy variable equal to 1 if a unit had an open station in 1891. The variable x_i includes first nature controls, second nature controls, and county fixed effects as all of these could affect the change in population. The instrument for having an LCP is again used to address endogeneity in 1891 station access. We now turn to a description of the LCP as it plays a crucial role.

4. The Least cost path and its properties

While there are various approaches to creating a least cost path (LCP), we develop one to fit our setting. As there was no official railway plan in England, the first step is to select town-pairs that will likely be connected by early railways. We start with all English and Welsh towns having a population greater than 5000 in 1801.³⁶ Their larger size meant they were almost certain to get at least one railway line connecting them with another town above 5000. But not all large town-pairs would be connected. A profit-seeking company would see little value in building a railway to connect distant towns of a moderate size. We use a simple gravity model to approximate the relative value of connecting all town-pairs each with a population above 5000. The gravitation value G_{ij} for town pairs i and j is $G_{ij} = (pop_i * pop_j) / dist_{ij}$, where $dist_{ij}$ is the straight line distance between town i and j . We ordered G_{ij} from largest to smallest and connect all pairs with a value greater than a threshold defined momentarily.

³⁶ The town population data come from Law (1967) and Robson (2006).

The second step is to identify routes connecting the selected town-pairs. We assume that in considering their routes, railway companies tried to minimize the construction costs considering distance and elevation slope. We use construction cost data for railways built in the 1830s and early 1840s. We also measure the distance of the lines and total elevation changes between towns at the two ends of the line. The construction cost is then regressed on the distance and the elevation change to identify the parameters (the details are in appendix A.1). Based on this analysis, we find a baseline construction cost per km when the slope is zero and for every 1% increase in slope the construction cost rises by three times the baseline (cost per km = $1 + 3 * \text{slope\%}$). We use this formula and GIS tools to identify the least cost path (LCP) connecting each town pair with a population above 5000 in 1801.

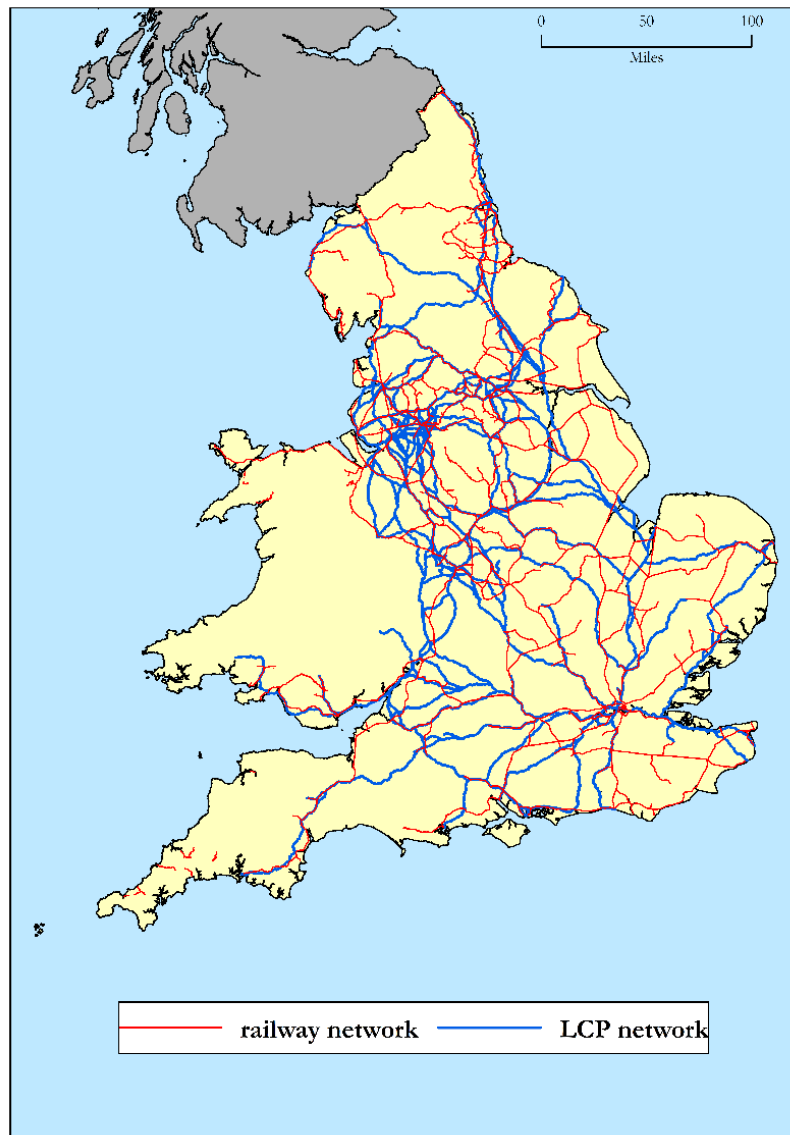
The third step is to identify the routes included in the rail LCP network. Our method is as follows. First, we start with the LCP route associated with the largest gravitational value G . Second, we add the LCP route associated with the second largest G . If the two routes are close to one another we combine duplicate sections. We continue in the same manner adding LCP routes until the total LCP network size equals the size of the 1851 network. For clarity, we label as 'LCP nodes' all town points selected to construct the LCP based on their gravitational value. In GIS, the nodes are points and thus the nodes will be close to our unit centers in some cases. We make extensive use of nodal locations below.

The LCP network and actual 1851 railway network are shown in figure 3. The overlap is almost exact in some cases. There is also statistical evidence for overlap in our units. There is a 0.323 correlation between an indicator for having railway lines pass through a unit in 1851 and an indicator for having the LCP pass through a unit. It should be noted that railway lines built after 1851 are close to the LCP too, but the overlap is weaker. For example, there is a 0.279 correlation between having a railway line in 1861 and having the LCP.

Another important fact is that many units with a station in 1851 also have an LCP. There is 0.251 correlation between the indicator for having the LCP and the indicator for having stations by 1851. The reason is that stations were so numerous along the railway line in England and Wales. On average there was one station for every 5.9 km of railway line in 1851. Also, our

data show that 24.4% of units had a railway line in 1851 and 10.7% had at least one station in 1851. Therefore about 1 out of every 2 or 3 units with railway lines also had stations.

Figure 3: The LCP network and 1851 rail network compared



Sources: see text.

We have more variables to instrument for stations. Building on the idea that railway companies often built stations near nodes of the pre-existing network, we use indicator for having a coaching inn in 1802. This comes from *Cary's New Itinerary*, which was a book for

travelers identifying routes and inns to rest. There were 1228 inns throughout England and Wales by 1802 and these have digitized and linked to GIS.³⁷ We interact inns in 1802 with the indicator for having the LCP in a unit (see table 2 for summary statistics). There is 0.228 correlation between this interaction variable and the indicator for having stations by 1851.³⁸

It is important for our analysis that LCP variables are not statistically related to population growth before railways arrived. To show this, table 3 report estimates from specifications where the dependent variable is the difference in log 1831 and 1801 population. In the columns (1) to (3) the indicator having the LCP is the main explanatory variable. In columns (4) to (6) it is log distance to the LCP. The specifications also differ based on the controls. In column (1) and (4) first nature variables and ln 1801 population density are included. In column (2) and (5) county fixed effects are added. In column (3) and (6) second nature controls are further added. The standard errors are clustered on counties in all specifications. The sample includes units whose center is more than 7 km from an LCP node for reasons explained momentarily. The specifications show that the indicator for having the LCP is not significantly associated with population growth from 1801 to 1831. The same for distance to the LCP. In the appendix we also show that the LCP interacted with coaching inns by 1802 is not statistically associated with the difference in log 1831 and 1801 population.

Table 3: Effects of stations and LCP on population growth in pre-railway era

Panel A	(1)	(2)	(3)	(4)	(5)	(6)
Estimator	OLS	OLS	OLS	OLS	OLS	OLS
Dependent variable:	$\Delta \ln \text{pop}, 1831-1801$					
LCP in unit	-0.00140 (0.00854)	0.0103 (0.00789)	0.00542 (0.00802)			
Log distance to LCP				0.00560 (0.00377)	-0.000461 (0.00397)	0.00361 (0.00413)
County FE	N	Y	Y	N	Y	Y
Second Nature	N	N	Y	N	N	Y
Observations	8,337	8,337	8,337	8,337	8,337	8,337
R-squared	0.064	0.110	0.116	0.065	0.110	0.116

Notes to table 2: Standard errors in parentheses are clustered on counties. *** $p < 0.01$, ** $p < 0.05$, *

³⁷ We thank Alan Rosevear for digitizing coaching inns from Cary.

³⁸ We also tried an instrument for the length of LCP divided by land area. But once we condition on having the LCP, this variable did not predict having an 1851 station.

$p < 0.1$. All models include first nature variables and In pop 1801 density as controls. For definitions of first and second nature variables see table 1. For definitions of county fixed effects see the text. All units less than 7 km from an LCP node are dropped.

