Railways and growth: evidence from mid-nineteenth century England and Wales

Dan Bogart[*], Eduard Alvarez[†], Max Satchell[‡], Leigh Shaw-Taylor[§] and Xuesheng You[¶]

Draft, May. 2017[‖]

Abstract

How do transport improvements affect the growth and spatial structure of population and employment? This paper answers these questions in the context of nineteenth century England and Wales. It analyzes the effect of railways on population and employment growth following a large expansion in the railway network during the 1840s. Endogeneity is addressed using a hypothetical network minimizing construction costs between large towns in 1801. Results show that population, secondary, and tertiary employment growth were significantly higher near railway stations and agriculture was significantly lower. Moreover, larger growth effects are found in higher density localities and with coal, suggesting railways magnified agglomeration economies and natural resource advantages. A final counterfactual shows population and employment growth would have been 7 to 9% lower from 1851 to 1881 if the network remained at its smaller 1841 level.

Keywords: Economic growth, railways, endowments, spatial development

JEL Codes: N4

[*]Associate Professor, Department of Economics, UC Irvine, dbogart@uci.edu

[†]Visiting Scholar, Department of History, University of Cambridge, edu.alvarez.palau@gmail.com

[‡]Research Associate, Department of Geography, University of Cambridge max.satchell@geog.cam.ac.uk

[§]Senior Lecturer, Department of History, University of Cambridge lmws2@cam.ac.uk

[¶]Research Associate, Department of History, University of Cambridge max.satchell@geog.cam.ac.uk

[‖]Data for this paper was created thanks to grants from the Leverhulme Trust grant (RPG-2013-093) Transport and Urbanization c.1670-1911 and NSF (SES-1260699), Modelling the Transport Revolution and the Industrial Revolution in England. We thank Walker Hanlon for comments on earlier drafts and seminar participants at UC Irvine, UC San Diego, NYU, Florida State, Trinity College Dublin, Queens Belfast, and the Economic History Association Meetings. All errors are our own.

1 Introduction

Large scale transport improvements have the potential to drive economic growth and change the spatial distribution of economic activity. Much of the discussion of transport improvements in the literature focuses on four main questions. (1) How large are the effects on growth? (2) Do locations with better access to transport attract certain types of employment, like manufacturing or services? (3) Do the effects of transport improvements differ depending on population density or endowments? (4) Do transport improvements really create growth or mainly reorganize existing production across space?[1] This paper examines these questions for the case of railways in mid-nineteenth century England and Wales.

Railways marked a major improvement over inland road and water transport in the mid-nineteenth century. Evidence for England suggests railways generated relatively large savings through lower freight rates, lower passenger fares, and faster travel speeds. The savings are thought to be especially important in the transport of coal and the movement of passengers.[2] The effects on local population and employment growth are far from clear however. Simmons (1986), one of the leading historians of railways, argues that "the railway did not necessarily produce growth, in population or business. It might take people or business away (p. 16)." [3] The existing empirical work on English railways finds that locations gaining a railway station experienced population growth on average.[4] However, there are reasonable concerns about whether these findings are robust considering the endogenous placement of railways. Moreover, the types of employment that were attracted to railways is largely unknown and also whether railways really caused local growth rather than taking it from elsewhere.

[1]See Hettigate (2006) for an example of the policy debate and Redding and Turner (2014) for the academic debate.

[2]For literature on the social savings of railways in England and Wales see Hawke (1970) and Leunig (2006).

[3]For works on railways and population growth in England see Gregory and Marti Hennenberg (2010), Alvarez et. al. (2013), and Casson (2013).

[4]For works on railways and population growth in England see Gregory and Marti Hennenberg (2010) and Alvarez et. al. (2013).

We provide new evidence by analyzing population and employment data in more than 9000 quasi-parish units. They are the smallest geographical units in England and Wales, which we construct and have consistent boundaries from 1801 to 1881. Population totals are observed for parish units every 10 years from 1801 to 1881 and the same for all male occupations in 1851 and 1881. Parish male occupations are classified into five general groups: (1) secondary, (2) tertiary, (3) agriculture, (4) extraction or mining, and (5) general labourer. Additional specialties within secondary and tertiary are also studied to better understand manufacturing and service employment.

This paper also uses new spatial data on transportation networks and endowments.[5] The data include GIS shapefiles with highly accurate information on locations of railway lines and stations, ports, the turnpike road network, inland waterways, exposed coal fields, elevation, coastal boundaries, and soils. These data are linked to parish units using GIS, yielding the possibility to study the connection between parish population, employment, endowments, and transport networks at different points in time. Our main empirical model analyzes the effect of distance to railway stations in 1851 on parish unit population and employment growth from 1851 to 1881. The year 1851 is useful because the railway network underwent a major expansions in the 1840s known as the railway mania. In a short-time some parishes gained greater access to railways when their distance to stations declined. Other parishes remained distant from stations, and had to rely on road and waterway networks from the early nineteenth century.

Although the baseline model includes a rich set of controls, the distance to railway stations could be correlated with omitted variables related to growth. To address this issue we construct a hypothetical railway network connecting large towns in 1801. The routes are chosen to minimize elevation changes and distance known as the least cost path (LCP). The LCP predicts a parish's distance to the actual railway network in 1851 very well. It is also

[5]For more details see the project on Transport, urbanization and economic development in England and Wales c.1670-1911 at http://www.campop.geog.cam.ac.uk/research/projects/transport/

a good instrument because it identifies parishes that were close to railway stations mainly because they were near favorable routes for connecting large towns.[6]

Our main findings are summarized as follows. Greater distance to railway stations in 1851 deceased population growth from 1851 to 1881. The baseline ordinary least squares (OLS) estimate implies that a 50% increase in distance reduced annual population growth by 0.074 percentage points. The instrumental variable (IV) estimate implies a 0.164 percentage point reduction in annual growth. Interpreting the IV as causal, the OLS estimates appear to under-state the effects of railways distance. In a more flexible model, with varying effects according to railway distance bins, we find that parishes within 10km gained population compared to parishes more than 20 km distant, while parishes 10 to 20 km experienced no gains or lost population in comparison. The sub-population of parishes beyond 20 km distance from a station is useful analytically because these units were arguably 'untreated' by railways. An accounting exercise with these estimates suggests that population losses caused by railways were small in comparison to the population gains.

We also find the effects of railway differed across employment types. Standard land use models predict that land intensive occupations like agriculture should rise with distance to railways, and labor intensive occupations should decline with distance. Our baseline OLS specifications confirm this prediction. They show that increasing distance to railways by 50% reduced annual secondary and tertiary employment growth from 1851 to 1881 by 0.08 and 0.123 percentage points. The same change is predicted to increase annual agricultural employment growth by 0.021 percentage points. The effects of railways were even larger in rapidly growing secondary and tertiary sub-sectors, like machine tools, iron & steel, finance, commerce, and administration. The IV estimates again suggest that OLS model under-states the effects of railways on secondary, tertiary, and agricultural employment growth.

There are several extensions. One tests whether railways had different effects depending

[6] Our methodology draws on the so-called inconsequential place approach and other studies which least cost paths as instruments for infrastructure. See Chandra and Thompson (2000), Michaels (2008), Faber (2014), Lipscombe et. al. (2013), and Adler (2017).

on initial population or employment density. The largest effects are found in parishes in the fourth quintile of 1851 population or employment density (i.e. the 60th to 80th percentile). One potential explanation is that railways fostered knowledge spillovers, which in many models tend to drive greater growth in medium density areas (see Desment and Rossi Hansberg 2009). Another extension shows that the effects of railways on population and secondary employment growth was larger in parishes with coal. Thus railways also favored growth in locations with natural resource advantages.

The last section of the paper investigates the quantitative significance of railways to aggregate population and employment growth. One comparison is with the transport networks that preceded railways. We find that railways generally had larger effects on population and employment growth than turnpike roads, inland waterways, and general ports. The exception is ports with steamships services to foreign destinations. The effects of railways are also gauged through a counterfactual exercise, where we assume the railway network had remained at its 1841 level. The rationale is that railways were still an experimental technology in the 1830s, and had they proved non-economic, railway building might have stopped as early as 1841. In the counterfactual calculation, population and employment growth in each parish are predicted with an 1841 network and then added up to get the aggregate growth. The estimates suggest that aggregate population, secondary, and tertiary employment growth between 1851 and 1881 would be 7.6, 9.3, and 8 percentage points lower with the 1841 network.

The findings contribute to several literatures. The first addresses the spread of industrialization during the nineteenth century, which in the English context involved a growth in the level of secondary employment, rather than a rising share of the labour force. Leading explanations for industrialization center around market access and endowments like coal (Crafts and Mulatu 2006, Wrigley 2010, Fernihough and Hjortshøj O'Rourke 2014, Crafts and Wolf 2014), while others have emphasized education and finance (Becker, Hornung, and

Woessmann 2011, Heblich and Trew 2015). In the British context, most of these theories have been tested using regions as the unit or focused studies on the textile industry. This paper uses much smaller units and covers a wider range of economic activities.

This paper also adds to a large number of historical studies on railways. Some works focus on social savings and the direct benefits of lower transport costs (Fogel 1964, Fishlow 1965, Hawke 1970, Crafts 2004). Others analyze their effects on local population density and agricultural income.[7] Two related studies to ours, Crafts and Mulatu (2006) and Gutlberlet (2014), examine the effects of falling transport costs on regional employment structures. Crafts and Mulatu (2006) are especially notable as they argue that falling transport costs had small effects on the location of British industry from 1871 to 1911. Our findings are from an earlier date (1851 to 1881) and imply that railways had significant effects. This and our companion paper on turnpike roads and canals shows that railways had a larger growth effects than earlier transport innovations (Bogart et. al. 2017).

Our paper also contributes to the broader empirical literature analyzing transport improvements and development. In the concluding section, we provide a comparison between our estimates and those from modern settings.[8] We also add new evidence on mechanisms. Our finding that railways contributed to growth more in larger density parishes suggests transport improvements magnify knowledge spillovers, which are generally greater in more dense areas. The finding that railways contributed to more growth if parishes had coal suggests transport improvements also magnify natural resource advantages. Finally, our setting also provides a rare opportunity to study growth versus reorganization effects, which is a key issue in this literature.

[7]See Herranz-Loncán (2006) for Spain, Donaldson (2014) for India, Jedwab et. al. (2015) for Africa, Berger and Enflo (2015) for Sweden, Tang (2014) for Japan, Hornung (2015) for Prussia, Donaldson and Hornbeck (2016) for the US, and Gregory and Marti Hennenberg (2010), Casson (2013), and Alvarez et. al. (2013) for Britain.

[8]A selection of such studies includes Baum-Snow (2007), Duranton and Turner (2012), Banerjee, Duflo, and Qian (2012), Faber (2014), Garcia-López et. al. (2015), and Storeygard (2016).

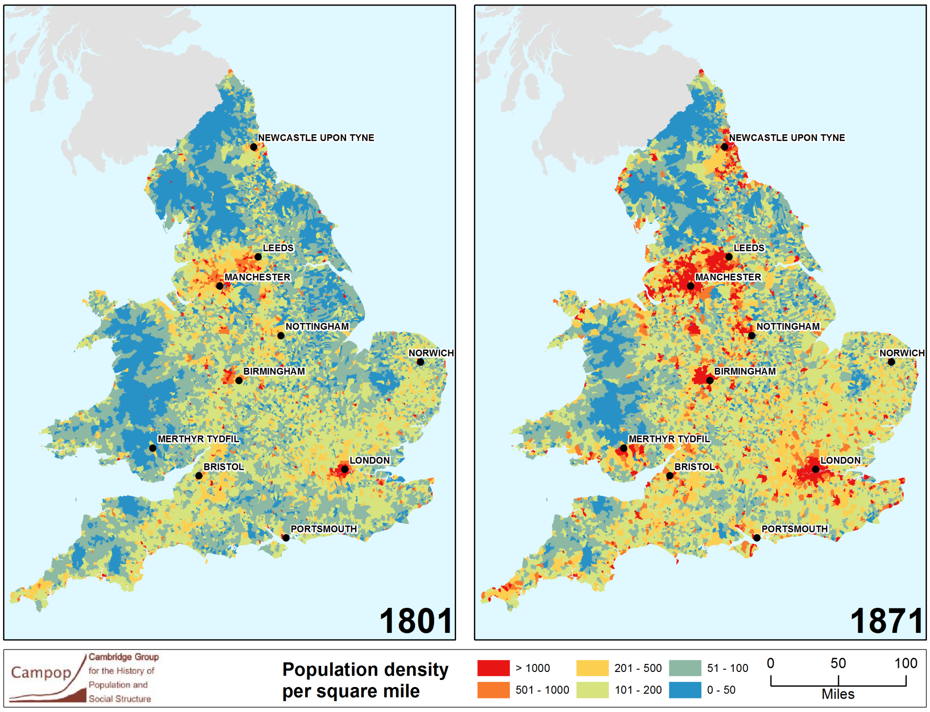
2 Background on population and employment

In the 1700s the English economy began experiencing regular population growth, but it was not until the 1800s that population growth truly accelerated. Census figures show that the English population increased from 8.6 million in 1801 to 17.0 million by 1851 and close to 22.3 million in 1881. The spatial distribution of the population changed substantially in the nineteenth century. The urban percentage of the population (people in towns of 5000 or more) rose from 29.5% in 1801 to 56.7% in 1871 (Shaw-Taylor and Wrigley 2014). London accounted for some of the urban growth with its percentage of the national population increasing from 11.2% in 1801 to 15.2% in 1871. The rest of urban growth mainly came from the northern industrial towns near Manchester, Leeds, Nottingham, and Birmingham, and in the coal mining districts near Newcastle and South Wales. Outside of the major towns, population generally declined. Individuals left the countryside for towns or emigrated to United States. Many urban migrants in England were positively selected (Long 2005).

