Transport development and urban population change in the age of steam: A market access approach

Xuesheng You, Dan Bogart, Eduard Alvarez, Max Satchell, Leigh Shaw-Taylor

Abstract

This article takes a market access approach to study the effect of a major transport development, the introduction of steam-powered transport, on urban population change for 415 towns in England and Wales between 1830 and 1911. The totality of the roads, inland waterways, coastal routes and railway networks recorded in available sources for these two dates were digitised with unprecedented accuracy, in order to build a multimodal transport network accordingly. Our baseline model produces the unexpected result that, on average, improved market access had a negative effect on urban population growth. More specifically, the elasticity of population change with respect to market access change was of approximately -0.25. We argue that this unexpected effect was due to the strong heterogeneous effects of market access according to initial town size. Indeed, our estimates show that while market access had a significantly positive effect on population growth in large towns, it had a significantly negative effect on small towns. We argue that this heterogeneous effect was due to the fact that improved transport networks reinforced agglomeration effects in large urban centres while diminished the dispersion effect that had previously protected small towns.

1. Introduction

This article investigates the effect of transport developments on urban population change in England and Wales between 1830 and 1911. These two time slices were chosen as the basis for analysis and comparison in order to reflect the fundamental technological change occurring in transport over this period -- the adoption of steam power. Informed by recent developments in economic theory, our research takes a market access approach to the question, and applies it to a multimodal transport network model. Market access captures the aggregate rather than localized effects of transport improvements by taking into account changes in all locations within a network. Units of analysis consist of towns that had reached a population size of at least 2,500 by 1841. We calculate the lowest-cost town-to-town freight route for both time slices based on the transport network in existence at the time. This in turn serves as the basis for our computation of each town's market access. Population change is then explained in terms of changes in market access along with other factors such as natural endowments and initial population size.

In doing so, this article engages with a large body of literature on the impacts of transport, while distinguishing itself in three major aspects. First, our empirical analysis achieves an unprecedented level of accuracy. Over the past ten years, our team members have fully digitized multiple historical transport modes such as roads, waterways and coastal routes, with entirely novel accuracy. Unlike previous studies, this enables us to include *all* transport modes in operation at the time in our multimodal transport network, as well as to take into account costs both *along* the modes and transhipment costs *between* the modes. This better reflects reality and thereby makes our estimates much more robust. Second, this article studies the effect of transport on population growth in a unique historical context that is significantly different from that of other studies. By 1830, England and Wales already had highly developed and urbanized economies with an extensive transport network. While there have been numerous studies on transport in relatively backward economies, the extent to which steam-powered transport can further affect population change in a well-developed economy is still relatively unknown. This article bridges this gap. Third, and perhaps most importantly, the research methodology outlined in this article can be applied to much earlier time periods to study the effects of fundamental changes in transport over the long-run in a systematic and consistent fashion. Most previous studies analysed the impact of transport by comparing socioeconomic outcomes across locations, but within the same regime of transport technologies. For example, a frequently asked question in the literature is that of the difference in population growth between locations with a railway station and those without. Valuable as it is, this methodology does not have the capacity to examine the effects of fundamental changes in transport over the long-run. For example, what were the benefits of transport development in the age of steam power, relative to those in the pre-steam age? This article makes a first step towards answering this important question.[1]

Two key findings arise from our analysis. Firstly, and to the best of our knowledge, our study is currently the only one in the field to find that, on average, market access had a negative impact on urban population growth. We attribute this counterintuitive finding to the unique historical context of our study. Indeed, by 1830, a clear urban hierarchy with large urban agglomeration centres had already been established in England and Wales thanks to transport developments in the preceding century. Further, by 1911, the transport network had become so dense that it did not exhibit a clear spatial bias. These spatially homogeneous improvements in transport implied that, while existing large agglomeration centres could attract economic resources and population more easily, many small to medium towns were now opened up to new and higher levels of competition from the agglomeration centres. These small to medium towns may have lost their appeal as migration destinations, or lost population directly to large agglomeration centres as a result. Hence, on average, market access would have had a negative effect on population growth. Secondly, and relatedly, we also find that market access exhibits clearly heterogeneous effects on population growth depending on initial population size. While improved market access had a significantly positive effect on population growth for large towns, the opposite is true for small towns. This is consistent with the fact that the urban hierarchy was reinforced over this period, with the standard deviation of population distribution amongst the towns in our sample rising overtime.

The rest of this article will be organized as follows. Section 2 describes the historical background on urbanization and transport development in England and Wales. Section 3 reviews the literature on the effects of transport on population change. Section 4 describes the data. Empirical strategies are set out in Section 5. Section 6 discusses the results. Section 7 concludes.

2. Background

2.1 Characters of urbanization in England and Wales

Before the mid-eighteenth century, rates of population growth were broadly similar across Britain and other European countries. However, from 1750 onwards, and especially after 1800, England and Wales underwent a population expansion that eclipsed that of its European counterparts. Between 1800 and 1850, England had a rate of population growth of 13.6 per 1000 per annum. The corresponding figure for France was only 4.3 (Shaw-Taylor and Wrigley 2014). Over the same period, the combined populations of France, Germany, Italy, the Netherlands, Spain and Sweden increased by 34.2 percent, while the English and Welsh population alone almost doubled from 9 million to 17 million. Population in England and Wales continued expanding at a brisk rate thereafter, and by the time of 1911 census, its population size had reached more than 36 million.

Underlying this radical population expansion was an equally remarkable rate of urban growth. By the beginning of the nineteenth century, England had already become one of the most urbanized countries in Europe, second only to the Netherlands. It is estimated that nearly 30 percent of the English population were urban circa 1800 (Wrigley 1987). The urban growth rate remained rapid throughout the nineteenth century, such that, by the mid-nineteenth century, nearly half of the English population was living in urban settlements (Waller 1983). By the time of the 1911 census, more than 70 percent of the total population in England and Wales were living in urban settlements of more than 10,000 inhabitants (Law 1967). Furthermore, this sharp urban growth accounted for

[1] We have another WP examining the effect of market access on urban population change between 1680 and 1830.

almost the entirety of total population growth in the period: between 1851 and 1911, the population living in towns of more than 10,000 inhabitants increased by about 17 million, while the total population increase in England and Wales over the same period was just above 18 million (Law 1967).

Underneath this impressive level of aggregate urban expansion lie different levels of population growth for different urban settlements over time. In the seventeenth century, London clearly dominated the urbanization process at a time when nearly all of the urban population growth and a third of the total population growth in England were to be found in the metropolis (Wrigley 1987). However, thanks to industrialization, the nature of urbanization changed dramatically in eighteenth-century England and Wales. London and other historic regional centres such as Norwich, Exeter and York continued to grow, but now did so at a rate similar to that of the country as a whole (Shaw-Taylor and Wrigley 2014). Instead, it was the industrial cities, mining centres, and major ports like Manchester, Newcastle and Liverpool that experienced the most rapid population growth, moving up the urban hierarchy at the expense of historic regional centres. This urban hierarchy, with London at the top followed immediately by industrial, commercial and transport cities, had been well established by the beginning of the nineteenth century, and the rank order remained remarkably stable thereafter.