We choose to restrict the sample to units more than 7 km from LCP nodes because at shorter distances having the LCP is positively associated with population growth from 1801 to 1831.³⁹ The sample restriction means losing 12% of observations and not studying the effects of stations very close to major 1801 towns, which are the nodes. Moreover, we cannot study station effects in the highest population density units, which also tended to be closer to LCP nodes. But importantly we still have observations in all 1801 population density deciles.⁴⁰

6. Estimates for baseline specifications

The baseline estimates for the effects of having stations by 1851 on the difference in log 1891 and 1851 population are shown in the top panel of table 4. The bottom panel shows the first stage coefficient for having an LCP. The specification in columns (1) and (2) include first nature controls and the log of population density in 1851, 1841 and 1831. Columns (3) and (4) add the county fixed effects and columns (5) and (6) further add the second nature controls. The standard errors are always clustered on counties. In all the specifications the Kleibergen-Paap F-stat is above 48, and so weak instruments are not a problem. The IV estimate for 1851 stations in (6) is 0.349. It is our preferred estimate as this specification controls for several confounding factors. When county fixed effects are omitted the IV estimates are much large (see column 2), which makes sense because the LCP is not evenly distributed across space and nor was population growth. When the second nature variables are omitted, the IV estimate are also larger (see column 4). Among these variables, distance to major 1801 cities and distance to an 1800 waterway, are most correlated with the LCP and they too affect population growth. After conditioned on them, we have a more credible causal estimate of the 1851 station effect.

³⁹ To illustrate, consider specifications similar to column (6) in table 2 where we add dummy variables for being 0 to 1 km from the LCP node, 1 to 2 km from the LCP node and so on up to 14 to 15 km from the LCP node and interactions between these 15 variables and the indicator for having the LCP in the unit. The coefficients and confidence intervals for the LCP and 15 interaction variables are shown in the appendix. Briefly they reveal that for some distances less than 7 km, having the LCP is significantly associated with higher population growth. No such effects are found for having an LCP at distances more than 7 km from the node.

⁴⁰ In our restricted sample, 5.1%, 9.7 and 10.3% are in the 10th, 9th, and 8th deciles of 1801 population density.

It is notable that the IV estimates are consistently larger than OLS. There are two possible explanations. First, getting stations in 1851 was associated with ‘worse’ 1851 to 1891 growth prospects after accounted for other factors, like population growth in the early 1800s. Second, there may be a local average treatment effect associated with having an LCP. Unfortunately, it is difficult to tell which explanation is more likely, but evidence below suggests IV is more of a local treatment effect.

Table 4: OLS and IV Cross-sectional estimates for effect of getting a station by 1851 on population growth from 1851 to 1891

OLS and IV for log difference 1891 and 1851 population						
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Station in unit by 1851	0.231*** (0.0292)	0.956*** (0.175)	0.178*** (0.0211)	0.473** (0.197)	0.166*** (0.0213)	0.349* (0.206)
County FE	N	N	Y	Y	Y	Y
Second Nature	N	N	N	N	Y	Y
Kleibergen-Paap F stat		95.428		57.289		48.939
Observations	8,341	8,341	8,341	8,341	8,337	8,341
R-squared	0.193		0.287		0.304	
First stage for Station in unit by 1851						
	(7) OLS	(8) OLS	(9) OLS			
LCP in unit	0.101*** (0.0146)	0.0800*** (0.0130)	0.0737*** (0.0132)			
County FE	N	Y	Y			
Second Nature	N	N	Y			
Observations	8,341	8,341	8,337			
R-squared	0.188	0.211	0.216			

Notes: Standard errors in parentheses are clustered on counties. *** p<0.01, ** p<0.05, * p<0.1. All specifications include first nature variables and 1851, 1841, and 1831 ln pop density as controls. For definitions of second nature variables see table 1. County FEs are described in the text. All units less than 7 km from an LCP node are dropped.

The IV estimates suggest large causal effects of railway stations. The station coefficient in column (6) is equivalent to 0.75 standard deviations of the dependent variable or in annual growth terms an increase of 0.87%. The effects of stations are also large compared to other variables (the coefficient estimates are in the appendix). For the specification in (6) the coefficients for having coal and being a coastal unit were 0.171 and 0.168, respectively. The coefficient on distance to the nearest top ten 1801 city in km is -0.0028. This is 1/124 of the IV

station effect, meaning getting a railway station in 1851 was equivalent to moving a unit 124 km closer to a major city, or like moving a unit from the midlands of England to near London.

The estimates for stations reported in table 4 are quite robust. If we restrict the sample to units more than 8 km from LCP nodes and re-run specification (6) in table 4, the IV coefficient for 1851 stations is 0.392 (S.E.0.205). Similar results are found for larger or smaller restrictions as shown in the appendix. We also estimate specifications with a second instrument: the LCP dummy interacted with the dummy for having stage coaching inns by 1802. In the first stage, we find that having the LCP and coaching inns increases the probability of getting an 1851 station by 0.086 (S.E. 0.036). In the second stage, the IV coefficient for 1851 stations is 0.300 (S.E.0.174, Kleibergen-Paap F-stat 24.77). The IV coefficient is a bit smaller than 0.349 from column 6 table 4, but the general conclusion is the same.

Table 5: OLS and IV Cross-sectional estimates for effect of getting a station at different dates on population growth from 1851 to 1891

	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Station by 1856	0.183*** (0.0205)	0.376* (0.222)						
Station by 1861			0.198*** (0.0177)	0.366* (0.218)				
Station by 1866					0.203*** (0.0145)	0.356* (0.209)		
Station by 1871							0.214*** (0.0161)	0.353* (0.202)
County FE	Y	Y	Y	Y	Y	Y	Y	Y
Second Nature	Y	Y	Y	Y	Y	Y	Y	Y
Kleibergen-Paap F stat		37.643		34.880		36.280		34.534
Observations	8,337	8,337	8,337	8,337	8,337	8,337	8,337	8,337
R-squared	0.309		0.317		0.321		0.328	

Notes: Standard errors in parentheses are clustered on counties. *** p<0.01, ** p<0.05, * p<0.1. All specifications include first nature variables and 1851, 1841, and 1831 ln pop density as controls. For definitions of second nature variables see table 1. County FEs are described in the text. All units less than 7 km from an LCP node are dropped.

In our main specification, the control group includes some units that did not have stations by 1851 but would get them later. If we use indicators for station open in 1856, 1861, 1866, and 1871 then some of these units will move into the treated group. In table 5 we show

estimates using indicators for having at least one station open in 1856, 1861, 1866, and 1871 instead of stations open in 1851. The OLS estimates get larger as station date increases. This makes sense because the control group has fewer units that would get stations later. However, the IV estimate is largely unaffected. For example, in IV having a station by 1871 is estimated to have caused a 0.353 increase in the difference of log 1891 and 1851 population. That is nearly identical to the estimated effect in table 3 column (6). Moreover, if we drop units that would get their first station after 1851, the IV estimate for 1851 station is 0.337 (S.E. 0.195), again very similar. Our explanation is that IV is probably capturing a local treatment effect.⁴¹

To help illuminate population changes over time we also estimate one specification like equation (2) using panel data from 1831 to 1891.⁴² The coefficients and 95% confidence intervals are plotted in figure 4. They show that 1881 is first year where having 1851 stations led to a significant difference with 1831 population equal to 43.4 log points. The difference becomes larger by 1891 reaching 52.6 log points. These estimates show that the gap between units with and without 1851 stations was widening with time and persisted into the 1880s.

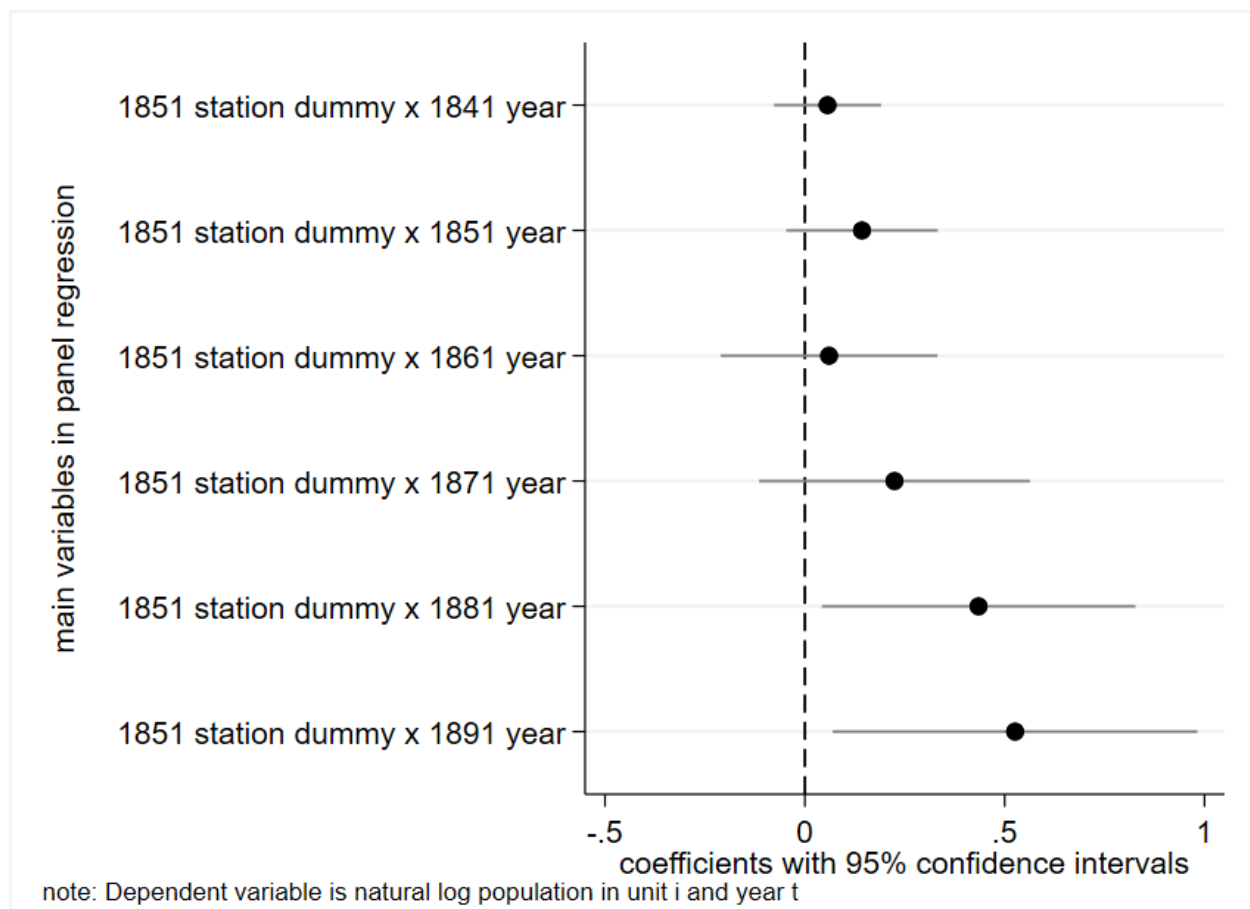
The estimates from our ‘change on change’ specification give a quantitative perspective on the equilibrium change in population between 1821 and 1891 caused by the change in station access over the same 70-year period. The sample includes all units more than 7 km from an LCP node and the controls include first nature variables, second nature variables and county fixed effects. The estimates are shown in the appendix. Briefly, the IV coefficient for the difference in station access between 1891 and 1821 is 0.605 (SE 0.274, Kleibergen-Paap F-stat 29.75). If converted into an annual growth effect, it implies that having a railway station increased population by 0.87% per year. Notice this is nearly identical to estimated increase in