As an illustration, the population density of the parish units we study below are shown in figure 1 for 1801 and 1871. Parishes shown in red have the highest levels of population density, blue the lowest. The growth of urban populations near London and the industrial towns is evident. So is the stagnant growth in some areas like north Wales, the southwest, the East Midlands, and the far northwest. The share of the total population living in the 100 largest parish units increased from 0.225 in 1801 to 0.323 in 1851 and 0.380 in 1881. Moreover, 22% of parish units experienced absolute population loss from 1801 to 1881, and between 1851 and 1881 the median parish lost population.

Employment also showed significant changes between 1801 and 1881. We focus on male employment in this and related papers because we have data for the entire 19th century. The aggregate data show that total male employment in England and Wales rose from 5.2 million in 1851 to 7.9 million in 1881. Male employment structure changed significantly. The three main occupational categories are primary, secondary, and tertiary. We divide primary

Figure 1: Population Density in England and Wales



Sources: see section 4 for details.

into agriculture and mining as they were quite different activities. Secondary refers to the transformation of the raw materials produced by the primary sector into other commodities, whether in a craft or a manufacturing setting. Tertiary encompasses all services including transport, shop-keeping, domestic service, and professional activities. Data from Shaw-Taylor and Wrigley (2014) shows that 36% of adult males in England and Wales worked primarily in agriculture in 1817, but in 1871 it was only 19%. The secondary sector absorbed a small share of the male labor force leaving agriculture. From 1817 to 1871 male secondary employment rose from 44% to 46% of the total. Male tertiary employment experienced the most change, increasing from 18% to 28% from 1817 to 1881. Mining or extractive employment also rose significantly from 3% to 6% in the same period.

The spatial trends in secondary employment match the trends in population. Already in the early nineteenth century, secondary employment was concentrated in the northern industrial towns and London. That concentration increased by the late 1800s, especially in the north. Tertiary employment density was low everywhere in the early 1800s, except for London. By 1881 tertiary was more common everywhere, but growth was a bit higher near the large northern towns. Agricultural employment densities were, of course, much lower than urban secondary and tertiary densities and declined in most locations.

2.1 Background on railways

England was a pioneer in railway technology and construction. Inventors like Richard Trevithick and George Stephenson developed steam locomotion in the early 1800s. Railway locomotives were far superior to coaches and wagons in both speed and cost. For example, rail freight rates in 1870 were one-tenth road freight rates in 1800 in real terms (Bogart 2014). Railways also had a competitive edge over barges on inland waterways, especially in speed. The only alternative transport sector that remained competitive with railways was coastal shipping. Steamships were invented in the early nineteenth century and transformed coastal and international shipping much like railways transformed land transportation (Armstrong 2009, Pascali 2016).

The first steam powered rail service open to the public was in the northern coal mining region between Stockton and Darling and opened in 1825. In 1830, the Liverpool and Manchester railway was opened to facilitate transportation of raw cotton and passenger traffic between the two large towns. It was promoted by local merchants and financiers who received authorization from parliament to build their line. Several other railways connecting nearby towns were promoted in the 1830s, but a national network had not yet formed.

The rail network substantially changed following what has been dubbed the 'railway mania' in the 1840s.[9] Several hundred railway companies were proposed by local groups and approved by parliament. Proposals called for nearly 15,000 km of railway track to be laid, but around 10,000 km were built in the following five years. The expansion of new routes was part of a strategy of the early railways to maintain their incumbent positions.

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

Figure 2: Growth in British rail miles and GDP

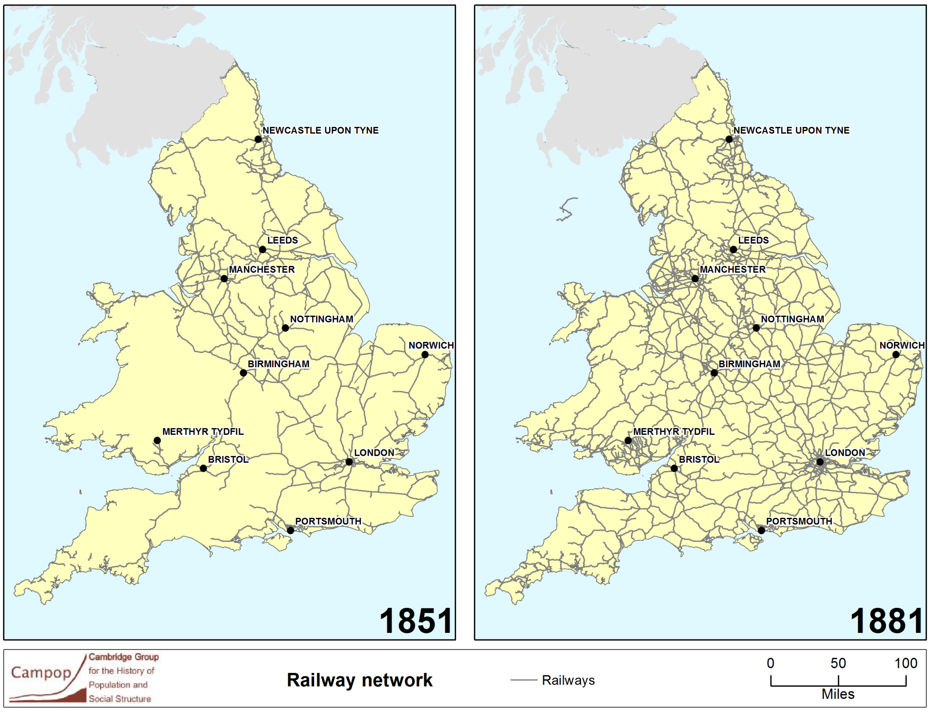
Sources: see text.

It was also driven by MPs who wanted to have railway stations in their constituency. The significance of the railway mania can be seen in figure 3, which plots the annual growth rate of British rail miles and British GDP. After an initial period of very rapid growth when the network was very small, rail mileage grew rapidly again in the mid-1840s. After 1850 railway mileage grew at much closer rate to GDP. The takeaway from this evidence is that the railway mania represented a 'shock' to the amount of railways in the English economy. It was also around the railway mania when some well known shippers and passengers switched to using railways.[10]

The railway mania is also significant because it produced the main trunk lines. The rail network in 1851 is shown in the left panel of figure 3. All of the major towns of England and Wales had a railway by this date, and most had a train service to London. However, there was still a diversity in railway access in 1851. In the parish units that we study below 50% were more than 6.9km distant from a railway station, 25% were more than 13.2km, and 10% were more 23.8km. We exploit this variation in our analysis below and in our main specification we estimate the effects of distance to railway stations in 1851 on parish-level

[10]For Maw (2015) for example documents that Pickfords, the well know shipping firm, switched from roads and canals to railways in the late 1840s.

Figure 3: Railway lines and stations in 1851 and 1881



Sources: see text.

population and employment growth between 1851 and 1881. Our approach does not capture railway building from 1851 to 1881, which as figure 3 shows continued further. We are less concerned with this issue because many of the areas with more railways in 1851 also had more railways in 1881. For the parish units we study below, the correlation between railway miles per square mile in 1851 and 1881 is 0.64.

Another advantage of our timing is that population growth from 1851 to 1881 cannot have directly caused railway building in the 1830s and 1840s, eliminating reverse causation. However, this does not imply that the expansion of railway stations by 1851 was unconnected with prior growth. In the parish units we study below population growth from 1831 to 1841 was 6% higher in parishes that increased the number of rail stations between 1841 and 1851 compared to those whose number of stations remained constant. In order to identify causal effects the endogenous placement of railways must be addressed.

3 Methodology and empirical strategy

Our empirical strategy is builds on previous studies analyzing urban growth and transportation (e.g. Baum-Snow 2007, Duranton and Turner 2012). Most works study the effects of network density or binary indicators of having infrastructure on population or employment growth. We differ in estimating the elasticity of population or employment growth with respect to distance to key infrastructures. The elasticity is grounded in a theory that transport infrastructure contributes to better consumer and producer access, which makes individuals and firms want to locate near these infrastructures (Redding and Turner 2014). The limited supply of land will act as a constraint on the growth. Thus, there should be a population gradient: areas closest to infrastructure should see the greatest growth and the areas furthest should see the least.

The baseline estimating equation is the following:

$$y_{i1881} - y_{i1851} = -\lambda y_{i1851} + \beta_1 r_{i1851} + \beta_2 x_i + \alpha_j + \varepsilon_{it} \tag{1}$$

where y_{it} is the natural of log population density in parish unit i in time period t, r_{i1851} is the log distance from parish unit i to its nearest railway station in 1851, x_i is a vector of geographic control variables including indicators for having coal, being coastal, and ruggedness. They serving as proxies for productivity and amenities. The controls also include variables for distance to the nearest large town in 1801 and distance to inland waterways, ports, and turnpike roads. α_j is a fixed effect for county j or registration district j. There are 59 counties in England and Wales and 616 registration districts. Our sample size of parish units is 9489 implying an average of 161 parishes per county and 15 parishes per registration district. The fixed effects control for external differences in market access (say being close to London versus Manchester) as well as other factors like unobserved productivity and amenities common among all units in a county or district. The standard errors are clustered on the fixed effect, for example the registration districts. The clustering addresses correlation in unobservables within districts.

Several alternative specifications to equation (1) are examined. One uses the average distance to the nearest and second nearest station rather than the nearest station only. It captures station density. A second replaces the natural log of station distance with dummy variables for distance to stations starting with 0-2 km, 2-4 km, and so on. A third includes a 3rd degree polynomial in the log of 1851 population to capture non-linearity in initial conditions. A fourth includes interactions between railways and quintiles for 1851 population or employment density and interactions with coal. A fifth includes a pre-trend variable for population growth from 1821 to 1851. This will partly address the issue that railways were placed in parishes that were growing prior.

Even with the wide range of controls there is still a concern that railways were placed in locations that for unobservable reasons were more or less likely to growth in the future. We address this issue using an instrumental variable (IV) for distance to railways stations. Our premise is that English railways were designed in large part to link larger towns that traded the most in the early nineteenth century. Some units along the route were close to railway stations simply because they were on the cost minimizing route connecting larger towns. In the literature on infrastructure, these are called inconsequential places.[11]

The first step in creating the instrument is to select the towns that will be connected by railways. We select towns with a population greater than 5000 in 1801. Recall that 1801 is before railways, so there is no direct connection. We use a simple gravity model (GM) equation to calculate the value of connecting English and Welsh towns with a population above 5000. The equation for town pairs i and j is $GM_{ij} = \frac{Pop_i Pop_j}{Dist_{ij}}$, where $Dist_{it}$ is the straight line distance between town i and j. We consider all town pair connections with $GM_{ij} > 10,000$.

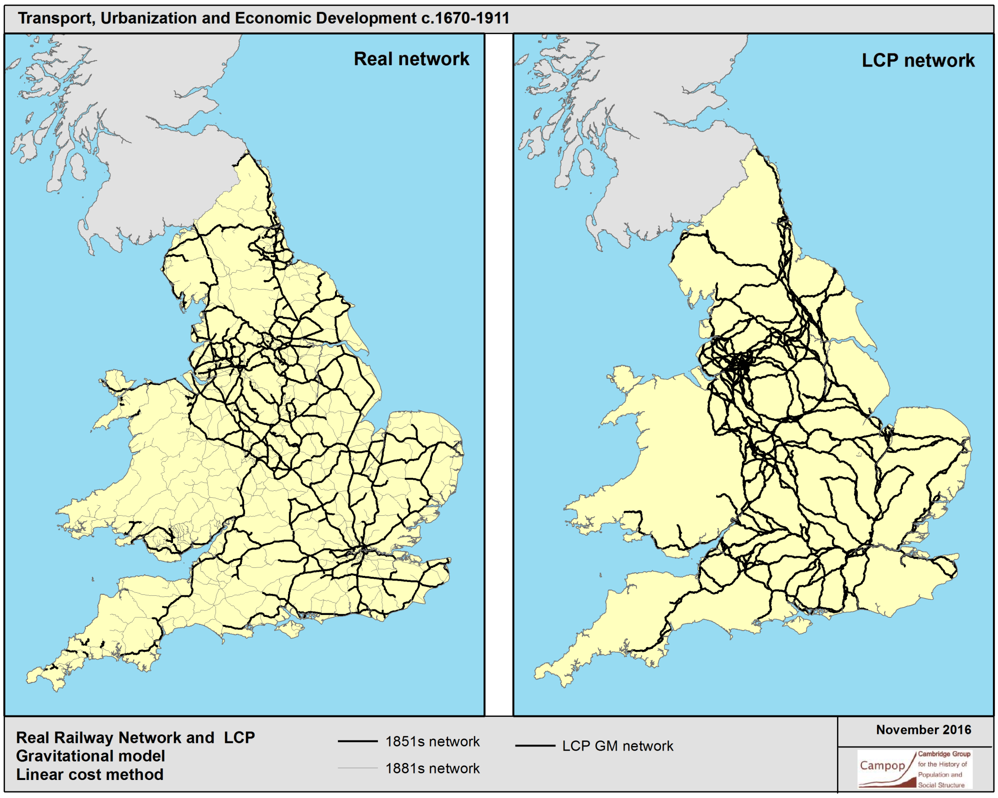
Next we identified a least cost path (LCP) connecting town pairs above the threshold.

[11] The inconsequential places approach has been used in other papers see Chandra and Thompson (2000), Michaels (2008), Faber (2014), Lipscombe et. al. (2013), and Adler (2017).

We assume that in considering their routes, railway companies tried to minimize the construction costs. The two main factors we consider are distance and building railways on sloping land. Terrains with higher slopes are those in which more earth-moving is required and, in consequence, their construction costs will be higher. We experimented with several approaches to modeling distance and slope. The details are given in appendix A. Our baseline model builds on a 19th century engineer Wellington (1877) who estimated the relationship between construction costs and elevation slope. Starting with a normalization of construction costs per km at zero slope to be 1, Wellington argued that construction costs per km increased by 2.96 for every 1% increase in slope. We use Wellington's formula to help identify the least cost path connecting our town pairs. Specifically, we used the ESRI least-cost-path python schema in order to run the spatial analysis using an elevation raster of England and Wales, which specifies elevation in 90 meter cells. The tool calculates a least cost path (LCP) from a destination point to a source. The end result is a network of hypothetical railway lines linking towns, which we call the LCP network. It is shown in the right hand panel of figure 4. The left hand panel of figure 4 shows the real railway network in 1851 in black and the lighter lines are the rail network in 1881. The overlap of the LCP and the rail network is very high, especially in 1851. Below we use the distance between a unit's location and the LCP as the instrument for distance to railway stations. Stations are so numerous by 1851 that distance to railway lines and distance to stations are similar.