The introduction of steam transport did not alter this urban hierarchy. Tables 1 and 2 demonstrate the aforementioned patterns of urbanization over time. Table 1 lists the ten largest cities in England and Wales in 1680 and tracks their ranks in the urban hierarchy over time. It is clear that most of these historic regional centres had already lost their prominence in the urban hierarchy by 1830. By contrast, Table 2 lists the ten largest cities in England and Wales in 1911 and tracks their ranks in the urban hierarchy backwards in time. While most of these cities had not been major urban centres at the beginning of the period, they came to occupy the prominent positions in the urban hierarchy by 1830 and remained there down to 1911.

Table 1. Top 10 largest cities in England and Wales, 1680

	Population 1680	Population 1830	Population 1911	Rank 1680	Rank 1830	Rank 1911
London	310.9	1737.5	6512.9	1	1	1
Norwich	14.2	61.1	124.1	2	12	24
York	14.2	26.3	82.3	3	30	40
Bristol	13.5	113.5	379.8	4	6	8
Newcastle	11.6	73.7	424.0	5	9	7
Oxford	11.1	20.6	62.0	6	42	49
Cambridge	10.6	20.9	55.8	7	40	57
Exeter	10.3	32.5	59.1	8	26	52
Ipswich	9.7	20.5	73.9	9	43	44
Yarmouth	9.2	24.5	55.9	10	33	56

Table 2. Top 10 largest cities in England and Wales, 1911

	Population 1680	Population 1830	Population 1911	Rank 1680	Rank 1830	Rank 1911
London	310.9	1737.5	6512.9	1	1	1
Liverpool	1.2	210.4	1101.1	175	3	2
Manchester	2.4	257.8	1034.7	56	2	3
Birmingham	2.7	151.2	910.9	44	4	4
Sheffield	2.1	90.7	474.4	74	7	5
Leeds	3.5	119.1	454.2	33	5	6
Newcastle	11.6	73.7	424.0	5	9	7
Bristol	13.5	113.5	379.8	4	6	8
Bradford	0.9	45.0	288.5	233	16	9
Hull	6.6	53.7	278.0	16	14	10

2.2 Improvement on transport networks in England and Wales

This article measures transport costs and market access mainly through freight transport. Hence, this sub-section will focus on the freight market. England and Wales already had a well-developed transport network before 1830, with extensive turnpike roads, inland waterways and coastal routes (Bogart 2017, Bagwell 1974). The introduction of railways after 1830 further revolutionized the transport network and, as such, it is not surprising that the Victorian era is frequently referred to as 'the Age of the Railway'. From then on, steam power increasingly became the dominant source of energy in moving people and goods around.

Railway development was initially slow, and only 400 miles of railway lines were built between the opening of the Manchester Liverpool Railway in 1830 and the mid-1830s. However, thanks to intensive investment and construction over the next eight decades, particularly during the three 'manias' of 1837-40, 1845-47, and 1862-65, the total railway mileage eventually reached nearly 20,000 miles by 1911 (Cobb).

Railways possessed obvious advantages over existing modes of transport in terms of cost, speed and capacity. Before the age of railways, inland waterways such as navigable rivers and canals were the most cost effective inland freight carriers, especially for bulky goods such as grain and coal. As such, waterways proved essentials in the development of inland industrial centres by providing cheap fuel and raw materials before the age of railways (Crafts and Mulatu 2006, Crafts and Wolf 2014). However, the coming of railways provided an even cheaper means of freight transport. Though the data is patchy, it has been estimated that by the mid-1840s the canal freight rate was about 3*d* per ton mile, while the railway rate was just 1.7*d* per ton mile (Freeman and Aldcroft 1988). Further, the railway freight rate fell over time, such that, by the eve of the Great War, it stood at only around 1*d* per ton-mile (Freeman and Aldcroft 1988). As a result, railways quickly took over waterways in inland freight transport: whereas in the 1840s, canals still carried more tonnage than railways, by the early 1850s, railways had surpassed canals and, by the mid-1850s, carried about twice the tonnage of canals. Four decades later, railways carried almost ten times the tonnage carried by canals (Bagwell 1974).

The displacement of roads by railways for freight transport, at least for medium- to long-distance travel, was even more apparent. Road transport was expensive: to give but one example, in the eighteenth century, moving coal ten miles by road would typically double its pit-head price (Flinn 1984: 146). In fact, road transport of goods with low value-to-weight ratios was so prohibitively costly that there is no evidence to suggest that roads might have carried bulky goods other than for the shortest distances, and road transport was only economically viable for goods with high value-to-weight ratios. By contrast, it is estimated that, in 1870, the rail freight rate was only one-tenth that of road freight in 1800 in real terms (Bogart 2014). The benefits of railways over roads also extended to speed, security and regularity. As a result, railways quickly overtook medium- to long-distance road transport (Bagwell 1974, Bogart 2009).

It is misleading, however, to argue that railways led to a swift wipe out of road and waterway transport. The sharp decline of pre-existing inland transport modes is hardly disputable and, in the wake of railways, they ceased to be the principal arteries of inland transport they had been until then. However, both roads and waterways took on a new functionality, complementing railways at the local level. Most canals acted as local feeders of traffic to the main railway lines. Roads connected lesser towns and villages to railway stations, and facilitated intra-locality traffic dispersion (Bagwell 1974). The importance of this new functionality is perhaps most vividly described by Thompson -- 'without carriages and carts the railways would have been like stranded whales' (Thompson 1970: 13).

While railways increasingly dominated the inland freight traffic, they had a worthy competitor offshore. Coastal shipping, thanks to its much lower costs than inland modes of transport, was already the dominant means of transport for bulky, low-value goods such as coal, grains and iron ore before the age of steam (Armstrong 2009: 17-21). Robinson estimated that the total tonnage of coastal shipping more than doubled from 212,000 tons in 1686 to 505,000 tons in 1776, an increase of 1.5 percent per annum (Robinson 1967, 1988). Freight transport by coastal shipping was thereby essential to industrialization and urbanization before the age of steam. Take the North East to London coal trade as an example: between 1700 and 1830, London's coal imports showed a four-fold increase to more than two million tons to feed its expanding population and industries, and it was coastal shipping that carried almost the entirety of these coal imports from the North East coalfield (Flinn 1984, 216-21). It is fair to argue that, without coastal shipping, the expansion of the North East coalfield and that of London would have been nearly impossible during this period. After 1830, coastal shipping remained competitive thanks to significant technological improvements such as the use of steam to propel vessels (Armstrong 2009: 100). Taking the London coal trade as an example once again, it was not until 1868 that the volume of railway borne coal brought into London exceeded that of sea borne coal. Moreover, thanks to further improvements in coastal shipping and port handling, the situation had reversed again by 1898. The importance of coastal shipping in the age of railways becomes even more apparent when taking into account haulage. In 1910, railways carried 330 million tons of goods as opposed to 75 million tons by coastal shipping. However, railways had an average haulage of only 40 miles as opposed to 251 miles for coastal shipping. Therefore, in ton-mile, coastal shipping's share of freight transport was even greater than that of railways, at 18 billion for the former and 13 billion for the latter. (Armstrong 1987).

In summary, railways and coastal shipping dominated medium- to long-distance freight transport in the age of steam. However, waterways and roads did not entirely lose their relevance. They took on a new function in filling in the lacunae of the railway and coastal shipping networks and connecting local areas to the more efficient transport modes. It was therefore the extended connectivity, offered collectively by the extensive network of railways, waterways, roads and coastal routes, that characterized the transport system in England and Wales in the age of steam.