⁴¹ Another robustness check uses propensity score matching with railway access in 1851 as the treatment variable and the log difference in 1891 and 1851 population is the outcome. The simplest specification matches on a single variable: the log of 1801 population density. The matched sample is balanced and yields a statistically significant treatment effect of 0.323 (S.E. 0.029), which is very similar to our 0.349 IV estimate in table 3 column 6. Unfortunately, we were unable to achieve balanced matching on many co-variables. But, if we match on all second nature variables or selected first nature variables, the treatment effects are similar.

⁴² The controls included unit fixed effects, year fixed effects, first and second nature variables interacted with year fixed effects, and county fixed effects interacted with year fixed effects. The variables of interest are the 1851 station access dummy interacted with year fixed effects post 1841. The instruments are indicators for having the LCP times the year fixed effects post 1841. As throughout the standard errors are clustered on counties.

annual population growth implied by the IV coefficient for 1851 station access in table 4 column 6. Using changes-on-changes specifications does not change the conclusion.

Figure 4. Panel estimates for effect of 1851 station on the log of unit population



Notes: the coefficients are from panel specifications ranging from 1831 to 1891. It includes unit fixed effects, year fixed effects, first and second nature variables interacted with year fixed effects, and county fixed effects interacted with year fixed effects. All units less than 7 km from an LCP node are dropped.

Next, we analyze the effect of 1851 stations on changes in male occupational structure. The specifications are like the cross-sectional approach in table 4, except here we use the difference in 1881 and 1851 male agricultural, secondary, or tertiary shares as the dependent variable. We also add controls for 1851 male shares in agricultural, secondary, tertiary, mining, or unspecified occupations to condition on occupational structure when railways were

beginning to open. The coefficient estimates are reported in table 6. Columns (1) and (2) show that getting stations in 1851 led to a significant decline in male agricultural shares. The IV coefficient -0.124 is equivalent to -0.80 standard deviations in the dependent variable. Columns (3) and (4) show that getting stations led to a significant increase in male secondary shares. The IV coefficient 0.066 is equivalent to 0.89 standard deviations. Columns (5) and (6) show smaller or less precise effects for tertiary shares. These estimates imply that railway stations also led to occupational change, reducing employment in land-intensive economic sectors and increased employment in labor-intensive sectors, mainly manufacturing.

Table 6: OLS and IV Cross-sectional estimates for effect of getting a station by 1851 on the difference in male occupational shares 1881 and 1851

Estimator	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
Dependent variable:	Δ male agriculture occupational share		Δ male secondary occupational share		Δ male tertiary occupational share	
Station in unit by 1851	-0.0422*** (0.00465)	-0.124** (0.0612)	0.0114*** (0.00286)	0.0667** (0.0339)	0.0253*** (0.00339)	0.0384 (0.0447)
Kleibergen-Paap F stat		48.139		48.139		48.139
Observations	8,337	8,337	8,337	8,337	8,337	8,337
R-squared	0.393		0.212		0.341	

Notes: Standard errors in parentheses are clustered on counties. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All specifications include county fixed effects, first nature variables, second nature variables, 1851, 1841, and 1831 \ln pop density as controls, and 1851 male shares in agricultural, secondary, tertiary, mining, or unspecified occupations. For definitions of first and second nature variables see table 1. The instrument for station in unit by 1851 is an indicator if unit has LCP in its boundaries. All units less than 7 km from an LCP node are dropped.

6. Heterogenous effects based on 1801 population

The effects of stations need not be uniform across units and with agglomeration economies they were likely to be greater if the population density was higher before railways. In this section, we estimate heterogenous effects depending on where the unit was in the 1801 population density distribution. It was more common for high 1801 density units to get 1851 stations, but it is reassuring that in our restricted sample, excluding units more than 7km from nodes, 72% of units in the top decile did not have a station by 1851. Recall that throughout the

population 90% did not have stations. We also observe some low-density units with stations in 1851. For example, 5.5% of units in the 1st, 2nd, and 3rd deciles had a station by 1851.

Agglomeration economics often stresses a density threshold, which result in different effects from transport improvements (see Lafourcade and Thisse 2011). Therefore, our preferred specification is to use binary measures of 1801 population density based on being above or below some percentile. However, it is more general to start with an interaction between 1851 stations and log 1801 density. Columns (1) and (2) in table 7 show these estimates for the baseline. The IV specification uses the LCP dummy interacted with the log of 1801 population density as the second instrument. The positive coefficients on the interaction term implies the effects of 1851 stations were greater with higher 1801 density. To interpret the coefficients, we predict the difference in log 1891 and 1851 population for units getting or not getting 1851 stations and according to different log 1801 population densities. At the median density having a station led to 0.21 increase. At the 25th and 75th percentiles the increases were 0.14 and 0.33. The appendix has a graph for the entire distribution.

Table 7: Heterogeneous effects of getting a station by 1851 on population growth from 1851 to 1891

Estimator	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) IV	(8) IV
Dependent variable:								
Station by 1851	-0.035 (0.101)	-0.668 (0.643)	0.214*** (0.0248)	0.555*** (0.195)	0.186*** (0.0261)	0.522*** (0.187)	0.609*** (0.227)	0.588*** (0.219)
Station by 1851* Ln 1801 pop density	0.051** (0.024)	0.250* (0.132)						
Station by 1851* Below 60 th pct. pop den. 1801			-0.108** (0.0414)	-0.497*** (0.172)			-0.494*** (0.221)	
Station by 1851* Below 70 th pct. pop den. 1801					-0.0420 (0.0379)	-0.350** (0.165)		-0.358 (0.220)
Drop units with more than 1 station	N	N	N	N	N	N	Y	Y
Kleibergen-Paap F stat		17.433		19.433		19.958	12.588	13.693
Observations	8,377	8,377	8,337	8,337	8,337	8,337	8,172	8,172
R-squared	0.305		0.307		0.307			

Notes: Standard errors in parentheses are clustered on counties. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. All specifications include county fixed effects, first nature variables, second nature variables, 1851, 1841, and 1831 ln pop density as controls. For definitions of first and second nature variables see table 1. All units less than 6 km from an LCP node are dropped. In (2) the instruments are has LCP and has LCP interacted with coal. In (4) we add the instrument has LCP interacted with dummy for below median 1801 population. In (6) we add the instrument has LCP interacted with distance to a major 1801 city.

The remaining columns in table 7 report specifications using binary measures of 1801 population density based on being above or below the 60th or 70th percentile. The LCP dummy interacted with the binary measures of density are the second instruments. In column (4), IV estimates imply the effect of getting stations was 0.497 lower if the unit had density below the 60th percentile. For these units the overall effect of stations was to increase the log difference by 0.059 (0.555-0.497), but it is not statistically different from zero. A null effect also cannot be rejected for units below the 70th percentile (see column 6).

It is possible that the heterogenous effects are due to getting multiple stations in high density units. While this is potential channel, it cannot entirely account for the heterogenous effect. To see this column (7) and (8) in table 7 restrict the sample to units with zero stations or only 1 station in 1851. The coefficients are strikingly similar, meaning that greater 1801 density increased the effects of stations even if only one station was open in 1851.

The heterogenous effects are robust to considering alternative dates for getting stations. In the appendix we report estimates from the change on change specification. It shows that getting a station by 1891 had a significantly larger effect on the log difference in 1891 and 1821 population if the unit had a higher 1801 density. Moreover, units in the bottom 60% of 1801 density did not experience significant gains from railways.⁴³ Overall, these findings are consistent with the logic of new economic geography models, which stress that peripheral areas will lose to core areas when transport costs begin to decline from high levels. We should also stress that accounting for heterogenous effects changes the quantitative interpretation. For units in the top 40% of 1801 density, getting stations led to an additional increase in annual population growth of 1.4%. Units in the bottom 60% got only a 0.1% increase in annual growth.