Beyond specification and identification, there is a question of interpretation in this model. As articulated by Redding and Turner (2014), the differences in population between a unit close to infrastructure and further away from infrastructure captures a combination of pure growth and reorganization. Pure growth involves raising the population of units near infrastructure without lowering populations further away. Reorganization takes population from units further from infrastructure and moves them to units closer to infrastructure. There is no easy way to separate pure growth from reorganization, but our context offers an op-

Figure 4: Rail network and least cost path network



portunity. The uneven spread of the railway network in 1851 meant that a good portion of our parish units were very far from railway stations. Approximately 20% were more than 16 km from a station. The distance is significant because if an individual wanted to walk to the market in 1851 and the maximum they were willing to travel was 4 to 6 hours they could only live around 15 to 20km from the market. Thus it is plausible that units more than 15 to 20 km from a railway station were largely untreated by the railway, while units within 15 to 20 km were treated. Firms offering cheaper goods near stations could entice people from parishes within 15km to travel to their location, but beyond 15km the walking distance meant those individuals were not contestable. In this approach if units just below 15 km from a station experience a loss in population relative to units more than 15km there would be some evidence for reorganization.

3.1 Employment growth model

Aside from identifying the effects of population growth we are also interested in whether railways attracted certain types of employment. One hypothesis is that employment growth will decrease with railway station distance in labor intensive employment categories, and increase with distance in land intensive employment categories. This prediction is consistent with spatial models with two differentiated goods sectors, and with differing labor production intensities.[12] We test for the effects of distance to railway stations on the five employment categories: (1) secondary, (2) tertiary, (3) agriculture, (4) extractive (mining, fishing, and forestry) and (5) general labourer. Our baseline estimating equation is:

$$y_{i1881}^{k} - y_{i1851}^{k} = -\lambda_1^k y_{i1851}^{k} + \lambda_2 y_{i1851} + \beta_1^k r_{i1851} + \beta_2 x_i + \alpha_j + \varepsilon_i \qquad (2)$$

where y_{it}^{k} is the log of employment density in category k in period t and as before y_{it} is the log of population density in period t. Our main coefficient of interest is β_1^k which measures the effect of log distance to 1851 railway stations in employment category k . Note the two controls for 1851 density are likely to have different effects. The expectation is that population density in 1851 y_{i1851} will be positively associated with employment growth in category k from 1851 to 1881 if there are 'urbanization' economies. By contrast, own category employment density y_{i1851}^{k} is predicted to be negatively associated with employment growth in category k due to technology diffusion and congestion forces, which promote dispersion of employment. As with the population models, the employment growth equation (3) is

[12] Fujita et. al. (2001), Rossi-Hansberg (2005), Desmet and Rossi-Hansberg (2009, 2014) write down spatial models which explicitly deal with two differentiated goods sectors, like agriculture and manufacturing, or manufacturing and services. Most of these models consider locations along a line with the central segments corresponding to central locations, say in the middle of an island, and distant segments corresponding to frontier locations. The two sectors differ in land or labor intensity, and there are productivity spillovers from being located near more employment in the same sector. The spillovers generate employment clusters or areas of specialization. The labor intensive sector is more concentrated and dense in employment than the land intensive sector. Its concentration and density depends on the level of transport costs. In cases of sufficiently high transport costs overall production will be low and there were will be multiple clusters. In cases of sufficiently low transport costs, overall production will be high with fewer employment clusters.

extended to examine various specifications and distance to railway stations is instrumented with distance to the LCP.

4 Data

4.1 Population and employment data

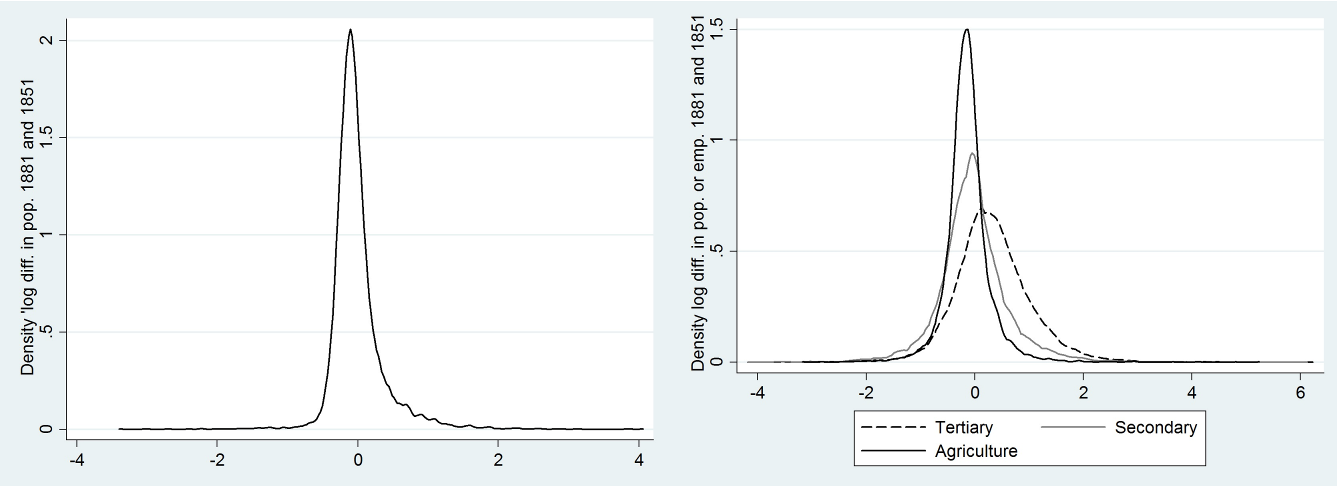
Population data from the census are available in every decade starting in 1801. They have been digitized at the parish level by the Cambridge Group for the History of Population and Social Structure (CamPop). Moreover, researchers at CamPop have created continuous parish units for the period between 1801 and 1891 and combined with them the census population data (Satchel et. al. 2016).

Detailed occupational data from the census are available at the parish level for all males in 1851 and 1881 through the Integrated Census Microdata (ICeM) project (Schürer and Higgs 2014). We focus on male occupations in part to link with earlier data. Male occupations are classified into 5 broad categories noted above using the primary, secondary, tertiary (PST) coding system.[13] Further classifications include 38 specific secondary occupations and 25 specific tertiary occupations. We identify 14 occupations with secondary and tertiary with above average growth for special analysis.

There is a challenge in merging the population and employment data in this paper. The continuous jurisdictional units for population from 1801 to 1891 are not always the same as the units recording occupations in 1851, 1881, and earlier sources. The sources report data sometimes in parishes, sometimes in townships within parishes, and sometimes in parishes that were later sub-divided. Using spatial matching techniques, we create a consistent set of boundaries for 9489 quasi-parish units that map population and employment data across the 19th century. For this paper we call them parish units for short. The details are provided

[13] The PST system is described in detail in Shaw Taylor et. al. (2014) and Wrigley (2015). Also see, http://www.campop.geog.cam.ac.uk/research/projects/occupations/categorisation/pst.pdf.

Figure 5: Distribution of population and employment growth across parish units 1881 and 1851



Sources: see text.

in the second appendix

The distribution of the log difference in population from 1851 to 1881 across units is plotted in the left hand side of figure 5. The median parish experienced negative population growth with a log difference of -0.073. Other parishes experienced significant population gains. At the top or 90th percentile the log difference is 0.377. Considering that English population increased from 17.0 million in 1851 to 22.3 million in 1881 (a log difference of 0.27) it is clear that population was concentrating in smaller numbers of units.

The distributions of the log difference in employment for 3 main categories are shown in the right hand panel of figure 5. Agricultural employment declined for most parishes, and the rate of decline was broadly similar across parishes as indicated by the tight distribution. Secondary employment also declined for most parishes, but there were some that experienced high secondary employment growth. The log difference in secondary employment is 0.69 at the 90th percentile, which implies a 100% growth in secondary employment. Tertiary employment grew in the median parish, but there was a wide dispersion. At the 90th percentile the log difference was 1.19 which implies a 228% growth.

4.2 Data on transport networks and parish unit centers

The paper uses new GIS data on transport networks in England and Wales.[14] Most importantly the data include the location of railway lines and stations with their opening and closing dates starting in the early nineteenth century.[15] Here we focus on the railway network of 1851 which is shown in figure 3 above. We also include data on distance to other transport infrastructure in the early to mid 19th century. These include all ports c.1840, ports with foreign steamship services c.1840, inland waterways c.1830, and turnpike roads c.1830. The list of all ports is comprehensive list of landing locations for ships given by *The Shipowner's and Shipmaster's Directory* from 1842. Steamships had become fairly common in most ports by the 1840s, but steamships destined for foreign ports were more limited. We draw on Langton and Morris (2002) who identify Hull, Liverpool, Bristol, London, and Southampton as the main ports with foreign steamship service. Turnpike road and inland waterway networks in 1830 are described in Rosevear et. al. (2017) and Satchell (2017).

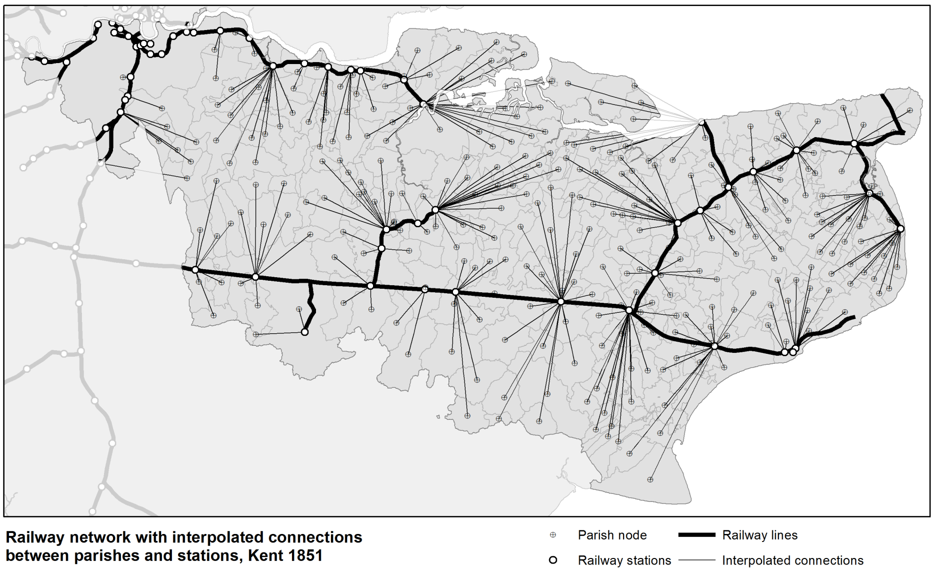
As discussed above, the empirical analysis studies the effects of distance between parishes and transport networks. For this purpose, it is necessary to define a parish unit center. Ideally the center would represent the main location of production and consumption in a parish. In many cases, the market place is ideal for this purpose. Therefore, if the parish had a market town at some point between 1600 and 1850 then the market town is taken as the center.[16] This applies to 746 of the 9489 units. Parishes that had no market town were likely to be rural and the geographic centroid is taken as the parish center. It should be noted that little error is introduced by using the centroid since the average parish unit was 15 square km. Figure 6 focuses on railways and stations in Kent to the southeast of London. It shows parish unit boundaries, centers (open circles), railway lines (black), and stations

[14]See http://www.campop.geog.cam.ac.uk/research/projects/transport/data/ for more details.

[15]See del Río, Martí-Henneberg, and Valentín (2008) for an initial description of the railways shape-file data. Additional upgrades were produced by the Cambridge group for the history of population and social structure (CamPop), see http://www.campop.geog.cam.ac.uk/research/projects/transport/data/railwaystationsandnetwork.html.

[16]See the following for a description of towns, http://www.campop.geog.cam.ac.uk/research/projects/transport/data/tow

Figure 6: Parish centers and distance to 1851 railway stations in Kent



Sources: see text.

(white circles). There is clearly much variation in the distance between parish centers and 1851 railway stations.

4.3 Data on endowments and summary statistics

The endowment data include shares of land in 11 soil types and indicators for being on exposed coal fields, and being on the coast. Soil categories are from Avery (1980) and Clayden and Hollis (1985) and were digitized by the National Soils Map of England and Wales.[17] The percentile in each category except one is included as a variable. Exposed coalfields are those where coal bearing strata are not concealed by rocks laid down during the Carboniferous Period.[18] Coastal units are identified using shapefiles for parish boundaries in England and Wales.

[17] The 11 soil categories are (1) Terrestrial raw, (2) Raw gley, (3) Lithomorphic, (4) Pelosols, (5) Brown, (6) Podzolic, (7) Surface-water gley, (8), Ground-water gley, (9) Man made, (10) peat soils, and (11) 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.

[18] The GIS does not capture a handful of tiny post carboniferous coal deposits, such as that at Cleveland (Yorkshire) which was worked in the 19th century. See Satchell and Shaw Taylor (2013) and http://www.campop.geog.cam.ac.uk/research/projects/transport/data/coal.html for more details.

The endowment data also include ruggedness measures like the average elevation, the average elevation slope in the parish, and the standard deviation in the elevation slope in the parish. The third appendix provides a description of these variables. A final control variable of note is the distance to the nearest large town in 1801 (Birmingham, Bristol, Leeds, Liverpool, London, Manchester, Newcastle, Plymouth, Portsmouth, Sheffield). It captures growth effects from having nearby urban centers.