3. Literature summary: impact of transport on population change

3.1 Social savings approach

There exists a long cliometric tradition among economists and economic historians that consist in attempting to quantify the causal effects of transport networks on a number of socioeconomic indicators. Railways have typically been the centre of attention in such endeavours. The starting point of this tradition is Fogel's classic work on American railways. Taking the social savings approach, Fogel used counterfactuals to argue that the economic benefit of railways could have been achieved through lesser transport modes such as canals with relative ease. As such, he estimated railways' contribution to American GNP in 1890 to be small, of no more than 2.7 percent (Fogel 1964). Hawke later applied the same approach to investigate the effect of railways on the British economy. He arrived at a more positive account and estimated that the social savings from the railways contributed 6 to 11 percent to the national income of England and Wales in 1865 (Hawke 1970). Illuminating as it is however, later studies have frequently highlighted the theoretical and empirical limitations of the social savings approach (McClelland 1968, White 1976, Leunig 2010).

3.2 Estimation of localized effects

Since Fogel's and Hawke's works, a growing number of studies have applied econometric analysis to the question. Estimation methods have become more and more sophisticated over time, enabling challenging issues such as endogeneity, spatial correlation and path dependence to be addressed. However, the various studies and methodologies continued to share the same fundamental approach – measuring the treatment effect of railways on local socioeconomic change through comparison with locations without railways. As such, the literature has largely focused on the question of railways' impact on population distribution, industries, and trades, as well as on related aggregates.

Varying estimates of the effects of railways have come out of these studies, for different temporal and spatial contexts. Crafts and Mulatu (2006) find that railways only had a very weak effect on the location choices of industries in nineteenth-century Britain. By contrast, Tang (2014) finds a strong effect of railways on the redistribution of industries to more efficient locations in Meiji Japan. Berger and Enflo (2015) find a strong distributional effect of railways leading to population expansion in towns with a railway connection in nineteenth-century Sweden, at the cost of nearby locations without. Frach and Morillas-Torné *et al* (2013) argue that places without a railway connection in nineteenth-century Spain tended to lose population, although the effect on places with a railway connection remains unclear. Different estimates have been produced even for the same temporal and spatial contexts. Alvarez and Franch *et al* (2013) find that places with high levels of railway coverage in nineteenth-century Britain tended to undergo considerable population growth while places with low levels of coverage tended to lose population, whereas Schwartz and Gregory *et al* (2011) estimate that railways actually slowed rural depopulation in Britain. (For other references, see Atack and Bateman *et al* 2010 and Atack and Margo 2011 for America; Banerjee and Duflo *et al* 2012 and Baum-Snow and Brandt *et al* 2017 for China; Casson and Shaw-Taylor *et al* 2013 and Casson 2013 for comparisons between three English counties; and Galizia and Martí-Henneberg 2013 for a pan-European comparison).

Studies of this nature are invaluable in quantifying the effects of transport developments. However, many of them share two common drawbacks. First, they overwhelmingly focus on one particular mode of transport such as railways while, in reality, it was the transport network comprising all modes that mattered. For instance, railways may have had inconsequential effects on some locations precisely because the provision of other transport modes to complement the railways' functionality remained limited. Second, these studies only capture the localized effects of transport modes while, in reality, the impacts of transport were also determined by changes elsewhere in the network. For example, one location may have seen a significant extension of its transport network over time and yet only experienced limited socioeconomic benefit because transport development in other, competing locations were relatively greater.

3.3 Market access approach

In order to address these two concerns, some studies have shifted their focus to market access in the context of multimodal transport networks (See for example, Kotavaara and AntiKainen *et al* 2011, Koopmans and Rietveld *et al* 2012, Zhang and Nian *et al* 2016, Mimeur and Queyroi *et al* 2018, and Donaldson 2018). Multimodal networks integrate and connect different transport modes. Each transport mode constituting the network has its own transport costs, and transition from one mode to another often entails extra costs. Market access, in essence, captures the socioeconomic 'attractiveness' of a location to consumers and producers, taking into account changes elsewhere. Its measurement has two major components. First, each location's observable aggregates (such as population size). Second, the cost of transport, derived from the multimodal network, between each location and all others. Market access for each location therefore reflects a transport cost weighted sum over all other locations' observable aggregate.

Within this paradigm, the most significant contribution is perhaps Donaldson and Hornbeck's recent work on American railways (2016). Building upon a general equilibrium model (Eaton and Kortum 2012), they arrive at the theoretical prediction that trade flows, land values and population size respond positively to increases in market access. Their empirical analysis, using a reduced-form expression, is consistent with these predictions. Furthermore, they highlight a valuable property of the market access approach: all geographical units' market access will adjust to changes anywhere along the transport network. Therefore, both the direct and indirect effects of transport developments can be fully captured by analysing changes in market access. As such, market access can capture all of the economic forces that make goods and factor markets interdependent across geographical units, and regression models on market access are not prone to estimation bias (Donaldson and Hornbeck 2016: 826-7).

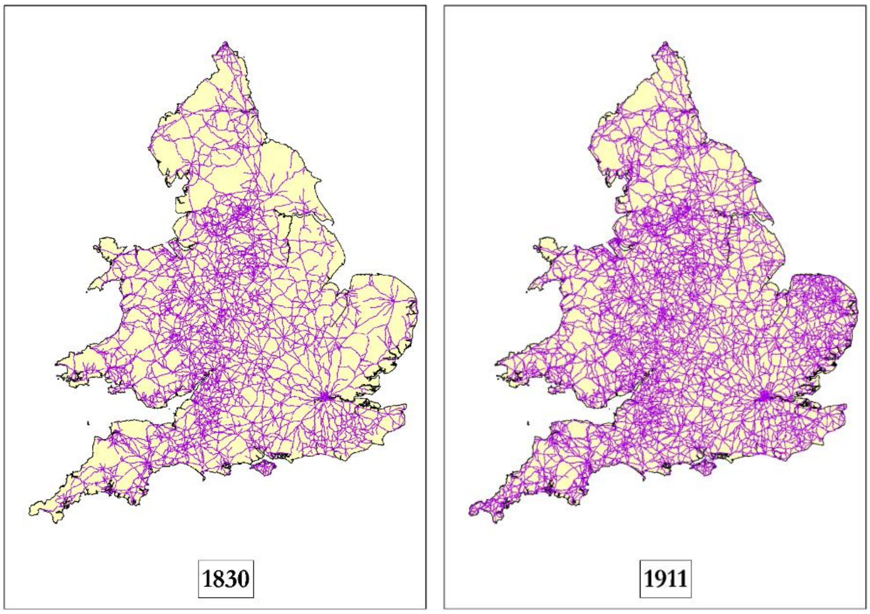
Given its desirable properties, this article adopts the market access approach to estimate the effects of transport on population change in an advanced economy, with an unprecedented level of accuracy. Before moving on to an outline of our empirical strategies, the next section describes the data sources and details how we constructed the multimodal network.

4. Data

4.1 Transport modes and digitization

Our multimodal network for 1830 consists of roads, waterways, coastal routes and ports. The transport modes constituting the 1911 multimodal network include roads, waterways, coastal routes, ports, railway lines and railway stations. The locations of each of the transport modes mentioned were identified from the most appropriate historical sources, and digitized in ArcGIS. To the best of our knowledge, this digitization and the subsequent network construction represent the most detailed and wide-ranging case study currently available in the field of multimodal network analysis.