⁴³ For the change on change specification, units in the bottom 70% experienced an 0.390 log difference in population. Units in the top 30% experienced an 0.78 log difference in population.

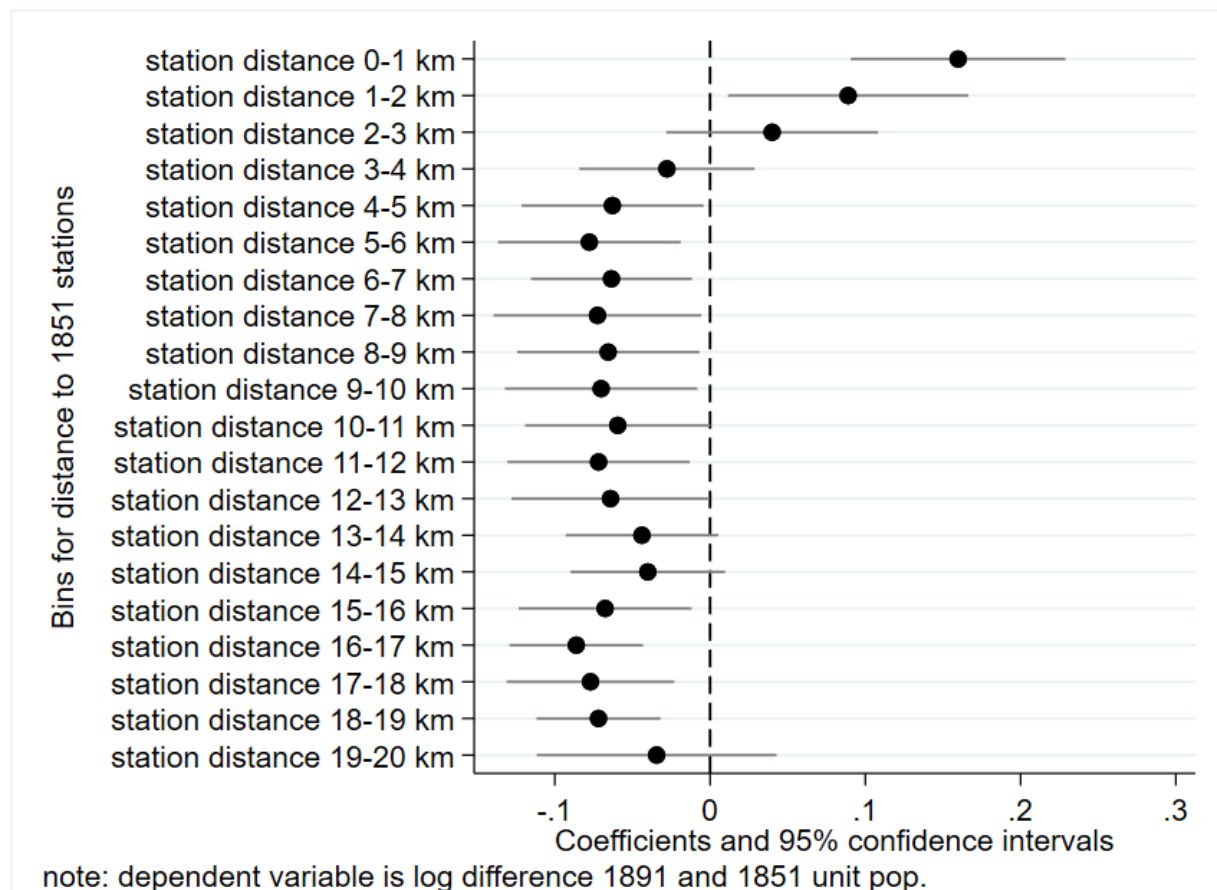
8. Local population displacement effects

An additional factor to consider is that population growth may have been different depending on how far units were from stations in 1851. Of most interest is the possibility that after some distance threshold growth was lower in units closer to stations compared to units farther from stations. We first test for local displacement effects using variables for 1851 station distance-bins like 0 to 1 km, 1 to 2 km and so on up to 19 to 20 km. The specification includes all control variables and as before drops units less than 7 km from an LCP node. The coefficients and 95% confidence intervals for each distance bin are plotted in figure 5. The interesting results are for units more than 4 km from stations. The estimates imply they had a 0.05 to 0.10 lower difference of log 1891 and 1851 pop. compared to units more than 20 km from stations. Thus, these estimates support a displacement zone between 4 and 20 km from an 1851 station.

In the appendix we show that the coefficient plot is similar if we use distance to 1861 or 1871 station bins rather than distance to 1851 station bins. The standard errors get a larger, but still they imply statistically lower growth between 4 and 20 km distance. Thus, the general finding of a displacement zone is not affected by when we measure distance to stations open at later dates.

A similar plot for the effect of station distance bins on the change in agricultural occupations is shown in figure 6. The most interesting result is that being more than 4 km and less than 10 km led to an increase in the agricultural share by around 0.03. These results suggest that in the population displacement zone, the occupational structure became more land intensive. Similar estimates for secondary and tertiary share do not show such a stark change around 4 km. This implies that in the displacement zone, neither sector experienced an especially large decline in their shares.

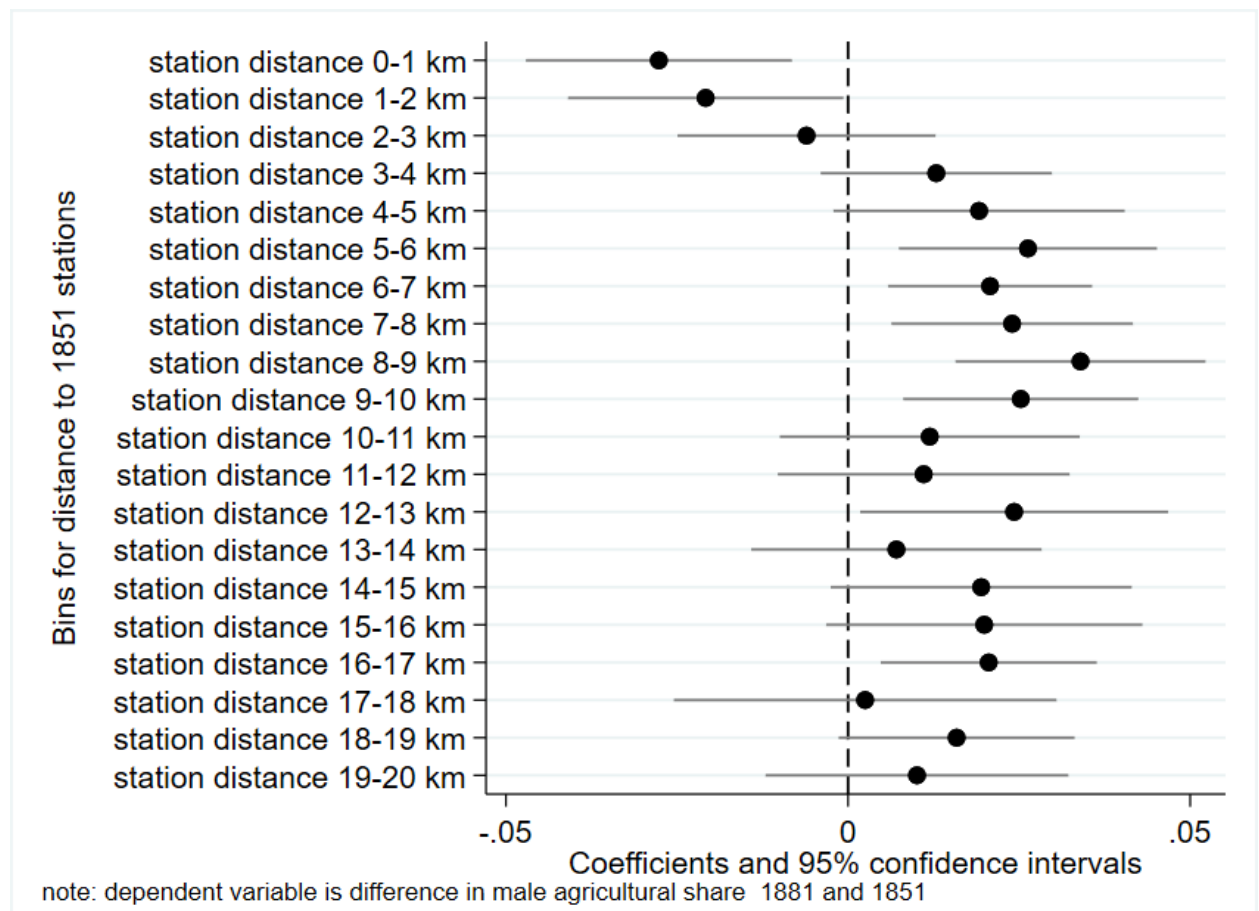
Figure 5. Effects of distance to 1851 stations on log difference in 1891 and 1851 population



Notes: the coefficients are from specifications that include county fixed effects, first nature variables, second nature variables, 1851, 1841, and 1831 ln pop density as controls. For definitions of first and second nature variables see table 1. All units less than 7 km from an LCP node are dropped.

As in earlier sections, endogeneity is a concern for interpreting these coefficients. However, the large number of station parameters in the previous two figures make it impractical to use instruments. Instead, we use log distance to stations and its square as the endogenous variable and log distance to LCP and its square as the instruments. The specification is otherwise identical to the baseline model studying the log difference in 1891 and 1851 population.