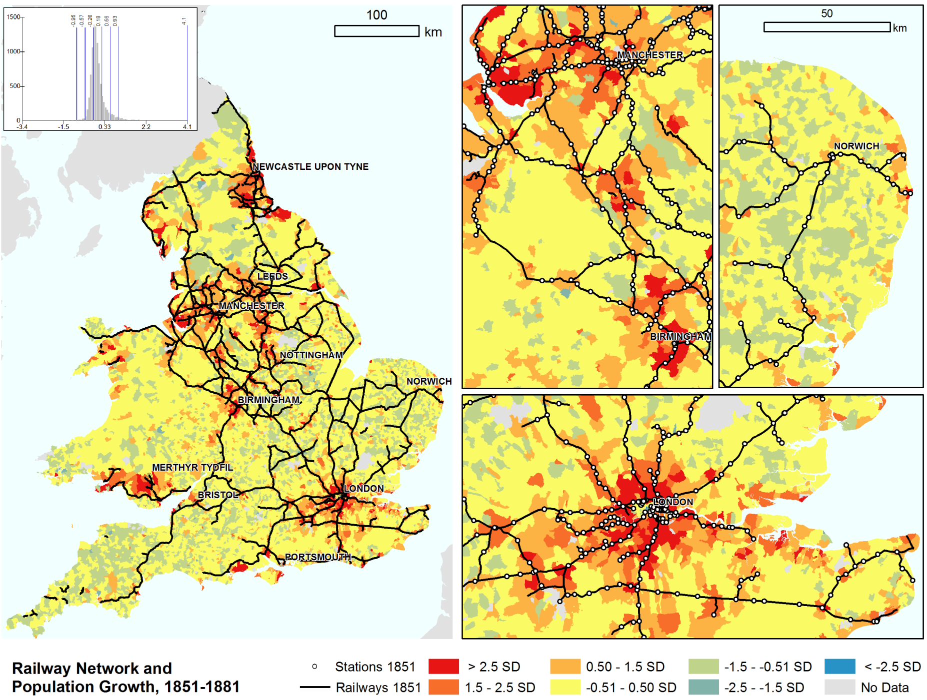
Table 1 reports summary statistics for the main variables. They are labeled as railways variables (1), population and occupation dependent variables (2), and the controls (3)-(6) described above. Note the 54 county and 616 registration district dummies are omitted.

5 Results

We begin with correlations and visual previews of our results. The correlation coefficient between the log distance to railway stations in 1851 and the log difference in population in 1881 and 1851 is -0.203 (p-value 0.00). Figure 7 maps this correlation. It shows the location of railway lines and stations in 1851 along with each parish unit's population growth from 1851 to 1881 measured in standard deviations. At a national level it is clear that areas which grew rapidly from 1851 to 1881 were close to railway stations. The same pattern is seen at a regional level near Manchester, Birmingham, and London. However, the East Anglia region in the upper right shows that having railway stations nearby did not guarantee high population growth. Our estimates suggest the effects in East Anglia were different in part because it had no coal and it had lower population density in 1851 than the Manchester and London regions.

Table 2 reports results for various regression specifications. Coefficients for geography and pre-railway infrastructure are not reported but will be described later. Column (1) is the most parsimonious and includes the log of 1851 population density, geography and pre-

Figure 7: Railways and local population growth from 1851 to 1881



Sources: see text.

railway transport infrastructure, and county fixed effects. It shows a significant negative effect of log distance to railway stations on population growth. The magnitudes will be discussed in detail below. Column (2) adds the log difference in population growth from 1821 to 1851 to address concerns about pre-trends. Population growth from 1821 to 1851 is found to be positively related to population growth from 1851 to 1881 as expected. Including this variable diminishes the effect of station distance, but it remains significant. The specification in column (3) replaces county fixed effects with registration fixed effects. The coefficient on station distance becomes more negative, implying that the narrower control for unobserved heterogeneity at the local level increases the magnitude of the station distance coefficient. Column (4) adds the higher order polynomial terms for log of 1851 population density. The estimated effect of station distance changes little. Column (5) reports the IV results for the same specification as in (4). The IV estimate is substantially larger than OLS. Columns (6) and (7) repeat the OLS and IV estimates dropping units with towns used in the construction of the LCP. Recall these were towns with a population greater than 5000 in 1801 and were predicted to get a railway connection to at least one large town by our gravitational model threshold. The IV estimates get larger in magnitude but the overall conclusion is the same.

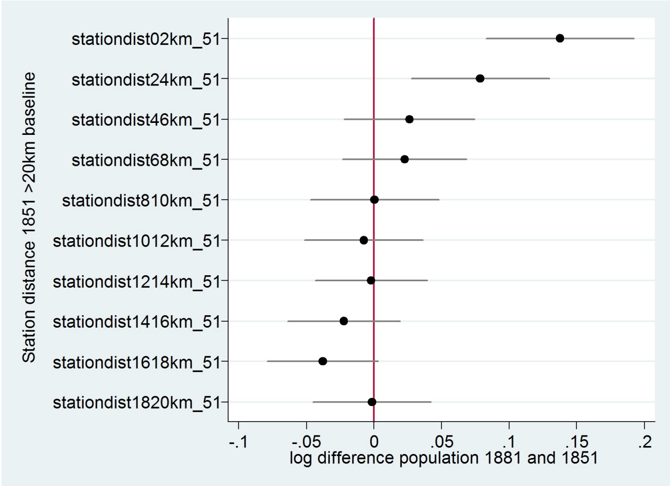
What do the coefficients imply? According to the OLS estimate from column (4) of table 2 a 10% increase in station distance reduces population growth from 1851 to 1881 by 0.54 percentage points. A 50% increase in distance reduces growth by 2.22 percentage points, or a 0.074% reduction in annual growth. According to the IV estimates a 50% increase in station distance reduces annual growth by 0.164% annually. For comparison the standard deviation for the annual population growth rate is 1.04%.

The larger magnitude of IV is not due to a weak instruments problem. The Kleibergen-Paap Wald F statistic indicates that the first stage is very strong. The higher IV coefficient is also consistent with a comparison of columns (2) and (3) in table 2, where the coefficient on railway distance is larger including registration district fixed effects than county fixed effects. This showed that controlling for more unobserved heterogeneity at the local level raises the coefficient.

There are several potential explanations for the downward bias in OLS. One is that the main objective of railway builders was to connect the large towns at the lowest cost. Therefore, railways were placed in some units because the cost of purchasing land was lower. This would imply that some parishes along the route and near stations were negatively selected. A related explanation is that the railway companies were pressured by politicians to put railway stations in their constituency, perhaps more so if the constituency had low growth potential. This would again imply a negative selection. Whatever the reasons, a similar finding of a downward bias in OLS will be found in other specifications below.

Railway access can be defined in other ways. One alternative is the average distance to the first and second nearest station in 1851. This approach addresses whether it was important to have more than one nearby station. Results using the log of the average

Figure 8: Effects on population growth for discrete distances to 1851 railway stations



Sources: see text.

distance to the first and second nearest station in 1851 are presented in table 3. They are qualitatively similar in all specifications, although the magnitudes differ.

Another alternative specification replaces log distance to stations with 10 distance bins ranging from 0-2 km to 18-20 km. The omitted group are parish units more than 20km from a railway station in 1851. The specification is otherwise identical to column (4) in table 2 and includes registration district FEs. The coefficients are plotted in figure 8 along with 95% confidence intervals. A striking pattern is found. Parishes within 0-2 km and 2-4 km had the largest population growth from 1851 to 1881, 13.7 and 7.9% percentage points higher than parishes more than 20 km. For parishes between 4 km and 20 km distant there is not strong evidence for any effect from railways. The one exception are the parishes between 16 and 18 km distant. Their population growth is estimated to be 3.8 percentage points lower than parishes more than 20 km from a station.

These estimates provides some insights on the magnitude of growth versus reorganization effects. If we assume that parish units more than 20km from a station were not treated, then we can interpret the effects in parishes 0-10 km distant as creating population growth and the effects in parishes 10-20 km distant as decreasing population growth through reorganization effects. We can also quantify the relative magnitude of each using the distribution of parishes in each distance bin, their population in 1851, and the coefficient estimates shown in figure 8. Railways increased population by 1.45 million in parishes between 0-10 km, or around 9.6%. Railways decreased population by 0.024 million in parishes between 10-20 km, or -1.2%. Thus, according to these calculations, the net effect of railways was to increase population form 1851 to 1881 by 1.43 million or 8.4%.

5.1 Employment growth results

The results for male employment growth are summarized in table 4. Column (1) analyzes the effects on the log difference in total male employment in 1881 and 1851 using OLS. The specification includes the log distance to the nearest 1851 station, a 3rd degree polynomial in log 1851 male employment, the log of 1851 population density, the log difference in population in 1851 and 1821, registration district fixed effects, geographic controls, and pre-railway infrastructure controls. The results show a negative and significant effect of station distance on male employment growth. Column (2) reports similar results using IV. The distance to station coefficient becomes more negative and remains significant, albeit with less precision. These two findings are nearly identical to the similar model for population growth (columns 4 and 5 in table 2), which is reassuring because male employment and population growth should respond similarly to station distance.

The advantage of studying employment is that we can test which types of occupations were attracted to parishes near railway stations. Column (3) in table 4 examines the effects on the log difference in secondary employment using OLS and column (4) does the same using IV. In both models a 3rd degree polynomial in 1851 secondary employment is included. The estimates are of similar magnitude to total male employment, although the standard error is larger in IV and makes the coefficient statistically insignificant.

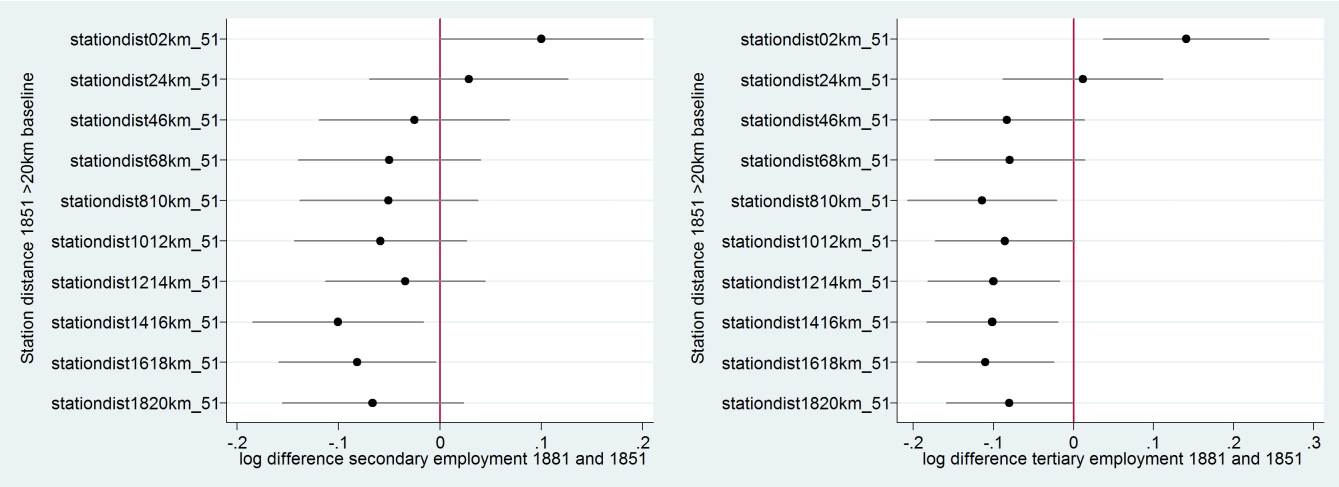
Columns (5) and (6) in table 4 repeat the same specifications for the log difference in tertiary employment. The coefficients are negative and significant in OLS and IV and are larger in magnitude than for total male employment. It appears that tertiary occupations were especially attracted to railway stations. Columns (7) and (8) show results for the log difference in agricultural employment. In OLS the coefficient on log station distance is positive and significant. In the IV the coefficient is larger although less precise. It is clear though that agricultural employment growth exhibits a very different relationship to railway access compared to secondary and tertiary employment growth.

The final two employment categories are extractive and labourer. Results for these are shown in columns (9) to (12). Labourer employment growth is broadly similar to secondary and tertiary although again the standard errors are larger in the IV. The effect of railways is less consistent on extractive employment growth with the sign of the IV coefficient changing. Extractive is arguably a unique sector because its location depends greatly on endowments like coal, and it was heavily dependent on transport. It is perhaps less straightforward to think of an inconsequential place getting a railway station and also having coal to mine.

Secondary and tertiary employment growth might have differed in close, medium and great distances from stations just at it did with population growth. Figure 9 plots the coefficients for both using the railway distance bins from 0-2 km, 2-4 km, etc. up to 18-20 km. The units more than 20 km are the omitted group. A similar striking pattern is found. Parishes within 4 km of railway stations grew significantly compared to parishes more than 20 km away. But parishes at medium distances, starting with 4 km up to 20 km generally had less secondary and tertiary employment growth than units more than 20 km. The negative growth effects for secondary are significant in two groups (14-16 and 16-18 km). In tertiary, the negative growth effects are present in many groups ranging from 8 to 18 km distance. The effects of railways on employment growth are clearly mixed across space, even though they are positive overall.

Secondary and tertiary are the most diverse of the 5 employment categories and so

Figure 9: Effects on secondary and tertiary growth for discrete distances to 1851 railway stations



Sources: see text.

it is useful to further examine their sub-categories. Rather than review all we chose to focus on secondary and tertiary employment categories that were growing more rapidly than the average from 1851 to 1881. Our list of rapidly growing secondary occupations includes printing, glass making, instrument making, chemicals, fuel, iron & steel, machine tools, electrical goods, gas equipment, and railway vehicles. Together these occupations represented 17% of secondary employment in 1851 and 26.5% of secondary employment in 1881. They accounted for 32% of all secondary employment growth from 1851 to 1881.