Figure 1. road networks in England and Wales, 1830 and 1911



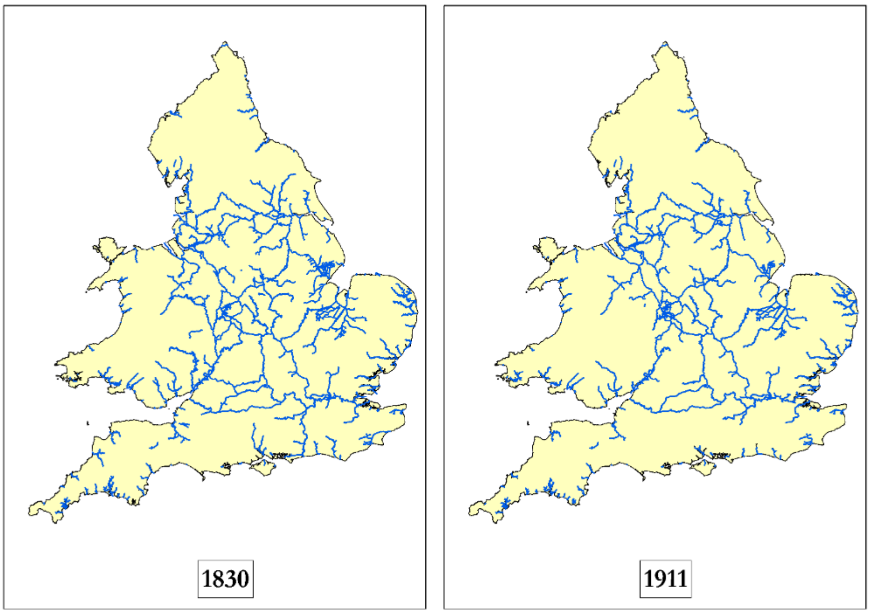
The 1830 road transport network is based on the turnpike network. Turnpike roads were toll roads built, maintained and operated by turnpike trusts authorized by Acts of Parliament. They built new roads or maintained existing roads by levying tolls on road users and issuing bonds mortgaged on the tolls (Bogart 2005, 2005, 2009). The 1830 turnpike network used in this article is based on our dynamic GIS database of turnpike roads covering the period 1667-1892 (Rosevear and Satchell *et al* 2017). This database reports the opening and closing dates of turnpike roads as well as the names of turnpike trusts whose authority each road fell into. When creating this GIS database, Max Satchell identified John Cary's *New Map of England and Wales and Part of Scotland* as the primary source behind an initial digitisation of the locations of turnpike roads. Further work based on sources such as parliamentary records, Acts of Parliament, milestones and tollhouses was undertaken by Satchell and Rosevear *et al* to improve the accuracy of the initial digitization.[2]

For inland waterways in both times slices, Max Satchell undertook ten years of painstaking labour to create a dynamic GIS database of rivers and canals between 1600 and 1948.[3] This database (Satchell, Newton and Shaw-Taylor 2017) utilizes a number of sources including, but not limited to, Richard Dean's *Inland Navigation: A Historical Waterways Map of England and Wales, the Ordnance Survey first edition* (surveyed 1840-1890)*, the Ordnance Survey Old Series* (surveyed 1789-c.1840), Hadfield's *The Canals of the British Isles Series* (11 vols.), T. S. William's *River Navigation in England 1600-1750,* the *Royal Commission on Canals and Waterways* (11 vols. 1906-1911), and H. de Salis' *Bradshaw's Canals and Navigable Rivers of England and Wales*. For this article, we selected the waterways operating in 1830 or 1911.

[2] The work to create this database was funded by grants from the NSF (SES-1260699) 'Modelling the Transport Revolution and the Industrial Revolution in England'; the Leverhulme Trust (RPG-2013-093) 'Transport and Urbanization c.1670-1911'; and the British Academy (SG121870) 'Riots and the Great Reform Act of 1832'. For more information, see Bogart, Rosevear, and Satchell 'Turnpike roads of England and Wales 1667-1892 GIS shapefile documentation'.

[3] The work to create this database was funded by grants from the ESRC (LCAG/080 RG43990) 'The Occupational Structure of 19th Century Britain'; the Leverhulme Trust (JJAG/078 RG51665) 'The Occupational Structure of England and Wales 1379-c.1729'; the Leverhulme Trust (RPG-2013-093) 'Transport, Urbanization and Economic Development in England c.1670-1911'. For more information, see Satchell, 'Navigable waterways of England and Wales 1600-1948 time dynamic GIS shapefile documentation'.

Figure 2. inland waterways in England and Wales, 1830 and 1911



With regards to the maritime connections in both time slices, we created a dynamic GIS database of ports and coastal routes covering the period 1650-1911 (Alvarez and Dunn et al 2017). We took a two-steps approach to create this database.[4] First, we identified and digitized the locations of ports from an extensive list of sources including Sacks and Lynch's *Ports 1540-1700*, Hargrave's *A Collection of Tracks Relative to the Law of England from Manuscripts*, Langton and Morris' *Atlas of Industrialising Britain 1780-1914*, Daniel's *The Shipowner's and Shipmaster's Directory to the Port Charges*, Steel's *Ship-Master Assistant and Owner's Manual*, and Hopwood's *Harbour Authorities*. Second, the coastal routes between ports were digitized based on the navigational knowledge of the era and on the physical geography of the coast in itself. Bathymetrical maps were used to determine the minimum distance the ships could navigate from the coast. For penetration routes from the cabotage line to the ports, we traced connections maximising depth and avoiding potential clogging in sandbanks (For more details, see Alvarez and Dunn, 2019).

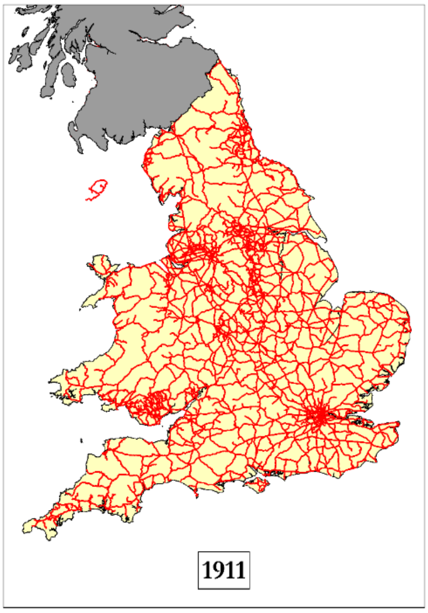
For railway transport in 1911, we used a dynamic GIS database of railway lines and stations covering the period 1825-1911 (Martí-Henneberg, Satchell and You *et al* 2017). This database maps the location of every railway line and station with their opening and closing dates. It is a verified and edited upgrade of the GIS mapping of railway lines and stations created by the team led by Jordi Martí-Henneberg at the University of Lleida (2006). The digitization done by Martí-Henneberg's team was based on the late Michael Cobb's definitive atlas *The Railways of Great Britain*. Our upgraded GIS increases the accuracy of Lleida's GIS. It added over 300 previously omitted stations and a few omitted lines, and corrected a significant number of digitization errors in the attribute data, particularly concerning the opening and closing dates.[5] For this article, we selected the railways lines and stations operating in 1911 from the dynamic GIS database.