Figure 6. Effects of distance to 1851 stations on change in 1881 and 1851 agricultural shares

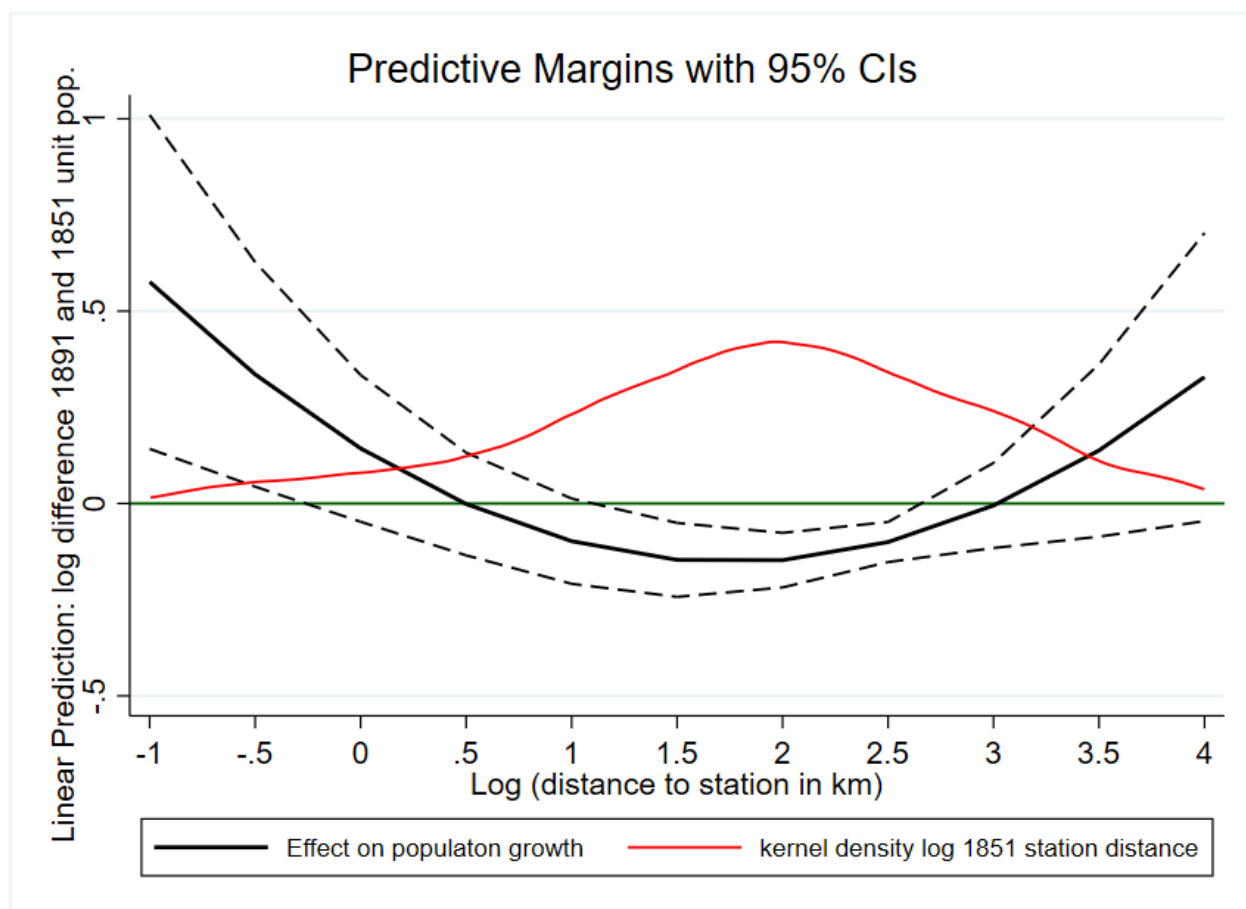


Notes: the coefficients are from specifications that include county fixed effects, first nature variables, second nature variables, 1851, 1841, and 1831 ln pop density, and occupational shares in five categories as controls. For definitions of first and second nature variables see table 1. All units less than 7 km from an LCP node are dropped.

The IV estimates for the effects of log station distance and its square support the finding of local population displacement. The findings are best summarized with a graph. Figure 6 shows the estimated effect of log 1851 station distance between -1 and 4 along with the kernel density of the distribution of log 1851 station distance. Population growth is large and statistically different from zero for all units less than 0.5 or around 1.6 km distance. Population growth becomes negative and statistically different from zero for all units between 1.25 and 2.7

or 3 to 15 km distance. Comparing these effects with the kernel density (in red) shows that based on distance only a minority of units experienced positive population effects from railways, around 15%. Many more, around 60%, experienced negative population effects.

Figure 6. IV estimates for effect of 1851 station distance on population growth



Sources: Author's calculation, see text.

9. Discussion: Economy-wide population change and productivity

One classic question concerns how much railways changed the whole of the British economy. The best attempts to measure the consumer surplus suggest the gains from shipment of freight were around 5 to 10% of GDP in 1890 (Foreman-Peck 1991) and the gains to passengers in money and time saved were around 5% in 1890 (Leunig 2006). What do our estimates imply for this debate? In this section we estimate how the population size and

distribution would have been different if railways had not been invented. Then we quantify some aggregate effects for labor income.

In the first step, total population levels in 1891 are predicted using the specification shown in column 4 of table 7 with the indicator for 1851 stations and its interaction with units in the bottom 60th percentile of the 1801 population distribution. For predictions, we switch to the full population rather than the restricted sample of units more than 7 km from nodes. The fit is reasonably good even in the full sample. The correlation between the actual population in 1891 and that predicted by our specification is 0.85.

In the second step, counter-factual population levels in 1891 are estimated if no unit had a station in 1851. The estimates imply that total 1891 population in England and Wales would have been about 22% lower if no units had stations by 1851. Also the population would have been less concentrated. Our estimates suggest the share of the 1891 population in the top 5% of units would have been 0.575 rather than 0.687. This estimate suggest that railways can account for nearly all the change in population concentration. Recall that between 1851 and 1891 the actual population share in the top 5% rose from 0.564 to 0.687

Population changes caused by the railway depended on the 1801 population distribution. In the counterfactual most units in the bottom 90% of 1801 would have had higher population in 1891 without railways. Our estimates suggest the median would have had 1.2% higher population in 1891. In the top decile the effect of railways varied much more. The median in the top 1801 decile would have 13.6% lower population in 1891.

In the third step, we estimate changes in occupational structure if there was no rail access in 1851. Males in agricultural occupations in 1881 are predicted using the specification shown in column 2 of table 6 with the indicator for 1851 stations. In cases where the predicted share is negative, the number of agricultural males is set to zero in a unit. The correlation between the actual number of males with agricultural occupations in 1881 and that predicted by our specification is 0.68, which suggest more noise than for population, but not too bad. Our estimates suggest there would have been 23.3% more males in agricultural occupations in 1881 if no units had railway stations in 1851. Also, there would have been 6.7% fewer males

with secondary occupations and 9.8% fewer males with tertiary occupations. In other words, much less structural transformation.

What are the broader implications for national labor income? With 22% lower population, GDP would be significantly smaller. One could estimate about 11% lower assuming a labor share in income of 0.5 as is common in national accounting (Crafts and Mills 2004). The lower concentration of population in large units had implications for productivity too. We use Leunig and Crafts (n.d.) estimate that the elasticity of labor productivity with respect to own population density was 0.025. We then calculate each unit change in productivity from the population change caused by no 1851 stations and then calculate weighted average using 1891 population weights. This calculation implies that by shifting the population to lower density units, labor productivity in the English and Welsh economy would have fell by 0.58%. This effect is significant but not too large compared to the national income loss from total population change.

However, the productivity implications of occupational change are much larger. If we use Boyer and Hatton's estimate that rural unskilled wages were 27.2% lower in real terms than urban wages, then a 23.3% increase in agricultural male workers would represent a $23.3 \times (-0.272) = -6.33\%$ loss in male wage income. We don't know how women's occupations changed in response to railways but if they were similar the total loss in wage income would have been above 5%. Broadly our estimate point to a large impact of railways on the English and Welsh economy mainly through less structural transformation.

7. Conclusion

In this paper, we study how railways led to population change and divergence in an already urbanized economy, England and Wales. We make use of detailed data on railway lines, stations, and population change in 19489 spatial units. Endogeneity is a major challenge in our context given that private companies built the network and every indication suggests that profit motives were central. To address this issue, we create a least cost path based on major 1801 towns and the length of the 1851 rail network. Our instrumental variable estimates show that having railway station in a locality by 1851 led to significantly higher population growth from

1851 to 1891 and shifted the male occupational structure away from agriculture. Moreover, in extensions, we estimate that having stations increased population growth more if localities had greater population density in 1801. Also, there were population losses for localities 5 to 15 km from stations, indicating a displacement effect. Overall, we find that railways reinforced the urban hierarchy of the early nineteenth century and contributed to further spatial divergence.