Table 5 reports the estimates for new and rapidly growing secondary categories. The specification is the same used to analyze the log difference in secondary employment in 1881 and 1851 (see table 4). The difference is that the controls include a 3rd degree polynomial in log 1851 own category employment, not a polynomial in log 1851 secondary employment. The results show that the log difference in employment is negatively related to station distance for 9 out of the 10 sub-categories. However, the coefficients are statistically significant in only two of the sub-categories, iron and steel and machine tools. One reason may be the small sample size for some of these new secondary categories. In column (11) we combine all males occupied in high growth secondary occupations into one sub-group. The effect of distance to stations has a significantly negative sign in the combined group. The same finding is reached using IV as shown in column (12).

Our list of rapidly growing tertiary occupations includes media, financial services, commercial and administrative services, and railway transport services. Together they represented 11.8% of tertiary employment in 1851 and 22.9% of tertiary employment in 1881, and account for 28.3% of all tertiary employment growth from 1851 to 1881. Table 6 reports the estimates for rapidly growth tertiary categories. There is a negative and significant effect on the growth of railway transport services, which makes sense given that railway stations required employees to operate. More notable is the negative and significant sign on employment growth in financial services, and in commerce and administration. Railways also contributed to non-rail tertiary employment growth.

Summarizing the results in this section, greater access to railways increased secondary employment growth. The OLS estimates, which are more precise than IV, suggest that increasing distance to railways by 50% reduces annual secondary employment growth by 0.08 percentage points. Railways had their clearest effects in rapidly growing secondary sectors, like iron and steel and machine tools. In the latter two increasing distance to railways by 50% reduces annual employment growth by 0.10 and 0.138 and percentage points.

Railways had larger effects on tertiary employment growth. The OLS and IV estimates are similar and imply that increasing distance to railways by 50% reduces annual tertiary employment growth by 0.123 to 0.162 percentage points. Some of this effect is due to the growth of railway employment near stations. But proximity also led to more growth in financial services and in commerce and administration. In the latter two the estimates imply that increasing distance to railways by 50% reduces annual employment growth by 0.077 and 0.123 percentage points. The last two service occupations are important because they contributed to growth in England and Wales well into the 20th century.

Finally, railways had very different effects on the spatial distribution of agricultural employment. In the OLS estimates, which are more precise than the IV, increasing distance to railways by 50% increases annual agricultural employment growth by 0.021 percentage points. We think the most likely explanation is that agriculture was not competitive in the land market near stations as railways attracted labor intensive industries. Many agricultural jobs were eliminated and those that remained shifted to parishes more distant from stations.

5.2 Heterogeneous effects based on initial density

In this section, we extend our baseline model to examine how the effects of station distance differ depending on a parish unit's 1851 population or employment levels. This extension is of interest for policy discussions as the distributional effects of transport improvements are often debated. It also has theoretical interest because railways may magnify the effects of knowledge spillovers which potentially increase growth for medium density areas. The following equation is estimated:

$$y^k_{i1881} - y^k_{i1851} = \sum_{j=2}^{5} \lambda_j \cdot 1851quintile^j_i + \beta_1 r_{i1851} + \sum_{j=2}^{5} \eta_j \cdot 1851quintile^j_i \cdot r_{i1851} + \beta_2 x_i + \alpha_j + \varepsilon_i \quad (3)$$

where the new terms are 4 dummy variables for 1851 population or employment density quintiles two, three, four, and five, and a series of interaction terms with distance to railway stations r_{i1851}. The elasticity of growth with respect to the distance for parishes in 1851 quintile $j = 2, 3, 4, 5$ is $\beta_1 + \eta_j$. The lowest density quintile 1 is omitted, and the effect of railways distance is β_1.

The results are summarized in table 7. Column (1) shows estimates for the population growth model. There is evidence for heterogeneous effects on population growth. The largest effect of station distance on population growth is for units in quintile 4 (i.e. the 60th to 80th percentile of 1851 population density). In this group increasing rail distance by 50%

lowers annual growth by 0.098 percentage points. The effect for quintile 1 is to reduce annual growth by 0.0547 percentage points. In quintiles 2, 3, and 5 the effect of railway distance is also greater but the coefficients are not as precisely estimated. There is no strong evidence for heterogeneity in the effects of station distance on secondary employment growth (see column 2, table 7). Railway distance had its largest effect in quintile 4, but the estimates are not precise. The same applies to other quintiles. There is more evidence for heterogeneity with respect to initial tertiary employment density (see column 3, table 7). The largest effects are found in quintile 4, but are also significant in quintiles 2 and 3. In quintile 4 increasing rail distance by 50% lowers annual tertiary employment growth by 0.098 percentage points. The effect for quintile 1 is to reduce annual growth by 0.060 percentage points.

Overall, railways seem to have their largest effects on population and employment in parishes with medium to large initial density, generally the 60th to the 80th percentile. These suggest that railways enhanced or reinforced knowledge spillovers, the main drivers of higher growth in medium to large density areas according to several models including Desmet and Rossi-Hansberg's (2014) spatial development model. The precise mechanisms require more research, but one theory is that railways facilitated information flows. A complementary argument is that units with the lowest initial employment density grow less if they were close to railway stations because they faced more competition with medium and large employment areas where productivity is higher.

5.3 Heterogeneous effects based on coal

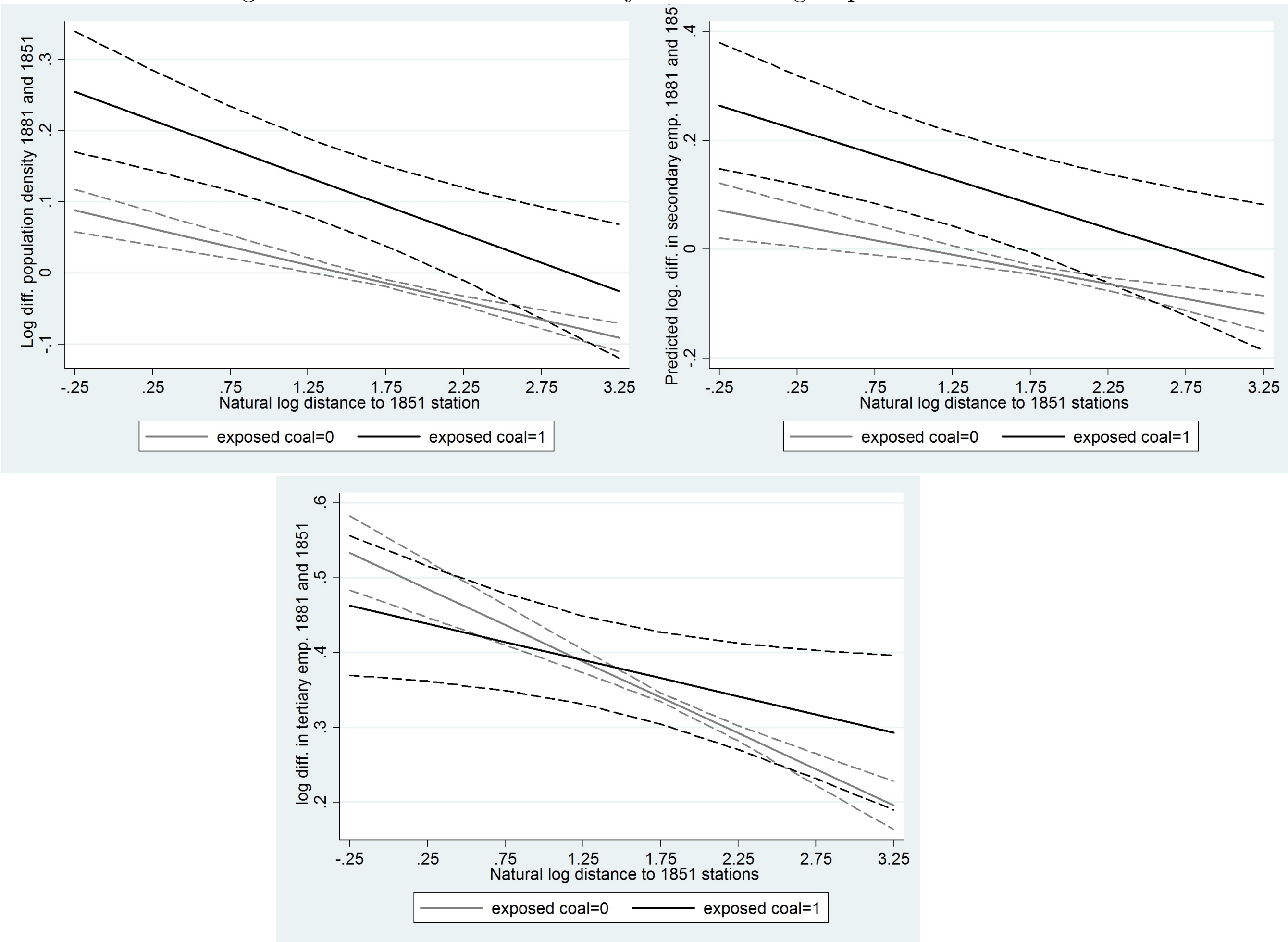
A second extension for heterogeneity considers whether railways had different effects depending on endowments of coal. Transportation of coal was a problem historically because it has a low value to weight ratio, and thus it did not pay to ship it over long distances (Wrigley 2010). Thus parish units with coal deposits had an energy cost advantage, and were often more likely to adopt early steam engines (Nuvolari et. al. 2011). It is an open question whether the advantages of having coal were enhanced by railways or other transport improvements. We examine this issue by studying interactions between log station distance and an indicator for parishes having exposed coal, which proxies for favorable coal endowments. The specification is otherwise the same as the baseline models (2) and (3). For simplicity, the interactions between station distance and 1851 population or employment density are omitted.

The results are shown graphically in figure 10 for population growth, and secondary and tertiary employment growth. They provide several insights. First, having coal endowments significantly increased population and secondary employment growth. On average population growth was 12 percentage points higher in parishes with coal, and secondary employment growth was 13 percentage points higher in parishes with coal. By contrast, tertiary employment growth was not any higher in parishes with coal. This makes sense because coal was not an important input in the production of services. Second, the estimates show some heterogeneity in the effects of railways and coal. For example, the effect of increasing station distance on population growth was one-half larger if parish units had a coal. The effects on secondary growth are two-thirds larger if a parish unit had coal. For tertiary there is no significant heterogeneity, but it is notable that the ordering of the slopes is different. The magnitude of the effect of railway station distance is smaller if a parish had coal.

5.4 Heterogeneous effects based on distance to nearest large town

As a final heterogeneity exercise, we created quintiles for distance to the nearest largest town. Recall that large towns are defined as the top 10 towns in 1801. In earlier specifications this distance was always included as a control variable. The quintiles for distance to a large town are now interacted with distance to railway stations as in the previous models. The analysis

Figure 10: The effects of railways and having exposed coal



Sources: see text.

was run for population growth, secondary employment growth, and tertiary employment growth from 1851 to 1881. In all cases, the interaction terms are not significant (results are available upon request). There are two implications. First, unlike what Faber (2016) finds for China in the 20th century, there was no secondary employment loss for units near large towns who were better connected with railways. Second, as we discussed earlier, one concern is that our variable for railways does not take into account the markets that were accessible by different railway stations. Presumably units close to large towns that were also near railway stations had access to larger markets, but this effect does not show up in our results.

5.5 Comparisons with earlier infrastructure

In this section, we compare the estimated effects of railways with the transport networks preceding them in history. With the exception of coastal shipping, railways were significantly more cost efficient than earlier forms of transport on roads and inland waterways. However, earlier transport networks, like turnpike roads, inland waterways, and ports, may have had persistent effects into the nineteenth century, even after their transport modes became outdated. Thus it is not obvious that railways had a larger effect on mid-nineteenth century growth than other transport networks. Also interesting is the comparison with ports serving steamships, as they were a recent innovation like railways and are estimated to have large effects on world trade (Pascali 2016). We measure the relative effects by calculating standardized coefficients for the distance to each transport network. For example, the standardized coefficient for the railways variable is its coefficient multiplied by the standard deviation of log distance to stations divided by the standard deviation in the log difference of population or employment. The comparison is informative because the distributions of the transport variable appear broadly similar, albeit with differences in means and variances.

Table 8 reports the results for specifications using the baseline OLS model from table 2. In general, log distance to railway stations has the first or second largest standardized coefficient and is the most consistent factor in explaining growth. Log distance to ports with steamships also has a large effect on growth, and emphasizes the importance of steam power for mid-nineteenth century growth. However, it is also surprising that distance to earlier transport networks have similar or larger effects in some cases. For example, distance to turnpike roads (in logs) is nearly as important as distance to railways in explaining secondary employment growth. Distance to inland waterways is nearly as important as distance to railways in explaining agricultural employment growth. Finally, the distance to general ports has a larger effect than railways in explaining extractive employment growth. It would appear that earlier forms of transport had persistent effects, perhaps by creating employment clusters which led to subsequent growth in the nineteenth century.

5.6 Counterfactual with 1841 network

As a final exercise we consider how growth would have been different if history unfolded differently. Railways were an experimental technology in the 1820s and 30s. It was only around 1840 that railways were proven to be cost effective substitutes for horse drawn wagons and canal boats (Dyos and Aldcroft 1969). Instead suppose that railways proved too costly, and no more railways were built after 1841. How would growth in England and Wales have been different from 1851 to 1881? We answer this question by comparing our model's predicted level of growth given the railway network of 1851 with our model's predicted level of growth assuming England and Wales kept its 1841 network.[19] In this case, we think it is appropriate to use a model that flexibly captures the effects of railways. Our model for predicting population growth comes from equation (4) below.

$$y^k_{i1881} - y^k_{i1851} = \sum_{j=2}^{5} \lambda_j \cdot 1851quintile^j_i + \beta_1 r_{i1851} + \sum_{j=2}^{5} \eta_j \cdot 1851quintile^j_i \cdot r_{i1851}$$

$$+\kappa coal_i + \tau r_{i1851} \cdot coal_i + \beta_2 x_i + \alpha_j + \varepsilon_i \tag{4}$$

The specification includes the standard controls, registration district fixed effects α_j, the 4 quintiles for 1851 population density $quartile^j_i$, and it allows for interactions between station distance and the 4 quintiles for 1851 population density and interactions between station distance and having exposed coal. The models for predicting employment growth is similar except it includes interactions with the quintiles for employment density.