---

[4] This work was funded by the Leverhulme Trust (RPG-2013-093) 'Transport, Urbanization and Economic Development in England and Wales 1670-1911'; the NSF (SES-1260699) 'Modelling the Transport Revolution and the Industrial Revolution in England'; and the Newton Trust 'Transport, Energy, and Urbanization c. 1670-1911'. For more information, see Bogart and Satchell, 'Ports of England and Wales, 1650-1911 GIS shapefile documentation'.

[5] For more information, see Satchell, 'England, Wales and Scotland Rail Lines 1807-1994 GIS shapefile documentation'.

Figure 3. Railways in England and Wales, 1911



4.2 *Towns and digitization*

This article focuses on the effect of transport on urban population change. We define urban locations using the population criteria. The population threshold by which 'urban' can be defined is unavoidably ambiguous (Smith, Bennett and Radicic 2018), and we therefore opt for an inclusive approach, choosing locations that had reached a population size of at least 2,500 either in the seventeenth century according to Langton's list of towns (2000), or in 1801 or 1841 according to Law (1967) and Robson (2006). In total, there are 415 locations satisfying this criterion and hence included in our analysis. For the purpose of this article, the benefits of taking this inclusive approach with a relative low threshold are clear. The mechanisms through which, and extent to which, transport development affected population change may well be significantly different for different types of urban locations. Our approach ensures that urban locations with a wide spectrum of characteristics, ranging from industrial centres to market towns, are all being considered. Therefore, the effects of transport development can be analysed and measured in a wide range of contexts. The locations of these 415 urban units were extracted from our GIS database of 1,746 candidate towns in England and Wales from c.1563 to 1911 (Satchell and Potter *et al* 2017). In this database, each candidate town is represented as a point. The point was digitized according to a hierarchy of central and enduring features of urban locations, such as market place, parish church, and inn.[6] The summary statistics of the 415 urban locations are presented in table 3. While, on average, the population size of the towns in our sample grew by c.2.4 times between 1830 and 1911, growth rates for different towns differed to a considerable degree.

Table 3. summary statistics of 415 towns, 1830 and 1911

	(1) N	(2) mean	(3) std. dev	(3) min	(4) max
Population 1830	415	15,057	87,622	1,670	1,737,535
Population 1911	415	55,172	333,691	2,334	6,512,914
Population growth	415	40,114	246,810	-59,056	4,775,379

[6] For more information on the method of digitization, see Satchell, 'Candidate Towns of England and Wales, c.1563-1911 GIS shapefile documentation'.

Growth rate %	415	240.1	383.7	-67.9	4837.5
Log MA 1830	415	12.83	0.39	11.51	13.88
Log MA 1911	415	16.18	0.09	15.35	16.34
Diff in log MA	415	3.23	0.37	2.26	4.58

4.3 Construction of multimodal network

A multimodal network is in essence an integrated network of different transport modes. It allows movement along each transport mode as well as changes between different modes subject to specified connectivity policies. Each of the transport modes has a cost parameter such as, in the case of this article, cost of moving a ton of goods per mile. Where change from one transport mode to another was possible, we specify an extra cost, such as the transhipment cost of moving a ton of goods from a railway station to the connecting road network. Measurements such as the least-cost-route and the total cost of moving along such a route between any origin and destination pair can then be calculated. The cost of moving along the least-cost-route forms an important part of the calculation of market access. For this article, we constructed multimodal networks for 1830 and 1911. The major difference between the two is that the latter incorporates railway network. However, the method of construction is similar.

The first step in constructing our multimodal networks was to topologically modify each transport mode so that it was ready for integration. Standard topological rules from ArcGIS such as 'Must Not Have Dangles' and 'Must Be Single Part' were applied to all transport modes in order to identify and fix disconnections arising from imprecise digitization within each transport mode. Appropriate topological rules were then applied to different transport modes to ensure that they reflected each transport mode's routing reality. For example, topological errors from 'Must Not Intersect' were universally fixed in the road network to make sure that a turn could be made whenever two road segments crossed each other. However, similar topological errors were only fixed when they coincided with stations or junctions in the railway network to make sure that there was no unrealistic turn when railway tracks crossed each other.

The next step was to connect and integrate all of the transport modes and towns together. We built the connections using a set of interpolated straight lines between transport modes and towns. We assumed that the 415 towns mentioned in the previous section were always connected to the road network. Therefore, straight lines were constructed to connect each town to its nearest road segment. Towns did not always fall into the catchment area of a port or a railway station. However, we assumed that, if there was a port or a railway station within a 2km radius around a town, that town was connected to the corresponding waterway or railway network. We then constructed a straight line to connect the town to its nearest port or railway station. With regard to ports, we assumed that they were always connected to the road network and waterway network. Therefore, we constructed straight lines to connect each port to its nearest road segment and waterway segment. However, we assumed that they were connected to the railway network only if there was a railway station within a 2km radius. In this case, we constructed a straight line to connect the port with its nearest railway station. With regard to railway stations, we assumed that they were always connected to the road network and therefore, constructed straight lines to connect each railway station to its nearest road segment.[7] The topologically cleaned transport modes, towns and interpolated straight lines constitute the participating elements in our multimodal networks. We then used the Network Database function in ArcGIS to build our multimodal network.[8] Within

[7] The work of integration was undertaken with a Python script. The script can be requested from the authors.

[8] Specific assignment of connectivity group and connectivity policy for each participating element can be requested from the authors.

this network, given the cost parameters of each transport mode and the transhipment cost between different transport modes,[9] the minimum cost of travelling between any pair of the 415 towns can then be calculated based on the built-in Dijkstra's algorithm in ArcGIS.[10] Snippets of our multimodal network for 1911 are presented in figure 4.

Figure 4. snippet of multimodal network, 1911

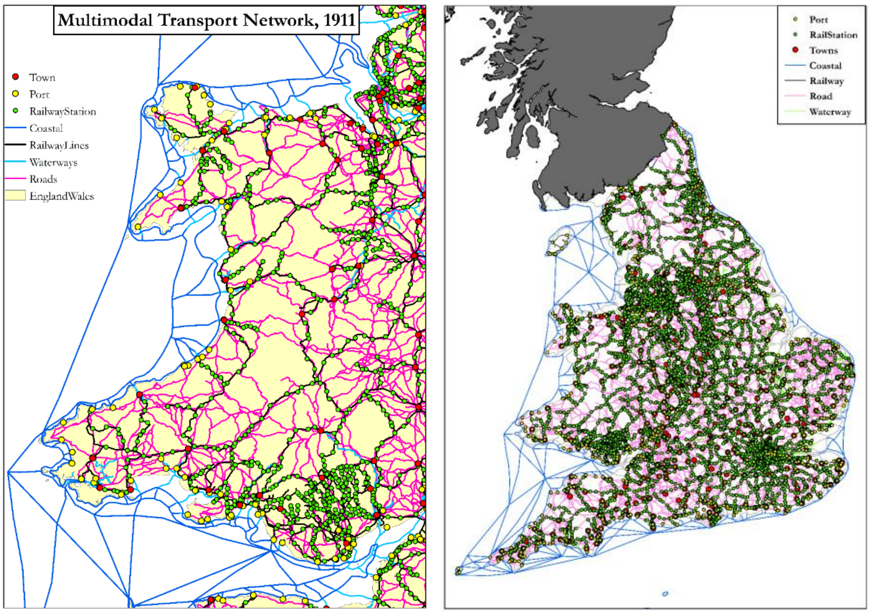


Figure 5 gives a sense of the degree of transport improvement between 1830 and 1911, using our multimodal networks and the corresponding cost parameters. The maps show the isocost regions of freight transport from London at each date. It is no surprise that with the introduction of steam power in railways and coastal shipping, by 1911 freight could travel much further and at a significantly reduced cost. Figure 6 shows another key feature of transport improvement over this period: in addition to the clear decrease in transport costs demonstrated in figure 5, variations in transport costs across urban units also decreased significantly over time. Whether we look at the maximum, minimum, or average transport costs from each town, the range over which towns differed from each other was significantly reduced between 1830 and 1911. It is therefore fair to argue that, by 1911, and regardless of the size and character of different towns, transport efficiency was by and large spatially homogeneous, and was not biased towards larger or smaller towns. This stylized fact has important implications in explaining our estimated results later on.