How do our findings relate to other studies in the comparative historical literature on railways? One of the few to also analyze finely grained spatial data is Buchel and Kyburz (2020), who estimate the effects of railways on population change in Switzerland. Employing a similar IV framework, these authors show that having a ‘first-wave’ station in the mid-nineteenth century increased a Swiss municipality’s annual population growth by 0.6%. Buchel and Kyburz also find evidence for local displacement 2 to 8 km from stations. By comparison, we find that in England and Wales having stations increased annual population growth by 0.87%, with displacement effects reaching 15 km in distance. We also find large effects from stations in localities that entered the nineteenth century with denser populations. Thus, our results show that railway contributed to more divergence in England and Wales. The greater strength of agglomeration economies in this economy is likely to be one reason for the difference.

Several studies in the comparative railway literature focus on industrialization and firms, but few examine occupational structure like here.⁴⁴ One exception is Berger (2019) who studies railways and occupational change in Sweden. Employing a similar IV framework, Berger shows that having a trunk railway line in a parish increased its manufacturing occupational share by 0.066. We find a nearly identical estimate for male secondary employment, which suggests that in two different environments, railways contributed to greater employment in manufacturing.

What are the implication for the evolution of the British economy? Through studying the effects on population concentration and occupational change, we find that railways effects on national income are larger than is suggested by methods relying only on estimating consumer surplus from lower freight rates and higher speeds. This suggests one needs to employ several approaches to identifying the full effects of a major transportation change like railways.

⁴⁴ Hornung (2015) studies number and size of firms, Attack, Haines, and Margo (2011) study factories, Tang (2014) studies firm capitalization.

References

Alvarez, Eduard, Xavi Franch, and Jordi Martí-Henneberg. "Evolution of the territorial coverage of the railway network and its influence on population growth: The case of England and Wales, 1871–1931." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 46.3 (2013): 175-191.

Alvarez, E, Dunn, O., Bogart, D., Satchell, M., Shaw-Taylor, L. , 'Ports of England and Wales, 1680-1911', 2017.
<http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html>

Armstrong, John. *The Vital Spark: The British Coastal Trade, 1700-1930*. International Maritime Economic History Association, 2009.

Atack, Jeremy, Fred Bateman, Michael Haines, and Robert A. Margo. "Did railroads induce or follow economic growth?." *Social Science History* 34, no. 2 (2010): 171-197.

Atack, Jeremy, Michael R. Haines, and Robert A. Margo. *Railroads and the Rise of the Factory: Evidence for the United States, 1850-70*. No. w14410. National Bureau of Economic Research, 2008.

Atack, Jeremy, and Robert A. Margo. "The Impact of Access to Rail Transportation on Agricultural Improvement: The American Midwest as a Test Case, 1850-1860." *Journal of Transport and Land Use* 4.2 (2011).

Avery, Brian William. *Soil classification for England and Wiles: higher categories*. No. 631.44 A87. 1980.

Bairoch, Paul, and Gary Goertz. "Factors of urbanisation in the nineteenth century developed countries: a descriptive and econometric analysis." *Urban Studies* 23.4 (1986): 285-305.

Baldwin, Richard E., and Philippe Martin. "Agglomeration and regional growth." *Handbook of regional and urban economics*. Vol. 4. Elsevier, 2004. 2671-2711.

Baum-Snow, Nathaniel. "Did highways cause suburbanization?." *The Quarterly Journal of Economics* 122.2 (2007): 775-805.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., & Zhang, Q. (2017). Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics*, 99(3), 435-448.

Beach, Brian, and W. Walker Hanlon. "Coal smoke and mortality in an early industrial economy." *The Economic Journal* 128.615 (2018): 2652-2675.

Berger, Thor, and Kerstin Enflo. "Locomotives of local growth: The short-and long-term impact of railroads in Sweden." *Journal of Urban Economics* (2015).

Bogart, Dan. "The Transport Revolution in Industrializing Britain," in Floud, Roderick, Jane Humphries, and Paul Johnson, eds. *The Cambridge Economic History of Modern Britain: Volume 1, Industrialisation, 1700–1870*. Cambridge University Press, 2014.

Boyer, George R., and Timothy J. Hatton. "Migration and labour market integration in late nineteenth-century England and Wales." *Economic History Review* (1997): 697-734.

Büchel, Konstantin, and Stephan Kyburz. "Fast track to growth? Railway access, population growth and local displacement in 19th century Switzerland." *Journal of economic geography* 20.1 (2020): 155-195.

Bureau of Railway News and Statistics. *Railway Statistics of the United States of America*. Chicago: R. R. Donnelley and Sons, 1913 and 1916.

Cameron, R. *Concise Economic History of the World* (New York: O.U.P., 1993) p. 193.

Campbell, Gareth, and John D. Turner. "Dispelling the Myth of the Naive Investor during the British Railway Mania, 1845–1846." *Business History Review* 86.01 (2012): 3-41.

Campbell, Gareth, and John D. Turner. "Managerial failure in mid-Victorian Britain?: Corporate expansion during a promotion boom." *Business History* 57.8 (2015): 1248-1276.

Cary, John. *Cary's New Itinerary: Or an Accurate Delineation of the Great Roads, Both Direct and Cross Throughout England and Wales; with Many of the Principal Roads in Scotland. From an Actual Admeasurement by---; Made by Command of His Majesty's Postmaster General, for Official Purposes. Under the Direction and Inspection of Thomas Hasker (etc.)*. Gosnell, 1802.

Casson, Mark. *The world's first railway system: enterprise, competition, and regulation on the railway network in Victorian Britain*. Oxford University Press, 2009.

Casson, Mark. "The determinants of local population growth: A study of Oxfordshire in the nineteenth century." *Explorations in Economic History* 50.1 (2013): 28-45.

Casson, Mark, A.E.M. Satchell, Leigh Shaw-Taylor, and E.A. Wrigley, "Railways and local population growth: Northampton and Rutland, 1801-1891" in Casson, Mark, and Nigar Hashimzade, eds. *Large databases in economic history: research methods and case studies*. Routledge, 2013.

Chandra, Amitabh, and Eric Thompson. "Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system." *Regional Science and Urban Economics* 30.4 (2000): 457-490.

Church, Roy, Alan Hall, and John Kanefsky. History of the British Coal Industry: Volume 3: Victorian Pre-Eminence. Vol. 3. Oxford University Press, USA, 1986.

Clayden, Benjamin, and John Marcus Hollis. Criteria for differentiating soil series. No. Tech Monograph 17. 1985.

Cobb, M. H. "The Railways of Great Britain: A Historical Atlas at the Scale of 1 Inch to 1 Mile. 2 vols." *Shepperton: Allen* (2006).

Cormen, Thomas H., Charles E Leiserson, Ronald L Rivest and Clifford Stein: Introduction to Algorithms, Cambridge, MA, MIT Press (3rd ed., 2009) pp.695-6.

Crafts, Nicholas, and Timothy Leunig. "Transport improvements, agglomeration economies and city productivity: did commuter trains raise nineteenth century British wages?. working paper, n.d.

Crafts, Nicholas, and Terence C. Mills. "Was 19th century British growth steam-powered?: the climacteric revisited." *Explorations in Economic History* 41.2 (2004): 156-171.

Crafts, Nicholas, Timothy Leunig, and Abay Mulatu. "Were British railway companies well managed in the early twentieth century? 1." *The Economic History Review* 61.4 (2008): 842-866.

Desmet, Klaus, and Esteban Rossi-Hansberg. "Spatial development." *The American Economic Review* 104.4 (2014): 1211-1243.

Donaldson, Dave. "Railroads of the Raj: Estimating the impact of transportation infrastructure." *American Economic Review* 108.4-5 (2018): 899-934.

Donaldson, Dave, and Richard Hornbeck. "Railroads and American economic growth: A "market access" approach." *The Quarterly Journal of Economics* 131.2 (2016): 799-858.

Duranton, Gilles, and Matthew A. Turner. "Urban growth and transportation." *The Review of Economic Studies* 79.4 (2012): 1407-1440.

Duranton, G., & Puga, D. (2014). The growth of cities. In *Handbook of economic growth* (Vol. 2, pp. 781-853). Elsevier.

Faber, Benjamin. "Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System." *Review of Economic Studies* 81.3 (2014): 1046-1070.

Fernihough, Alan, and Kevin Hjortshøj O'Rourke. Coal and the European industrial revolution. No. w19802. National Bureau of Economic Research, 2014.

Foreman-Peck, James. Railways and late Victorian economic growth. Cambridge University Press, 1991.

Fujita, Masahisa, Paul R. Krugman, and Anthony Venables. The spatial economy: Cities, regions, and international trade. MIT press, 2001.

Gibbons, Stephen, Teemu Lyytikäinen, Henry G. Overman, Rosa Sanchis-Guarner. "New road infrastructure: the effects on firms." *Journal of Urban Economics* 110 (2019): 35-50.

Gourvish, Terence Richard. Railways and the British economy, 1830-1914. Macmillan International Higher Education, 1980.

Gregory, Ian N., and Jordi Martí Henneberg. "The railways, urbanization, and local demography in England and Wales, 1825–1911." *Social Science History* 34.2 (2010): 199-228.

Hanlon, W. Walker. "Coal smoke, city growth, and the costs of the industrial revolution." *The Economic Journal* 130.626 (2020): 462-488.

Hsiang, Solomon M. "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America." *Proceedings of the National Academy of sciences* 107.35 (2010): 15367-15372.

Hawke, Gary Richard. Railways and economic growth in England and Wales, 1840-1870. Clarendon Press, 1970.

Heblich, Stephan, Stephen J. Redding, and Daniel M. Sturm. The Making of the Modern Metropolis: Evidence from London. No. w25047. National Bureau of Economic Research, 2018.