[19]Specifically we calculate $\widehat{y_{i1881} - y_{i1851}}(rail1851)$ which is the predicted log difference in growth for each unit using the rail network of 1851. We then take exponential of the predicted growth which gives the predicted ratio for population or employment: $\widehat{\frac{Y_{i1881}}{Y_{i1851}}}$. We then multiply by the 1851 value $Y_{i1851}$to get each unit's predicted population or employment level in 1881 $\widehat{Y_{i1881}}$. Finally we sum over all units to the national predicted population or employment. The same calculation is done for$\widehat{y_{i1881} - y_{i1851}}(rail1841)$ which is the predicted log difference in growth for each unit using the rail network of 1841.

The results are summarized in table 9. With the 1851 network, population is predicted to grow by 44.3 percentage points, and with the 1841 network, population is predicted to growth by 36.3 percentage points. The difference in population growth is -7.6 percentage points. Turning to employment, we find that secondary employment is predicted to grow 9.3 percentage points less, tertiary employment 8.0 percentage points less, and agriculture 2.2 percentage points less with the 1841 network. Based on these estimates population and employment growth would have continued if railway technology proved non-economic in the 1830s. However, their growth would have been significantly lower, indicating that railways were a key driver of development in mid-nineteenth century England and Wales.[20]

6 Conclusion

How do transport improvements affect the growth and spatial structure of employment? This paper answers these questions in the context of nineteenth century England and Wales where urbanization increased and secondary and tertiary employment rose. Our empirical analysis studies the effects of greater distance to railway stations in 1851 on population and employment growth. To address endogeneity of railway placement, we use an instrument that identifies locations close to railways because they were on a least cost path minimizing elevation changes and distance between large towns. The OLS estimates are generally smaller in magnitude than IV. In this conclusion we discuss the OLS noting that they understate the effects of railways.

The main estimates implies that the elasticity of population growth with respect to distance to railways is -0.055. The estimated elasticities of secondary, tertiary, and agri-

[20] Hawke (1970) estimated that without railways freight and passengers would have faced higher transport costs resulting in approximately a 6 percent loss in GDP by 1865. Our estimates are more in line with Leunig (2006) who revises Hawke's estimates for time and money savings to railway passengers and finds them to be around 7 percent of GDP by 1880. Note that we arrive at our estimate from a very different methodology by using parishes more distant from railway stations as our counter-factual for parishes close to stations.

cultural employment growth with respect to distance to railways are -0.058, -0.089, and 0.015. These estimates are based on a 30 year horizon and so in annual percentage growth terms the distance elasticity estimates are -0.189, -0.198, -0.31, and 0.049 for population, secondary, tertiary, and agricultural employment.

How do the effects of railways compare with transport improvements in other contexts? One comparison is with modern-day China. Banerjee, Duflo, and Qian (2012) estimate the elasticity of annualized GDP growth in Chinese localities with respect to distance to railways in China. For comparison, we convert their estimates into a growth effect after 30 years. Their estimates for population growth, secondary growth, tertiary growth, and primary growth are -0.06, -0.194, -0.036, and -0.009. We find a larger effect of railways on tertiary employment growth and a smaller effect on population and secondary employment growth. Perhaps one reason is that England had already started its structural transformation in secondary employment prior to railways, but it was still in the midst of its early transition to tertiary. In China the transition to secondary occurred at the same time as the large increase in railways.

Duranton and Turner (2012) study the effects of highways on population growth of U.S. cities from the 1980s to the 2000s. They find that a 1% increase in highway density near a city would raise population growth by 0.15% over the next twenty years. In annual terms, their estimated elasticity of population growth with respect to highway density is 0.701. Our elasticity estimate for distance to stations is not directly comparable, but the magnitudes appear larger in the US. Thus more modern transport improvements from the automobile era seem to be generating greater growth effects.

Our findings also contribute to a better understanding of the mechanisms by which transport improvements affect growth. The result that railways had the largest effect on growth in medium to large density areas suggests that railways enhanced knowledge spillovers, which are a key factor in many urban and trade models. There are also policy implications from the findings that medium to large density areas had the largest growth effects from railways. If the aim of policy makers is to enhance growth in areas with less population or employment, then transport investments may not be as effective.

References

1. Anon. The Shipowner's and Shipmaster's Directory to the Port Charges, All the Depths of Water ... at the Various Places for Loading and Discharging Vessels in Great Britain and Ireland, Together with Similar Information Respecting Many of the Principal Foreign Ports, Etc. J. Daniel & Company, 1843.

2. 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.

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

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

5. Banerjee, Abhijit, Esther Duflo, and Nancy Qian. On the road: Access to transportation infrastructure and economic growth in China. No. w17897. National Bureau of Economic Research, 2012.

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

7. Becker, Sascha O., Erik Hornung, and Ludger Woessmann. "Education and catch-up in the industrial revolution." American Economic Journal: Macroeconomics (2011): 92-126.

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

9. 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.

10. Bogart, Dan, Max Satchell, Eduard J Alvarez-Palau, Xuesheng You, Leigh Shaw Taylor, "Turnpikes, canals, and economic growth in England and Wales, 1800-1850," Working paper 2017.

11. 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.

12. 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.

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

14. 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.

15. 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.

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

17. 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.

18. Crafts, Nicholas. "Productivity growth in the industrial revolution: A new growth accounting perspective." The Journal of Economic History 64.02 (2004): 521-535.

19. Crafts, Nicholas, and Abay Mulatu. "How did the location of industry respond to falling transport costs in Britain before World War I?." The Journal of Economic History 66.03 (2006): 575-607.

20. Crafts, Nicholas, and Nikolaus Wolf. "The location of the UK cotton textiles industry in 1838: A quantitative analysis." The Journal of Economic History 74.04 (2014): 1103-1139.

21. Del Río, Eloy, Jordi Martí-Henneberg, and Antònia Valentín. "La Evolución de la red ferroviaria en el Reino Unido (1825-2000)." Treballs de La Societat Catalana de Geografia 65 (2008): 654-663.

22. Desmet, Klaus, and Esteban Rossi-Hansberg. "Spatial growth and industry age." Journal of Economic Theory 144.6 (2009): 2477-2502.

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

24. Donaldson, Dave. Railroads of the Raj: Estimating the impact of transportation infrastructure. No. w16487. National Bureau of Economic Research, 2010.

25. Donaldson, Dave, and Richard Hornbeck. Railroads and American Economic Growth: A" Market Access" Approach. No. w19213. National Bureau of Economic Research, 2013.

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

27. Dyos, Harold James, and Derek Howard Aldcroft. British transport: an economic survey from the seventeenth century to the twentieth. Leicester University, 1969.

28. Faber, Benjamin. "Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System." The Review of Economic Studies (2014): rdu010.

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

30. Fishlow, Albert. American Railroads and the Transformation of the Ante-bellum Economy. Vol. 127. Cambridge, MA: Harvard University Press, 1965.

31. Fogel, R. "Railways and American Economic Growth." Baltimore: Johns Hopkins Press.(1979):"Notes on the Social Saving Controversy," The Journal of Economic History 39.1 (1964): 1-54.

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

33. 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.

34. 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.

35. Gutberlet, Theresa. "Cheap Coal versus Market Access: The Role of Natural Resources and Demand in Germany's Industrialization." (2014).

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

37. Heblich, Stephan, and Alex Trew. "Banking and Industrialization." (2015).

38. Herranz-Loncán, Alfonso. "Railroad impact in backward economies: Spain, 1850–1913." The Journal of Economic History 66.04 (2006): 853-881.

39. Hettige, Hemamala. "When do rural roads benefit the poor and how?: an in-depth analysis based on case studies." (2006).

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

41. 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).

42. Kitson, P.M., Shaw-Taylor, L., Wrigley, E.A., Davies, R.S., Newton, G., and Satchell, A.E.M. 2012. The creation of a census of adult male employment for England and Wales for 1817. Cambridge Working Papers in Economic and Social History, 4. Available at: http://www.econsoc.hist.cam.ac.uk/working_papers.html

43. 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.

44. Langton, John, and Robert John Morris. Atlas of industrializing Britain, 1780-1914. Routledge, 2002.

45. 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.

46. Long, Jason. "Rural-urban migration and socioeconomic mobility in Victorian Britain." Journal of Economic History (2005): 1-35.

47. 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.

48. Nuvolari, Alessandro, Bart Verspagen, and Nick Von Tunzelmann. "The early diffusion of the steam engine in Britain, 1700–1800: a reappraisal." Cliometrica 5.3 (2011): 291-321.

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

50. Pascali, Luigi. "The wind of change: Maritime technology, trade and economic development." forthcoming, American Economic Review.

51. 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.

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

53. Redding, Stephen J., and Matthew A. Turner. Transportation costs and the spatial organization of economic activity. No. w20235. National Bureau of Economic Research, 2014.

54. 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.

55. Rosevear, A., Satchell, A.E.M., Bogart, D, and Sugden, K., Shaw Taylor, L., 'Turnpike roads of England and Wales,' 2017.

56. Satchell, A.E.M., "Navigable waterways and the economy of England and Wales 1600-1835," Working Paper 2017.

57. Satchell, A.E.M., Kitson, P.M.K., Newton, G.H., Shaw-Taylor, L., Wrigley E.A., 1851 England and Wales census parishes, townships and places (2016). http://www.geog.cam.ac.uk/resea

58. Satchell, A.E.M. and Shaw-Taylor, L., Exposed coalfields of England and Wales (2013).

59. 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.

60. 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.

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

62. Storeygard, Adam. "Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa." The Review of Economic Studies (2016): rdw020.

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

64. 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.

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

66. Wrigley, E. A. "The PST system of classifying occupations," Working paper 2015. http://www.geog.cam.ac.uk/research/projects/occupations/britain19c/papers/paper1.pdf

Table 1. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
(1) railway vars.					
Distance to 1851 rail station in km	9489	10.4564	11.0657	0.0215	73.1296
Log distance to 1851 rail station in km	9489	1.8401	1.0979	-3.8407	4.2922
Distance to GM LCP km (IV)	9489	11.8619	16.5488	0.0001	116.3862
(2) Population and occupation dependent vars.					
Ln diff Pop. Density 1851 to 1881	9489	-0.0087	0.3765	-3.4044	4.0585
Ln diff secondary emp. 1851 to 1881	9061	-0.03	0.6569	-4.1897	6.1203
Ln diff tertiary emp. 1851 to 1881	9321	0.3328	0.7035	-3.7612	6.2385
Ln diff agriculture emp. 1851 to 1881	9403	-0.1382	0.4228	-3.1781	5.2364
Ln diff extractive emp. 1851 to 1881	3385	0.2752	1.1165	-3.912	7.5374
Ln diff labourer emp. 1851 to 1881	8231	0.4024	1.1718	-3.8067	5.247
(3) 1851 population and occupation controls					
Ln pop. density 1851	9489	4.2425	1.3673	0.8088	11.6253
Ln secondary emp. 1851	9222	1.3039	1.7556	-3.2755	9.6566
Ln tertiary emp. 1851	9362	0.9765	1.7621	-3.4681	10.1004
Ln agriculture emp. 1851	9449	2.2543	0.7663	-3.1699	7.7996
Ln extractive emp. 1851	4358	-0.7515	1.9174	-4.8644	6.62
Ln labourer emp. 1851	8586	0.1948	1.7672	-3.7992	8.7426
(4) Pre-trend controls					
Ln diff Pop. Density 1851 to 1821	9489	0.1695	0.2690	-1.0492	4.7950
(5) Geographic controls					
Distance to nearest large city in 1801 km	9489	136.3901	67.9921	0	418.7408
Indicator exposed coal	9489	0.0802	0.2716	0	1
Indicator coastal unit	9489	0.1479	0.355	0	1
Elevation	9489	89.721	74.025	-1.243	524.38
average elevation slope within unit	9489	4.7675	3.6157	0.4849	37.4272
SD elevation slope within unit	9489	3.4324	2.7174	0	23.1755
Perc. of land with Raw gley soil	9489	0.0847	1.3279	0	76.4964
Perc. of land with Lithomorphic soil	9489	8.6151	19.8301	0	100
Perc. of land with Pelosols soil	9489	8.2038	20.6374	0	100
Perc. of land with Podzolic soil	9489	4.6249	14.3262	0	99.5655
Perc. of land with Surface-water gley soil	9489	24.6329	29.4604	0	100
Perc. of land with Ground-water gley soil	9489	10.1871	20.1177	0	100
Perc. of land with Man made soil	9489	0.36384	3.2621	0	94.9904
Perc. of land with Peat soil	9489	1.1875	5.2798	0	91.4403
Perc. of other soil	9489	0.5354	1.9668	0	65.1538
(6) Pre-railway transport infrastructure					
Distance to nearest inland waterway 1830 km	9489	7.2316	6.5016	0	48.3873
Distance to nearest steamship port 1840 km	9489	85.0676	44.058	0	267.7452
Distance to nearest general port km	9489	30.2513	22.9766	0.0592	99.7121
Distance to nearest turnpike road km	9489	1.2302	1.4749	0	15.3485

Sources: see text.