[9] For our cost parameters and sources, please refer to the data note.
[10] For more information on the details of Dijkstra's algorithm, see ESRI website.

Figure 5. Isocost of freight transport from London, 1830 (left) and 1911(right)

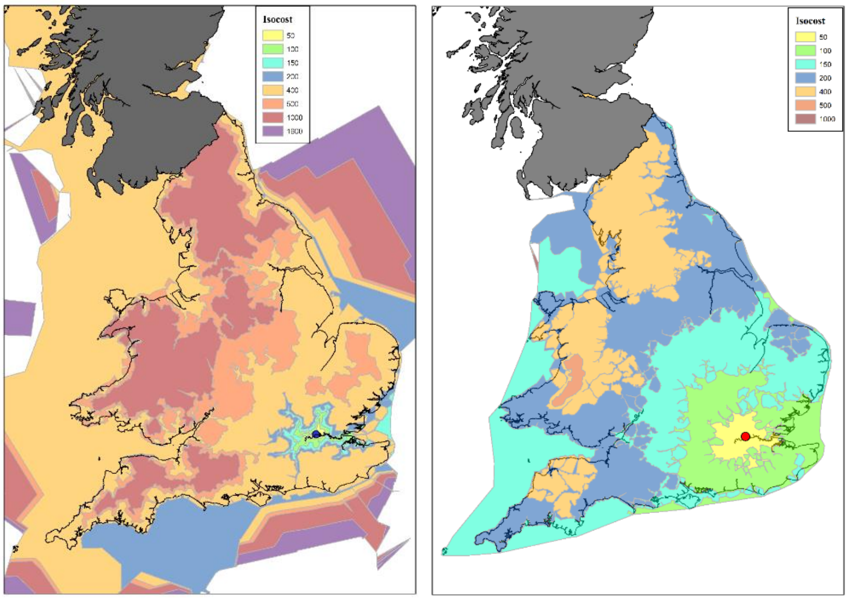
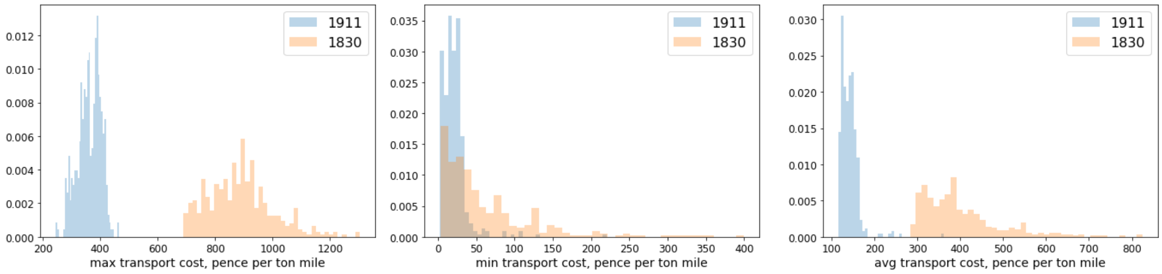


Figure 6. Minimum, maximum, and average transport costs from each town, 1830 and 1911



5. Empirical Specification

5.1 Measurement of market access

Our measurement of market access is based on Donaldson and Hornbeck's first-order approximation formula (2016). It is expressed as

$$MA_{i,t} = \sum_{j=1,\neq i}^{J} \frac{Pop_{j,t}}{\tau_{ij,t}^{\theta}}$$

Where $MA_{i,t}$ is market access for location *i* in year *t*, and Pop_{jt} is the population size of location *j* in year *t*. The latter is drawn from the population figures in Law and Robson's lists of towns (Law 1967, Robson 2006). $\tau_{ij,t}$ measures the transport costs between locations *i* and *j* in year *t*. It can take different forms, such as travel time for passenger traffic or cost for freight traffic, depending on the context of analysis. For the purpose of this article, it is measured as the ratio of the cost of a ton of coal at *j* over the cost of a ton of coal at *i*. The freight traffic of coal was selected on the assumption that it reflects the most relevant indicator of economic efficiency for the period. It is calculated based on the multimodal network and cost parameters detailed in the previous section. The parameter θ is known as trade elasticity. It is inversely related to the standard deviation of productivity across locations. The smaller θ is, the larger the incentives to specialize and trade. The true value of θ is difficult to estimate and often depends on the empirical context.

We therefore opted for a method that consists in fitting different values of θ into our regression model and choosing the value that yields the largest maximum likelihood ratio. The values of θ we tested ranged between 1 and 10, at 0.1 increments. After analysis, the chosen value of θ was 1.6. It is reassuring to know that, using the same method for our market access paper for an earlier time-period, we estimated a larger θ value of 2. Indeed, this is consistent with our expectations – different local economies in England and Wales became more specialized over time. It should be noted that when θ is equal to one, the expression for market access is equivalent to the older concept of market potential (Harris 1954).

The expression for market access reflects a trade-cost weighted sum over population. Within this framework, changes to a location's market access will arise not just from transport improvement in that particular location but also from changes in all other locations within the network. For example, economic events outside location *i*, if captured by population sizes $Pop_{j,t}$, affect the market access of location *i*. Similarly, the transport cost $\tau_{ij,t}$ depends on changes in transport provision in location *i*, location *j*, and anywhere in between. This expression of market access therefore enables analysis of the aggregate effects of transport development including spillovers, rather than simply its localized effects. That what happens in one location depends on changes in all other locations reflects the most fundamental aspect of a network. Donaldson and Horbeck, by solving a general equilibrium model, arrive at a prediction of a log linear relationship between market access and population. This will serve as the basis for our empirical strategy.

5.2 Model Specifications

Our main goal is to estimate the effect of market access on population change in 415 urban units in England and Wales between two time slices, 1830 and 1911. The baseline specification is:

$$\ln Pop_{i,t} = \alpha + \delta Year_t + \theta Z_i + \beta \ln MA_{i,t} + \varepsilon_{it} \qquad (1)$$

Where *t* is the year 1830 or 1911, $\ln Pop_{i,t}$ is the log of the population size for unit *i* in year *t*, $\ln MA_{i,t}$ is the log of market access for unit *i* in year *t*, $Year_t$ is an indicator equal to one if the observation is in year 1911 and zero otherwise, and Z_i captures the unobservable unit fixed effects.