Hodgson, Charles. "The effect of transport infrastructure on the location of economic activity: Railroads and post offices in the American West." *Journal of Urban Economics* 104 (2018): 59-76.

Hornung, Erik. "Railroads and growth in Prussia." *Journal of the European Economic Association* 13.4 (2015): 699-736.

Garcia-López, Miquel-Àngel, Adelheid Holl, and Elisabet Viladecans-Marsal. "Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities." *Journal of Urban Economics* 85 (2015): 52-67.

Ghani, Ejaz, Arti Grover Goswami, and William R. Kerr. "Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing." *The Economic Journal* 126.591 (2016): 317-357.

- Holl, Adelheid. "Highways and productivity in manufacturing firms." *Journal of Urban Economics* 93 (2016): 131-151.
- Jedwab, Remi, Edward Kerby, and Alexander Moradi. "History, path dependence and development: Evidence from colonial railroads, settlers and cities in Kenya." *The Economic Journal* (2015).
- Jarvis A., H.I. Reuter, A. Nelson, E. Guevara (2008). Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT), available from <http://srtm.csi.cgiar.org>.
- Kellett, John R. *The impact of railways on Victorian cities*. Routledge, 2012.
- Lafourcade, Miren, and Jacques-François Thisse. "New economic geography: the role of transport costs." *A handbook of transport economics*. Edward Elgar Publishing, 2011.
- Law, Christopher M. "The growth of urban population in England and Wales, 1801-1911." *Transactions of the Institute of British Geographers* (1967): 125-143.
- Leunig, Timothy. "Time is money: a re-assessment of the passenger social savings from Victorian British railways." *The Journal of Economic History* 66.3 (2006): 635-673.
- Lipscomb, Molly, Mushfiq A. Mobarak, and Tania Barham. "Development effects of electrification: Evidence from the topographic placement of hydropower plants in Brazil." *American Economic Journal: Applied Economics* 5.2 (2013): 200-231.
- Long, Jason. "Rural-urban migration and socioeconomic mobility in Victorian Britain." *The Journal of Economic History* 65.1 (2005): 1-35.
- Maw, Peter. *Transport and the Industrial City: Manchester and the Canal Age, 1750 1850*. Manchester University Press, 2013.
- Martí-Henneberg, J., Satchell, M., You, X., Shaw-Taylor, L., Wrigley E.A., 'England, Wales and Scotland railway stations 1807-1994 shapefile' (2017). <http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html>
- Michaels, Guy. "The effect of trade on the demand for skill: Evidence from the interstate highway system." *The Review of Economics and Statistics* 90.4 (2008): 683-701.
- Mitchell, Brian R. *Economic development of the British coal industry 1800-1914*. Cambridge University Press, 1984.
- Mitchell, Brian R. *British Historical Statistics*, Cambridge University Press, 1988.

Odlyzko, Andrew. "Collective hallucinations and inefficient markets: The British Railway Mania of the 1840s." University of Minnesota (2010).

Pascual Domènech, P. (1999). Los caminos de la era industrial: la construcción y financiación de la red ferroviaria catalana, 1843-1898 (Vol. 1). Edicions Universitat Barcelona.

Pooley, Colin, and Jean Turnbull. Migration and mobility in Britain since the eighteenth century. Routledge, 2005.

Poveda, G. (2003). El antiguo ferrocarril de Caldas. Dyna, 70 (139), pp. 1-10.

Purcar, Cristina. "Designing the space of transportation: railway planning theory in nineteenth and early twentieth century treatises." Planning Perspectives 22.3 (2007): 325-352.

Redding, Stephen J., and Matthew A. Turner. "Transportation costs and the spatial organization of economic activity." Handbook of regional and urban economics. Vol. 5. Elsevier, 2015. 1339-1398.

Riley, S. J., S. D. Gloria, and R. Elliot (1999). A terrain Ruggedness Index that quantifies Topographic Heterogeneity, Intermountain Journal of Sciences, 5(2-4), 23-27.

Robson, Brian T. Urban growth: an approach. Vol. 9. Routledge, 2006.

Rosevear, A., Satchell, M., Bogart, D., Shaw Taylor, L., Aidt, T. and Leon, G., 'Turnpike roads of England and Wales, 1667-1892', 2017.

<http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html>

Satchell, M. and Shaw-Taylor, L., 'Exposed coalfields of England and Wales' 2013.

<http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html>

Satchell, M., Shaw-Taylor, L., Wrigley E.A., '1830 England and Wales navigable waterways shapefile' (2017).

<http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html>

Satchell, M., Potter, E., Shaw-Taylor, L., Bogart, D., 'Candidate Towns of England and Wales, c.1563-1911', 2017. A description of the dataset can be found in M. Satchell, 'Candidate Towns of England and Wales, c.1563-1911 GIS shapefile'

<http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html>

Shaw-Taylor, Leigh, and Xuesheng You. "The development of the railway network in Britain 1825-1911." N.d.

Schurer, K., Higgs, E. (2014). Integrated Census Microdata (I-CeM), 1851-1911. [data collection]. UK Data Service. SN: 7481, <http://doi.org/10.5255/UKDA-SN-7481-1>.

Schürer, K., and Joe Day. "Migration to London and the development of the north–south divide, 1851–1911." *Social History* 44.1 (2019): 26-56.

Storeygard, Adam. "Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa." *The Review of Economic Studies* 83.3 (2016): 1263-1295.

Shaw-Taylor, Leigh, E.A. Wrigley, Peter Kitson, Ros Davies, Gill Newton and Max Satchell. *The occupational structure of England and Wales c.1817-1881*. Working Paper, 2010.

Shaw-Taylor, L. and Wrigley, E. A. "Occupational Structure and Population Change," in Floud, Roderick, Jane Humphries, and Paul Johnson, eds. *The Cambridge Economic History of Modern Britain: Volume 1, Industrialisation, 1700–1870*. Cambridge University Press, 2014.

Simmons, Jack. *The railway in town and country, 1830-1914*. (1986).

Sugden, Keith, Sebastian Keibek and Leigh Shaw-Taylor. "Adam Smith revisited: coal and the location of the woollen manufacture in England before mechanization, c. 1500-1820", *CWPESH* no. 33, 2018.

Tang, John P. "Railroad expansion and industrialization: evidence from Meiji Japan." *The Journal of Economic History* 74.03 (2014): 863-886.

Wellington, A.M. *The Economic Theory of the Location of Railways: An Analysis of the Conditions Controlling the Laying Out of Railways to Effect the Most Judicious Expenditure of Capital*. Ed. J. Wiley & sons, 1877.

Wrigley, Edward Anthony. *Energy and the English industrial revolution*. Cambridge University Press, 2010.

Wrigley, Edward Anthony, and Edward Anthony Wrigley. *The early English censuses*. Oxford University Press, 2011.

Appendix A.1: The least cost path instrument

In this appendix, we describe how we identify the LCP connecting our nodes. The main criteria used to plan linear projects is usually the minimization of earth-moving works. Assuming that the track structure (composed by rails, sleepers and ballast) is equal for the entire length, it is in the track foundation where more differences can be observed. Thus, terrains with higher slopes require larger earth-moving and, in consequence, construction costs become higher (Pascual 1999, Poveda 2003, Purcar 2007). The power of traction of the locomotives and the potential adherence between wheels and rails could be the main reason. Besides, it is also important to highlight that having slopes over 2% might imply the necessity of building tunnels, cut-and-cover tunnels or even viaducts. The perpendicular slope was also crucial. During the construction of the track section, excavation and filling have to be balanced in order to minimize provisions, waste and transportation of land. Nowadays, bulldozers and trailers are used, but historically workers did it manually. It implied a direct linkage between construction cost, wages and availability of skilled laborers. In fact, it is commonly accepted in the literature that former railways were highly restricted by several factors. The quality of the soil, the necessity of construction tunnels and bridges or the interference with preexistences (building and land dispossession) were several. Longitudinal and perpendicular slope were the more significant ones and we focus on these below.

Slopes are determined using elevation data. Several DEM rasters have been analyzed in preliminary tests, but we finally chose the Shuttle Radar Topography Mission (SRTM) obtained in 90 meter measurements (3 arc-second). Although being a current raster data set, created in

2000 from a radar system on-board the Space Shuttle, the results offered in historical perspective should not differ much from the reality. The LCP tool calculates the route between an origin and a destination, minimizing the elevation difference (or cost in our case) in accumulative terms. The method developed was based on the ESRI Least-Cost-Path algorithm, although additional tasks were implemented to optimize the results and to offer different scenarios. The input data was the SRTM elevation raster, converted into slope. This conversion was necessary in order to input different construction costs.

The next step is to specify the relationship between construction costs and slope. One approach is to use the historical engineering literature. Wellington (1877) discusses elevation slope (i.e. gradients), distance, and operational costs of railways, but this is not ideal as we are interested in construction costs. We could not find an engineering text that specified the relationship between construction costs and slopes. As an alternative we use historical construction cost data. The following details our data and procedure.