Table 2: Effect of distance to nearest railway station on log difference population density 1881 and 1851

	(1) OLS coeff. (p-value)	(2) OLS coeff. (p-value)	(3) OLS coeff. (p-value)	(4) OLS coeff. (p-value)	(5) IV coeff. (p-value)	(6) OLS coeff. (p-value)	(7) IV coeff. (p-value)
Log dist. to nearest 1851 rail station km	-0.0419	-0.0377	-0.0652	-0.0548	-0.1191	-0.0573	-0.1597
	(0.000)	(0.000)	(0.000)	(0.000)	(0.013)	(0.005)	(0.003)
Log pop. density 1851	0.0041	-0.0169	-0.0222	-0.9459	-0.8903	-0.9617	-0.8797
	(0.548)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)	(0.002)
ΔLog pop. density 1821 to 51		0.2778	0.2094	0.1923	0.1794	0.1868	0.1659
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log pop. density 1851* Log pop. density 1851*				0.1752	0.1641	0.1772	0.1609
				(0.000)	(0.000)	(0.002)	(0.002)
Log pop. density 1851* Log pop. density 1851* Log pop. density 1851*				-0.0099	-0.0094	-0.0099	-0.0091
				(0.000)	(0.000)	(0.002)	(0.002)
County fixed effects	Y	Y	N	N	N	N	N
Registration district fixed effects	N	N	Y	Y	Y	Y	Y
Controls for geography and pre-railway transport	Y	Y	Y	Y	Y	Y	Y
Include units with town nodes in LCP	Y	Y	Y	Y	Y	N	N
Kleibergen-Paap rk Wald F statistic					58.271		49.042
R-square	0.3008	0.3308	0.4550	0.4818	0.4724	0.4816	0.4336
N	9489	9489	9489	9489	9489	9390	9390

Notes: Standard errors are clustered on county in specifications (1)-(2) and on registration district in (3)-(7). Geographic controls include indicators for exposed coal, coastal, elevation, average elevation slope and standard deviation within parish, distance to nearest large city and share of soil types. Pre-railway transport includes distance to nearest inland waterway, port, steamship port, and turnpike road. The instrument is distance to the LCP connecting large towns in 1801. see text for more details on instrument.

Table 3: Effect of average distance to first and second nearest railway station on log difference population density 1881 and 1851

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	OLS	OLS	OLS	IV	OLS	IV
	coeff.	coeff.	coeff.	coeff.	coeff.	coeff.	coeff.
	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
Log av. dist. to first and second nearest 1851 rail station km	-0.0538	-0.0482	-0.0962	-0.0846	-0.1833	-0.0894	-0.2382
	(0.000)	(0.000)	(0.000)	(0.000)	(0.012)	(0.005)	(0.002)
Log pop. density 1851	0.0047	-0.0163	-0.0205	-0.9522	-0.9044	-0.9684	-0.9031
	(0.492)	(0.010)	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)
ΔLog pop. density 1821 to 51		0.2781	0.2088	0.1902	0.1750	0.1845	0.1612
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Log pop. density 1851*				0.1770	0.1680	0.1790	0.1669
Log pop. density 1851*				(0.000)	(0.000)	(0.000)	(0.002)
Log pop. density 1851*							
Log pop. density 1851*				-0.0100	-0.0096	-0.0100	-0.0095
Log pop. density 1851*				(0.000)	(0.000)	(0.000)	(0.002)
County fixed effects	Y	Y	N	N	N	N	N
Registration district fixed effects	N	N	Y	Y	Y	Y	Y
Controls for geography and pre-railway transport	Y	Y	Y	Y	Y	Y	Y
Include units with town nodes in LCP	Y	Y	Y	Y	Y	N	N
Kleibergen-Paap rk Wald F statistic					50.421		45.223
R-square	0.3002	0.3303	0.4545	0.4818	0.4724	0.4821	0.4602
N	9489	9489	9489	9489	9489	9390	9390

Notes: Standard errors are clustered on county in specifications (1)-(2) and on registration district in (3)-(7). Geographic controls include indicators for exposed coal, coastal, elevation, average elevation slope and standard deviation within parish, distance to nearest large city and share of soil types. Pre-railway transport includes distance to nearest inland waterway, port, steamship port, and turnpike road. The instrument is distance to the LCP connecting large towns in 1801. see text for more details on instrument.

Table 4: Effect of average distance to nearest railway station on log difference male employment density 1881 and 1851

	All		Sec.		Tert.	
	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	IV	OLS	IV	OLS	IV
	coeff.	coeff.	coeff.	coeff.	coeff.	coeff.
	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
Log dist. nearest 1851 rail station km	-0.0445	-0.0967	-0.0589	-0.0900	-0.0897	-0.1173
	(0.000)	(0.076)	(0.000)	(0.185)	(0.000)	(0.066)
3rd deg. poly. in log. 1851 own emp. den.	Y	Y	Y	Y	Y	Y
1851 pop. den. and pop. growth 21 to 51	Y	Y	Y	Y	Y	Y
Registration district FE	Y	Y	Y	Y	Y	Y
Geography and pre-rail transp. controls	Y	Y	Y	Y	Y	Y
Units with town nodes in LCP included	Y	Y	Y	Y	Y	Y
Kleibergen-Paap rk Wald F stat.		56.035		59.358		58.801
R-square	0.4386	0.4337	0.3983	0.3976	0.5199	0.5194
N	9488	9488	9061	9061	9321	9321
	Agric.		Extract		Labour	
	(7)	(8)	(9)	(10)	(11)	(12)
	OLS	IV	OLS	IV	OLS	IV
	coeff.	coeff.	coeff.	coeff.	coeff.	coeff.
	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
Log dist. nearest 1851 rail station km	0.0152	0.0513	-0.0456	0.1978	-0.0990	-0.0773
	(0.019)	(0.133)	(0.110)	(0.211)	(0.000)	(0.377)
3rd deg. poly. in log. 1851 emp.	Y	Y	Y	Y	Y	Y
1851 pop. den. and pop. growth 21 to 51	Y	Y	Y	Y	Y	Y
Registration district FE	Y	Y	Y	Y	Y	Y
Geo. and transport controls	Y	Y	Y	Y	Y	Y
Units with town nodes in LCP	Y	Y	Y	Y	Y	Y
Kleibergen-Paap rk Wald F stat.		62.721		40.420		66.264
R-square	0.5613	0.5589	0.4765	0.4578	0.5687	0.5686
N	9403	9403	3385	3385	8231	8231

Notes: Standard errors are clustered on registration district in all specifications. Geographic controls include indicators for exposed coal, coastal, elevation, average elevation slope and standard deviation within parish, distance to nearest large city and share of soil types. Pre-railway transport includes distance to nearest inland waterway, port, steamship port, and turnpike road. Own employment applies to male employment in columns (1) and (2), and to the employment category in all other columns. (The instrument is distance to the LCP connecting large towns in 1801. see text for more details.

Table 5: Effect of average distance to nearest railway station on log difference male employment density in rapidly growing secondary categories

	Printing	Glass making	Instrum. making	Chemical	Fuel	Iron & steel
	(1)	(2)	(3)	(4)	(5)	(6)
	coeff.	coeff.	coeff.	coeff.	coeff.	coeff.
	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
Log dist. nearest 1851 rail station km	-0.0373	-0.1523	-0.0374	-0.1298	-0.0982	-0.0727
	(0.422)	(0.502)	(0.238)	(0.165)	(0.158)	(0.000)
3rd deg. poly. in log. 1851 own emp. den.	Y	Y	Y	Y	Y	Y
1851 pop. den. and pop. growth 21 to 51	Y	Y	Y	Y	Y	Y
Registration District FE	Y	Y	Y	Y	Y	Y
Geography and pre-rail transp. controls	Y	Y	Y	Y	Y	Y
Units with town nodes in LCP included	Y	Y	Y	Y	Y	Y
R-square	0.7132	0.7502	0.6511	0.6905	0.7168	0.4460
N	1223	264	1424	594	820	7358
	Machine tool	Electrical goods	Gas equip.	Rail vehicle	All high gr. sec. OLS	All high gr. sec. IV
	(7)	(8)	(9)	(10)	(11)	(12)
	coeff.	coeff.	coeff.	coeff.	coeff.	coeff.
	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
Log dist. nearest 1851 rail station km	-0.1006	0.1945	-0.0233	-0.2171	-0.1039	-0.2081
	(0.001)	(0.457)	(0.901)	(0.302)	(0.000)	(0.021)
3rd deg. poly. in log. 1851 own emp.	Y	Y	Y	Y	Y	Y
1851 pop. den. and pop. growth 21 to 51	Y	Y	Y	Y	Y	Y
Registration district FE	Y	Y	Y	N	Y	Y
Geo. and transport controls	Y	Y	Y	Y	Y	Y
Units with town nodes in LCP	Y	Y	Y	Y	Y	Y
Kleibergen-Paap rk Wald F stat.						58.756
R-square	0.5451	0.8941	0.8487	0.3659	0.4510	0.4452
N	2617	243	369	99	7647	7647

Notes: Standard errors are clustered on registration district in all specifications. Geographic controls include indicators for exposed coal, coastal, elevation, average elevation slope and standard deviation within parish, distance to nearest large city and share of soil types. Pre-railway transport includes distance to nearest inland waterway, port, steamship port, and turnpike road. Own employment applies to male employment in columns (1) and (2), and to the employment category in all other columns. (The instrument is distance to the LCP connecting large towns in 1801. see text for more details.

Table 6: Effect of average distance to nearest railway station on log difference male employment density in high growth tertiary sectors

	Media	Financial services	Comm. and admin.	Railway transp.	All high gr. tert. OLS	All high gr. tert. IV
	(1)	(2)	(3)	(4)	(5)	(6)
	coeff.	coeff.	coeff.	coeff.	coeff.	coeff.
	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)	(p-value)
Log dist. nearest 1851 rail station km	-0.0337	-0.0565	-0.0893	-0.1510	-0.1781	-0.2980
	(0.688)	(0.046)	(0.000)	(0.000)	(0.000)	(0.040)
3rd deg. poly. in log. 1851 own emp.	Y	Y	Y	Y	Y	Y
1851 pop. den. and pop. growth 21 to 51	Y	Y	Y	Y	Y	Y
Registration district FE	Y	Y	Y	Y	Y	Y
Geo. and transport controls	Y	Y	Y	Y	Y	Y
Units with town nodes in LCP	Y	Y	Y	Y	Y	Y
Kleibergen-Paap rk Wald F stat.						37.358
R-square	0.7652	0.6108	0.5980	0.6070	0.4998	0.4949
N	737	1938	2566	2758	4498	4498

Notes: Standard errors are clustered on registration district in all specifications. Geographic controls include indicators for exposed coal, coastal, elevation, average elevation slope and standard deviation within parish, distance to nearest large city and share of soil types. Pre-railway transport includes distance to nearest inland waterway, port, steamship port, and turnpike road. Own employment applies to male employment in columns (1) and (2), and to the employment category in all other columns. (The instrument is distance to the LCP connecting large towns in 1801. see text for more details.

Table 7: Heterogeneous effects on growth between 1851 and 1881 depending on 1851 population and employment density

VARIABLES	Pop dens. (1) coeff. (p-value)	Secondary emp. (2) coeff. (p-value)	Tertiary emp. (3) coeff. (p-value)
log distance nearest 1851 rail station in km	-0.0402	-0.0467	-0.0441
	(0.000)	(0.121)	(0.100)
Quintile 2 for 1851 density (pop. or emp.)	-0.0960	-0.282	-0.436
	(0.000)	(0.000)	(0.000)
Quintile 3 for 1851 density (pop. or emp.)	-0.0920	-0.319	-0.639
	(0.001)	(0.000)	(0.000)
Quintile 4 for 1851 density (pop. or emp.)	-0.0404	-0.287	-0.729
	(0.207)	(0.000)	(0.000)
Quintile 5 for 1851 density (pop. or emp.)	0.00316	-0.283	-0.845
	(0.918)	(0.000)	(0.000)
Quintile 2*log distance nearest 1851 rail station in km	-0.00283	-0.00979	-0.0465
	(0.789)	(0.729)	(0.061)
Quintile 3*log distance nearest 1851 rail station in km	-0.0115	-0.0231	-0.0486
	(0.336)	(0.409)	(0.081)
Quintile 4*log distance nearest 1851 rail station in km	-0.0315	-0.0350	-0.0602
	(0.021)	(0.253)	(0.040)
Quintile 5*log distance nearest 1851 rail station in km	-0.00837	-0.00563	-0.00450
	(0.494)	(0.850)	(0.875)
Registration district fixed effects	Y	Y	Y
Controls for geography and pre-railway transport	Y	Y	Y
Observations	9,489	9,061	9,321
R-squared	0.467	0.320	0.426

Notes: Standard errors are clustered on registration districts. Geographic controls include indicators for exposed coal, coastal, elevation, average elevation slope and standard deviation within parish, distance to nearest large city and share of soil types. Pre-railway transport includes distance to nearest inland waterway, port, steamship port, and turnpike road.

Table 8: A comparison of the effect of different transport variables on population and employment growth 1851 to 1881: standardized coefficients

	1 Pop. growth Stand. coeff. (p-value)	2 Sec. growth Stand. coeff (p-value)	3 Tert. Growth Stand. coeff (p-value)	4 Agric. Growth Stand. coeff (p-value)	5 Extract. Growth Stand. coeff (p-value)
Log distance to nearest 1851 rail station in km	-0.1530	-0.0927	-0.0630	0.0242	-0.0551
	(0.000)	(0.000)	(0.000)	(0.066)	(0.050)
Log distance to nearest 1830 inland waterway in km	-0.0350	-0.0240	0.0017	0.0250	0.0132
	(0.008)	(0.091)	(0.080)	(0.002)	(0.616)
Log distance to nearest 1840 steamship port in km	-0.1760	-0.1056	0.0839	0.0450	-0.0264
	(0.003)	(0.074)	(0.679)	(0.313)	(0.767)
Log distance to nearest general port in km 1840	-0.0537	-0.0061	0.0116	0.0344	-0.0737
	(0.062)	(0.825)	(0.189)	(0.164)	(0.037)
Log distance to nearest turnpike in km 1830	-0.0724	-0.0849	-0.0176	0.0084	0.0081
	(0.000)	(0.000)	(0.000)	(0.042)	(0.653)
3rd deg. poly. in log. 1851 pop. or emp.	Y	Y	Y	Y	Y
1851 pop. den. and pop. growth 21 to 51	Y	Y	Y	Y	Y
Registration district FE	Y	Y	Y	Y	Y
Geo. and transport controls	Y	Y	Y	Y	Y
Units with town nodes in LCP	Y	Y	Y	Y	Y
R-square	0.4835	0.3998	0.5216	0.5630	0.4777
N	9474	9046	9306	9389	3374

Notes: for explanations of variables see tables 2-9.