As our panel data consists of two time periods, we use first-difference estimators (FD estimators) to estimate the effect of market access. The baseline estimation equation becomes:

$$\Delta \ln Pop_i^{1911-1830} = \delta + \beta_1 \Delta \ln MA_i^{1911-1830} + \epsilon_i \qquad (2)$$

Where $\Delta \ln Pop_i^{1911-1830} = \ln Pop_{i,1911} - \ln Pop_{i,1830}$ and $\Delta \ln MA_i^{1911-1830} = \ln MA_{i,1911} - \ln MA_{i,1830}$.

We also extend the baseline specification to control for first nature and second nature covariates that may also affect population growth. Our first nature controls include variables such as elevation, slope, temperature, rainfall, and distance to existing transport modes. Our second nature controls include variables such as occupations and population before 1830. Controlling for first nature and second nature covariates, our estimation equation becomes:

$$\Delta \ln Pop_i^{1911-1830} = \delta + \beta_1 \Delta \ln MA_i^{1911-1830} + \beta_2 FirstNature + \beta_3 SecondNature + \epsilon_i \quad (3)$$

We also hypothesize that there may be heterogeneous effects of market access on population growth depending on initial population size. Given the well-established urban hierarchy by 1830, large cities such as London, Manchester and Liverpool were already agglomeration centres. With the introduction of steam transport and denser transport networks, they could attract population and resources at an even greater rate than before as well as than other locations. Meanwhile, the dispersion effect that lesser towns once enjoyed would have diminished once they became better connected to the wider network. For those smaller urban units, transport improvement may have opened up competition in areas in which they did not hold comparative advantage. As a result, they may have grown at a slower rate or even lost population. To account for these heterogeneous effects, we estimate:

$$\Delta \ln Pop_i^{1911-1830} = \delta + \gamma_1 lnPop_{1830} + \gamma_2 \Delta \ln MA_i^{1911-1830} * lnPop_{1830} + \beta_1 \Delta \ln MA_i^{1911-1830} + \beta_2 FirstNature + \beta_3 SecondNature + \epsilon_i$$

6. Results

Table 4: baseline estimates, all 415 towns in the sample

	(1)	(2)	(3)
diff lnMA	-0.512***	-0.279***	-0.220
	(0.0968)	(0.0831)	(0.128)
Constant	1.636	-4.750**	-3.416
	(1.449)	(1.929)	(2.944)
Observations	415	414	414
R-squared	0.207	0.361	0.419
FE	Yes	Yes	Yes
First Nature	No	Yes	Yes
Second Nature	No	No	Yes

Robust standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1
The dependent variable is the natural log difference in town population between 1911 and 1830. The coefficient measures the elasticity of population with respect to market access. Specifications include town fixed effects, year fixed effects, nine region-specific time trends, and year specific cubic polynomials in latitude and longitude. Standard errors are clustered on region.

Our baseline model produces a surprising and even a counterintuitive estimate. It shows that on average, market access had a significant negative effect on population growth: the better connected a town was to larger markets with more consumers and producers, the slower its population growth. More specifically, our baseline results show that, on average, a 1% increase in a town's market access led to a c.0.6% decrease in its population size.

Our estimate is, to the best of our knowledge, the first in the literature to show a negative effect of market access on population growth. This appears to go against some of the most standard economic assumptions and reasoning since Adam Smith highlighted the importance of market size to economic growth. At face value, this result would seem to suggest that enlarged market size, reduced transaction costs, and increased access to larger bases of natural and/or human capital have a negative effect on growth. The key to reconcile this seemingly counterintuitive result with standard economic reasoning lies in the balance between dispersion and agglomeration effects in the context of new economic geography. We will return to this issue at a later stage when discussing the effect of market access according to towns' initial sizes.

Specifications 2 and 3 in Table 4 add first and second nature controls to our baseline model. As alluded to before, one key element driving changes in a town's market access is the improvements in its connection to the rest of the transport network. Transport improvement is not randomly located and the intensity of transport improvement is unlikely to be exogenous.

Instead, it is likely to be influenced by geographical as well as socioeconomic factors. For example, favourable geographies such as flatter slopes make it easier to build new transport links or extend the existing transport network. Locations that enjoyed greater degrees of economic development in the past may also stimulate an intensification of the existing transport network. Conversely, the most significant improvements in transport could happen in places with lesser degrees of previous economic development in an effort to tap into larger growth potentials. The first and second nature variables are added to the baseline model to control for these likely correlations between transport improvement and various geographical as well as socioeconomic factors. Our list of first nature variables includes coal deposits, coastal line, elevation, slope, rainfall level, temperature, and wheat suitability. Our list of second nature variables includes: each town point's distance to the nearest waterways in 1680 and 1760; distance to the nearest Ogilby road, turnpike roads in 1725, and 1750; the population size of the consistent geographical unit each town point is located within in 1801, 1811 and 1821; as well as the share of the adult male labour force in the primary, secondary and tertiary sectors in these same units in 1817.

The absolute values of the negative coefficients from specifications with first and second nature variables are smaller than those from our baseline model. This suggests that the locations and intensities of transport improvement, and hence changes in market access, are negatively correlated with existing favourable conditions. There are two possible explanations behind this observation. Firstly, as alluded to before, locations with previously limited economic growth may be endogenously targeted for their cheaper land and labour. Secondly, and perhaps more plausibly in our case, the transport network may become so dense that many places 'accidentally' receive an improvement despite previously limited economic growth. This point can be visually demonstrated by our railway network in 1911. As a matter of fact, it can be shown that none of the towns in our sample was more than 7 km from a train station, despite their varying socioeconomic characteristics. It is therefore unsurprising that market access improved most significantly in places with previously limited access to the network as the age of steam led to the formation of this dense transport network. However, despite the decrease in the magnitude of the effect after controlling for these correlations, the fundamental story from our baseline model remains the same – on average, market access had a negative effect on population growth.

Table 5: baseline estimates, excluding 20 largest towns in 1830

	(1)	(2)	(3)
diff lnMA	-0.570***	-0.373***	-0.378**
	(0.122)	(0.106)	(0.128)
Constant	0.975	-4.003*	-3.872
	(1.387)	(1.947)	(2.514)
Observations	395	394	394
R-squared	0.248	0.317	0.354
FE	Yes	Yes	Yes
First Nature	No	Yes	Yes
Second Nature	No	No	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
The dependent variable is the natural log difference in town population between 1911 and 1830. The coefficient measures the elasticity of population with respect to market access. Specifications include town fixed effects, year fixed effects, nine region-specific time trends, and year specific cubic polynomials in latitude and longitude. Standard errors are clustered on region.

In an attempt to further address the issue of endogenous transport placement, we excluded the 20 largest towns in 1830 from our regression. Given their large population sizes, it can be reasonably expected that those 20 towns were regional and national socioeconomic centres. New transport networks were therefore likely to be developed to connect these centres first before 'filling the gaps' in later stages, a point which is supported by the chronology of British railway development. In that sense, transport improvement around those 20 largest towns is clearly endogenous. Meanwhile, the other towns in our dataset may be regarded as 'inconsequential units' that received treatment 'by accident' because they were located along the corridors connecting regional centres.