A select committee on railways in 1844 published a table on the construction costs of 54 railways. See the Fifth report from the Select Committee on Railways; together with the minutes of evidence, appendix and index (BPP 1844 XI). The specific section with the data is appendix number 2, report to the lords of the committee of the privy council for trade on the statistics of British and Foreign railways, pp. 4-5. There were 45 with a clear origin and destination, to which we can measure total elevation change along the route (details are available). For these 45 railways we calculate the distance of the railway line in meters and the total elevation change (all meters of ascent and descent). We then ran the following regression of construction costs on distance in 100 meters and the elevation change in meters. This

regression produces unsatisfactory results, with total elevation change having a negative sign. We think the main reason is that the sample includes railways with London as an origin and destination. Land values in London were much higher than elsewhere and thus construction costs were higher there. Therefore, we omit railways with a London connection. We also think it is important to account for railways in mining areas as they were typically built to serve freight traffic rather than a mix with passenger.

Our extended model uses construction costs for 36 non-London railways. We regress construction costs on a distance in 100 meters, elevation change, and dummy for mining railways. The results imply that for every 100 meters of distance construction costs rise by £128.9 (st. err 45.27) and holding distance constant construction costs rise by £382.6 (st. err. 274.5) for every 1 meter increase in total elevation change. Construction costs for mining railways are £340,418 less (st. err. 179,815). For our LCP model we assume a non-mining railway, re-scale the figures into construction costs per 100 meters, and normalize so that costs per 100 meters are 1 at zero elevation change. The formula becomes:

$$\text{NormalizedCostper100meters} = 1 + 2.96 * (\text{ElevationChangeMeters} / \text{Distance100meters})$$

The elevation change divided by distance can be considered as the slope in percent, in which case our formula becomes $\text{Cost} = 1 + 2.96 * \% \text{slope}$. We think this is a reasonable approximation of the relationship between construction costs, distance, and elevation slope.

The LCP algorithm is implemented using ESRI python, using as initial variables the elevation slope raster, the reclassification table of construction costs, and the node origin-destination nodes. We implemented the least-cost-path function to obtain the LCP corridors.

These corridors were converted to lines, exported, merged and post-processed. Maps of our preferred LCP are shown in the text.

Appendix A.2: Elevation, slope, and ruggedness variables

The aim of this appendix is to explain the creation of the elevation variables, including the original sources and method we followed to estimate them. There are several initiatives working on the provision of high-resolution elevation raster data across the world. The geographical coverage, the precision of the data and the treatment of urban surroundings concentrate the main differences between databases.

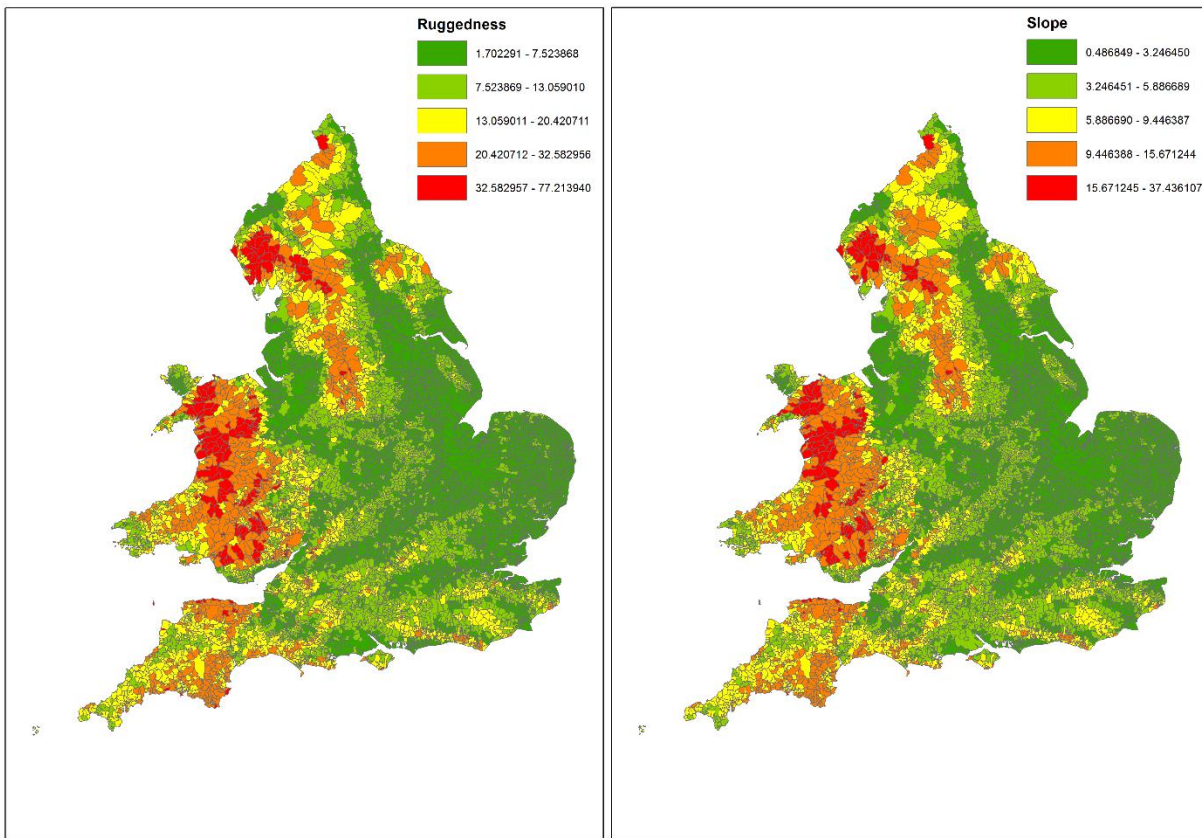
We obtained several elevation DEM rasters, preferably DTM , covering the entire England and Wales. In decreasing order in terms of accuracy, the most precise one database was LIDAR (5x5m.), Landmap Data set contained in the NEODC Landmap Archive (Centre for Environmental Data Archival). In second instance, we used EU-DEM (25x25m.) from the GMES RDA project, available in the EEA Geospatial Data Catalogue (European Environment Agency). The third dataset was the Shuttle Radar Topography Mission (SRTM 90x90m), created in 2000 from a radar system on-board the Space Shuttle Endeavor by the National Geospatial-Intelligence Agency (NGA) and NASA. And finally, we have also used GTOPO30 (1,000x1,000m) developed by a collaborative effort led by staff at the U.S. Geological Survey's Center for Earth Resources Observation and Science (EROS). All those sources have been created using satellite data, which means all of them are based in current data. The lack of historical sources of elevation data obligate us to use them. This simplification may be considered reasonable for rural places but it is more inconsistent in urban surroundings where the urbanization process

altered the original landscape. Even using DTM rasters, the construction of buildings and technical networks involved a severe change in the surface of the terrain. Several tests at a local scale were conducted with the different rasters in order to establish a balance between precision and operational time spend in the calculations. Total size of the files, time spend in different calculations and precision in relation to the finest data were some of the comparisons carried on. After these, we opted for SRTM90.

As stated in the text, the spatial units used as a basis for the present paper were civil parishes, comprising over 9000 continuous units. In this regard, we had to provide a method to obtain unique elevation variables for each unit, keeping the comparability across the country. We estimated six variables in total: elevation mean, elevation std, slope mean, slope std, ruggedness mean and ruggedness std. Before starting with the creation of the different variables, some work had to be done to prepare the data. In order to obtain fully coverage of England and Wales with SRTM data, we had to download 7 raster tiles. Those images were merged together, projected into the British National Grid and cut externally using the coastline in ArcGIS software.

Having the elevation raster of England and Wales, we proceed to calculate the first two variables: the elevation mean and its standard deviation. A python script was written to split the raster using the continuous units, to calculate the raster properties (mean and standard deviation) of all the cells in each sub-raster, and to aggregate the information obtained in a text file. These files were subsequently joined to the previous shapefile of civil parishes, offering the possibility to plot the results.

Appendix Figure 1: Slope and ruggedness measures



The second derivative of those results aimed to identify the variability of elevation between adjacent cells. In this regard, two methods were developed to measure this phenomenon: ruggedness and slope. Ruggedness is a measure of topographical heterogeneity defined by Riley et al (1999). In order to calculate the ruggedness index for each unit, a python script was written to convert each raster cell into a point keeping the elevation value, to select the adjacent values using a distance tool, to implement the stated equation to every single point, to spatially join the points to their spatial units and to calculate aggregated indicators (mean and standard deviation) per each continuous units.

In order to calculate the slope variable for each unit, a python script was written to convert the elevation into a slope raster, to split the raster using the continuous units, to calculate the raster properties (mean and standard deviation) of all the cells in each sub-raster, and to aggregate the information obtained in a text file. The obtained results for both ruggedness and slope are displayed at the end of this note. As the reader will appreciate, the scale of the indices is different (1 - 2 times) but the geographical pattern is rather similar. In this regard, we used for the paper those variables derived from slope measures because the time spend in calculations was rather lower.

Appendix A.3: Exposed coal

The shapefile of exposed coalfields of England and Wales c. 1830 was created by Max Satchell using the Digital Geological Map Data of Great Britain 1: 625,000 bedrock produced by the British Geological Survey (BGS). Exposed coalfields can be defined as those sections of coalfields where coal-bearing strata are not concealed by geologically younger rocks. They may, however, be overlain by natural (and man-made) sediments of the Quaternary period where they would form overburden in the exposed coalfield. Quaternary deposits are often unconsolidated sediments comprising mixtures of clay, silt, sand, gravel, cobbles and boulders. Exposed coalfields are of major historical importance because they were places where coal seams crop out at or near the ground surface making coal easiest to both discover and mine. For more details see

<https://www.campop.geog.cam.ac.uk/research/occupations/datasets/catalogues/documentation/exposedcoalfieldsenglandandwales1830.pdf>