Table 9: National Counter-factual assuming England &Wales kept rail network of 1841

Model predicted population growth 1851 to 1881 in % with 1851 network	44.0
Counter-factual population growth 1851 to 1881 in % with 1841 network	36.3
Change in population growth in %	-7.6
Model predicted secondary emp. growth 1851 to 1881 in % with 1851 network	59.0
Counter-factual secondary emp. growth 1851 to 1881 in % with 1841 network	49.7
Change in secondary employment growth in %	-9.3
Model predicted tertiary emp. growth 1851 to 1881 in % with 1851 network	85.6
Counter-factual tertiary emp. growth 1851 to 1881 in % with 1841 network	76.7
Change in tertiary employment growth in %	-8.0
Model predicted agricultural growth 1851 to 1881 in % with 1851 network	-10.0
Counter-factual agricultural growth 1851 to 1881 in % with 1841 network	-12.3
Change in agricultural employment growth in %	-2.2

Notes: for sources and details of calculations see text.

A Appendix on the least cost path instrument

The main criteria used to plan linear projects has always been 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-movings and, in consequence, construction costs become higher. According to the literature, increases in tenths of slope lead to great differences in the construction costs. For instance, according to Wellington (1877), changes in one percent of longitudinal gradient might imply to triple the total cost of the works. Other authors referred to the complexity of building railways in terrains with slopes over 2% (Pascual 1999, Poveda 2003). 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. Longitudinal and perpendicular slope were the more significant ones. However, the quality of the soil, the necessity of construction tunnels and bridges or the inference with preexistences (building and land dispossession) were also important. The aim of this methodological note is to explain the development of an instrument to predict the design of linear infrastructure minimizing the construction costs. Using elevation data and implementing a Least Cost Path algorithm, we created counterfactual railway networks regarding different hypothesis.

Several DEM rasters have been analysed in preliminary tests, but we finally chose the Shuttle Radar Topography Mission (SRTM) obtained 90 meter (3 arc-second). Although being a current raster dataset, 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 optimise 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 according to the real cost of construction obtained from secondary sources. This slope raster was reclassified using different hypothesis with the aim to define different scenarios. Scenario 1 represented a construction cost directly proportional to slope. Scenario 2 considered a ratio 1/3 between slope and construction costs according to Wellington (1877) estimates. Scenario 3 considered a graded increase in construction costs according to larger slope intervals. From 6% onward, it assumed the necessity to construct a tunnel, so costs remained constant.

Table A.1: Construction costs vs elevation slope (in relation to the cost of building in a flat area: slope = 0%)

slope %	scenario 1	scenario 2	scenario 3
0	0	1	1
0-1	1	4	1
1-2	2	7	1
2-3	3	10	4
3-4	4	13	7
4-5	5	16	11
5-6	6	19	15
6-7	7	22	19
7-8	8	25	19
8-9	9	28	19
9-10	10	31	19
>10	...	34	19

The LCP algorithm was implemented using ESRI python, using as initial variables the elevation slope raster, the reclassification table of construction costs, and the node origin-destination nodes.

For the election of origins and destinations, we selected all the towns with a population over 5,000 inhabitants in 1801 and applied a gravitational model: $GM_{ij} = \frac{Pop_i Pop_j}{Dist_{ij}^2}$, where GM_{ij} is the gravitational potential between town i and j, Pop_i is the 1801 population of town i, and $Dist_{it}$ is the straight line distance between i and j. The set of origins were the town pairs with a $GM_{ij} > 10,000$. All those relationships were included in the model.

Having defined these variables, we calculated the cost distance and the back-link rasters using the formulation below:

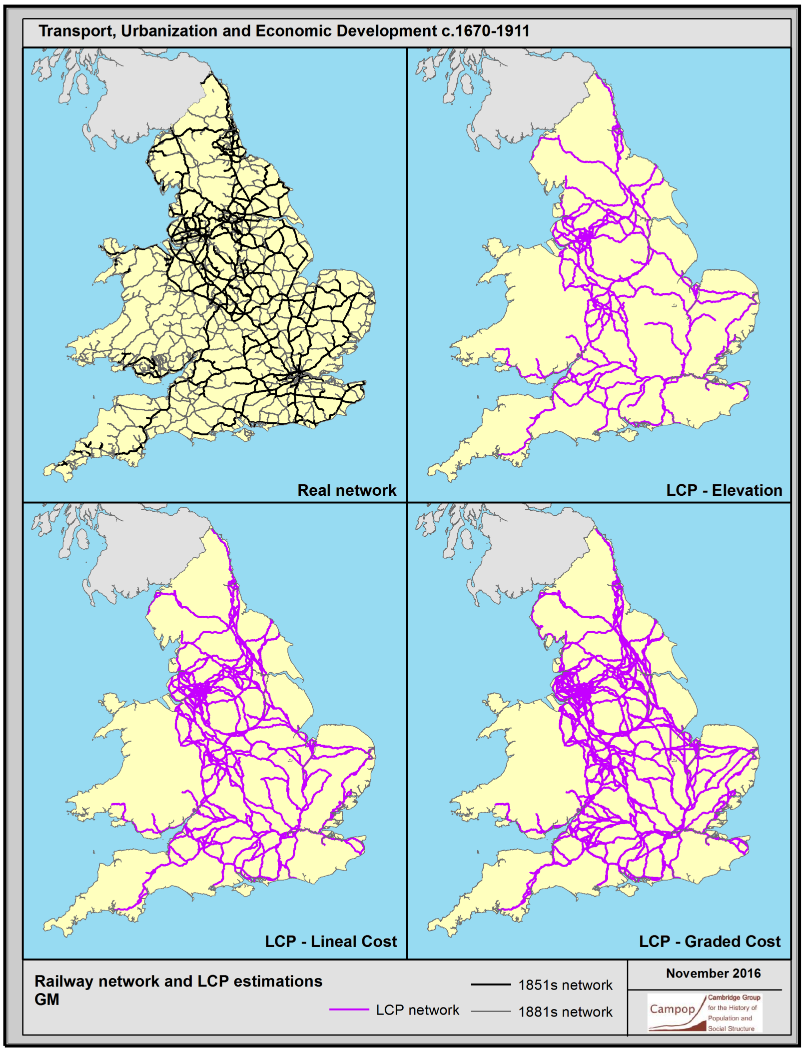
$GM_{ij} = (\frac{(CostSurface(a)*HF(a))+CostSurface(b)*HF(b))}{2})*SurfaceDistance(ab)*VF(ab)$, where $CostSurface(j)$ is the cost of travel for cell j, $HF(j)$ is the horizontal factor for cell j, $SurfaceDistance(ab)$ is the surface distance for a to b, and $VF(ab)$ is the vertical factor from a to b. Note that the division by 2 of the friction of the segments is deferred until the horizontal factor is integrated. Finally, we implemented the least-cost-path function to obtain the LCP corridors. These corridors were converted to lines, exported, merged and post-processed. Results of the methods are shown in the following figure.

A Appendix on creating mappable units

The English administrative units display highly inconsistent features. Several different hierarchical systems can coexist at the same time; different region can use different nomenclature; different systems can exist at different time slices; and boundaries of individual units within each system can change over time. Even though boundaries were never redrawn from scratch, different administrative system over time and boundary changes of individual units within any given systems over time mean that it would be difficult to carry out any analysis, either econometrically or cartographically, without having the data in a set of consistent geographical units.

This problem becomes even more apparent in the context of this paper and our larger project. This paper draws evidence from several datasets at different slices: the 1851 census

Figure 11: Rail network and least cost path network

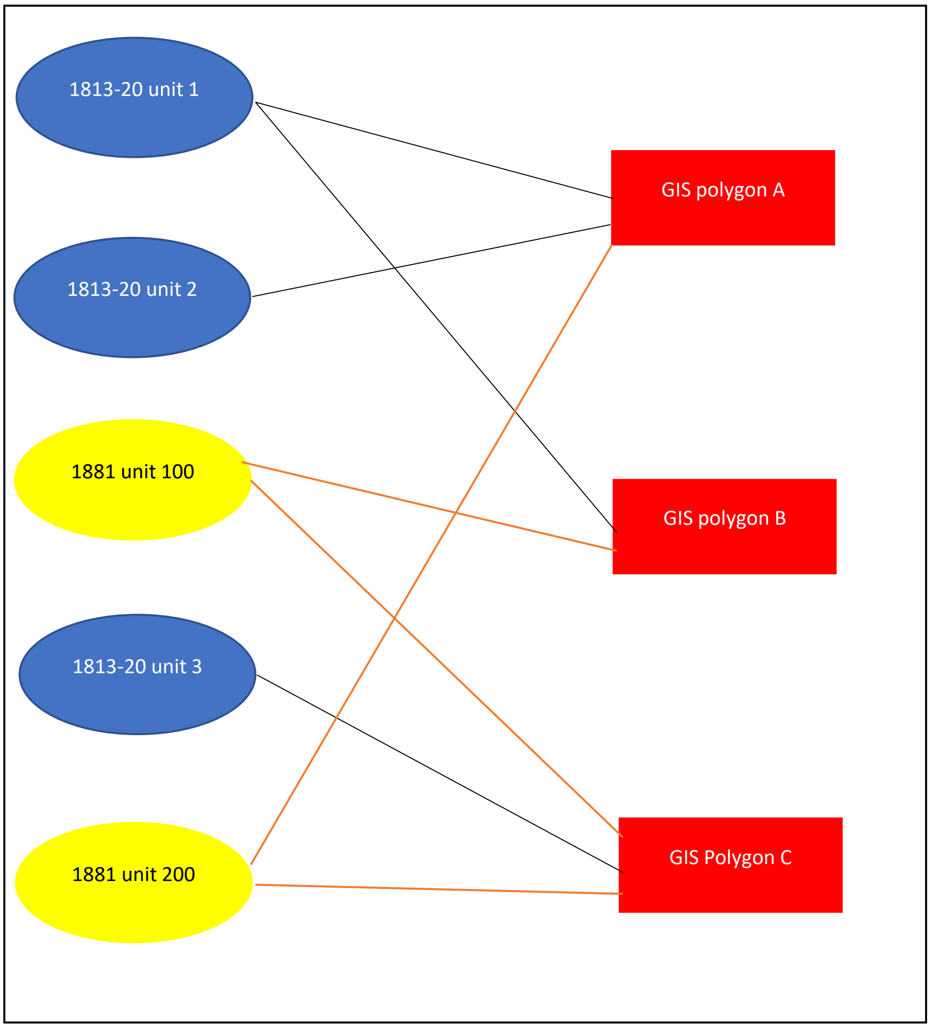


data, the 1881 census data, and the population data between 1801 and 1891. The larger project also draws on baptism data between 1813 and 1820. Each of these datasets have data at different geographical unit. The name and the number of geographical units in each dataset are presented in the table below.

	Name of the geographical unit	Number of the geographical unit
1813-20 Baptism data	Ancient parish	11,364
1851 census data	Civil parish	16,397
1881 census data	Civil parish	15,299
1801-91 population data	Continuous unit	12,750

The method of creating a set of consistent geographical units based on the units in each dataset involves two steps. Firstly, we made spatial match between parish level Geographical Information System (GIS) polygons and geographical unit from each dataset. The spatial match essentially made connections between the parish level GIS polygons and administrative units from each dataset through nominal linkage. The parish level GIS has c. 23,000 polygons. A separate note on the parish level GIS polygons can be found at Satchell et. al. (2016). Part of spatial match process can be carried out automatically, but there are cases where spatial matches can not be made automatically and require manual linkage. Ms Gill Newton and Dr Max Satchell, both of the Cambridge Group for the History of Population and Social Structure (Cambridge Group), University of Cambridge, managed the process of spatial matching based on an approach suggested by Dr Peter Kitson, previously of the Cambridge Group. A number of students from the University of Cambridge also provided research assistance during the process. A brief account of the spatial match process can be found in Kitson, P., et al, (2012). It should be noted that the nominal link between GIS polygons and administrative units from each dataset generated by the spatial match process can not be used directly for mapping purpose. This is due to the fact that a particular GIS polygon can be linked to more than one administrative units from each given dataset. But the spatial match process is essential for the second step we need to create a set of consistent geographical units over time.

The process presented above is the main function of Transitive Closure. When more datasets are added to the study, the situation becomes more complicated. But the basic idea remains the same. For example, imagine the following hypothetical situation:



If we are only dealing with 1813-20 baptism dataset, we can group polygons A and B together to form one mappable unit to represent units 1 and 2; and polygon C becomes a mappable unit on its own to represent unit 3. But once we add more datasets with different geographical units, in this case 1881 census data, we need to generate mappable units that are consistent across different datasets, i.e. over time as well. In this hypothetical case, Transitive Closure will group polygons A, B, and C together to form a single mappable unit. When dealing with 1813-20 baptism dataset, this mappable unit will draw data from units 1, 2 and 3. When dealing with 1881 census dataset, this mappable unit will draw data from units 100 and 200. In this way, the Transitive Closure process makes sure we are presenting and comparing observations from the same geographical units over time.

Transitive closure is a concept widely used in graph theory; for a formal definition and how to compute it, see for instance Cormen et. al. (2009). Ms Gill Newton, of the Cambridge Group, developed the Python code for Transitive Closure as part of the research project 'The occupational structure of Britain, 1379-1911' based at the Cambridge Group. Dr Xuesheng You, also of the Cambridge Group, implemented this code for this particular paper.

A Appendix on 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.

In order to carry on this work, we have downloaded 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 Dataset 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, although the involved contradictions. This simplification may be considered reasonable for rural places but it is more inconsistent in urban surroundings where the urbanisation 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 appendix on mappable units, 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.

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.

Slope was an alternative measure of topographical heterogeneity. 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.

Figure 12: Slope and ruggedness measures