We repeat our regression model on the sample without those 20 largest towns in 1830. The results are presented in Table 5. It is clear that, with this smaller sample, our baseline model with or without first and second nature controls is still showing the same result – on average, market access had a significant negative effect on population growth. Excluding those 20 largest towns from our sample however, the magnitudes of the negative coefficients in each specification become even larger. A comparison of the results in Tables 4 and 5 therefore suggests that the effect of market access on population growth varies based on towns' initial population sizes. While on average market access had a negative effect on a town's population growth, it may build upon existing agglomeration effects and further spur population growth in towns that are already much larger than others. It is to this hypothesis we turn to next.

Table 6: estimates controlling for initial population size in 1830

	(1)	(2)	(3)
diff lnMA	-0.220	-0.247	-2.594***
	(0.128)	(0.134)	(0.750)
lnpop1830		0.0681	-0.774**
		(0.0441)	(0.290)
diff lnMA*lnpop1830			0.260**
			(0.0824)
Constant	-3.416	-3.538	3.332
	(2.944)	(2.853)	(3.859)
Observations	414	414	414
R-squared	0.419	0.423	0.436
FE	Yes	Yes	Yes
First Nature	Yes	Yes	Yes
Second Nature	Yes	Yes	Yes

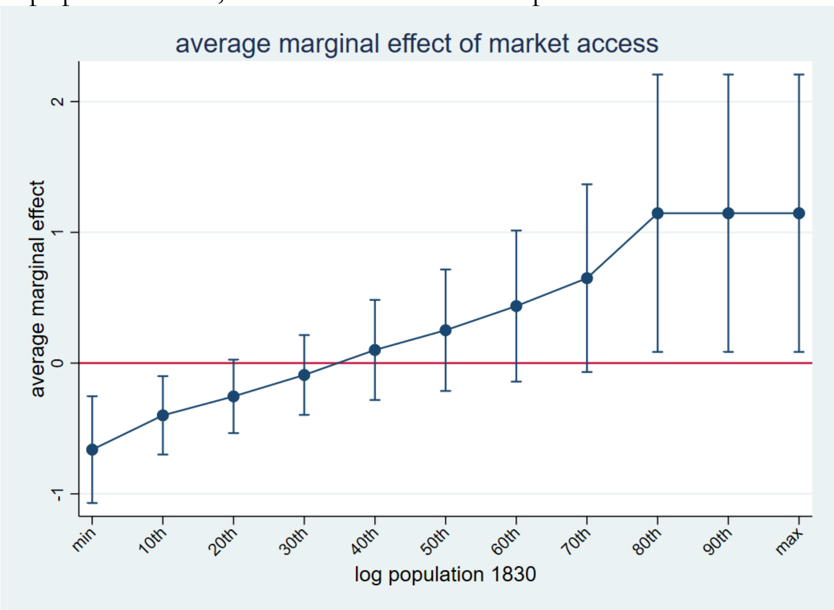
Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the natural log difference in town population between 1911 and 1830. The coefficient measures the elasticity of population with respect to market access. Specifications include town fixed effects, year fixed effects, nine region-specific time trends, and year specific cubic polynomials in latitude and longitude. Standard errors are clustered on region.

We introduce the natural logarithm of town population in 1830 to control for each town's initial conditions. The estimates are presented in Table 6. Column (2) shows that, although it affects the precision of the estimate, the incorporation of population levels in 1830 by and large does not alter the estimated effect of market access on population growth relative to our baseline model with first and second nature controls. It still shows that, on average, a 1% increase in market access led to a decrease of just over 0.2% in a town's population. The more interesting and informative results, however, can be found in our preferred specification in column (3). In this specification, we introduce an interaction term between market access and initial population size to allow for the possibility that the effect of market access on population growth is heterogeneous depending on initial population size. The estimation shows that the coefficients of market access and initial population size are both significantly negative. The coefficient of the interaction term, however, is significantly positive. Hence, there may exist an initial population threshold beyond which the marginal effect of market access on population growth becomes positive.

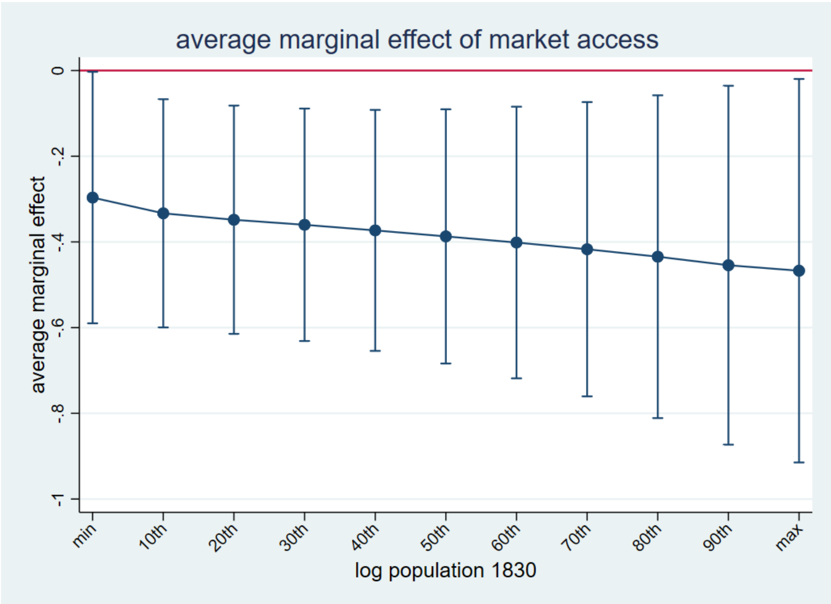
Figure 7 shows the marginal effect of market access on population change between 1830 and 1911 against different levels of initial population size. The estimated marginal effect of market access increases with a town's initial population size. For towns whose initial population size fell below the 20th percentile, market access has a significant negative marginal effect on population growth. For towns whose initial population size in 1830 fell between the 30th and 40th percentile, the estimated marginal effect of market access on population growth becomes positive. However, the estimation is not precise enough. As a result, for towns whose initial population sizes fell between the 30th and 70th percentile, we cannot reject the hypothesis that the marginal effect of market access on population growth is zero at the 5% significance level. It is only for towns whose initial population sizes in 1830 were above the 70th percentile that we find a significant, and positive, marginal effect of market access on population growth. More specifically, for these towns, the elasticity of population growth with respect to increases in market access is greater than 1.

Figure 7: marginal effect of market access based on model with interaction between market access and initial population size, all 415 towns in the sample



Our model cannot delineate the population dynamics between different towns. However, the heterogeneous marginal effects of market access presented so far hint at the possibility that the population of larger towns grew at the cost of smaller towns. Improvements in transport and market access create better and cheaper access to markets for consumption and production factors. On the one hand, this can lead to population growth as people move into local economies with better economic prospects and social amenities. On the other hand, better connections can generate an outflow of economic resources such as population seeking more secure employment and higher living standards elsewhere. The overall outcome for each town's population is likely to reflect the balance between these two forces. In the context of our analysis, towns with initially large population sizes, say above the 80th percentile, must have enjoyed higher degrees of agglomeration effects than smaller towns. Improvement in transport and increase in market access would have further enlarged the agglomeration effects for these towns as the initial condition of a larger economy with a larger population, facilitated by better market access, enabled them to attract more economic activity and population. By contrast, the dispersion effect that small towns had previously enjoyed was likely to diminish as transport connections improved. With better transport connections, economic activity, resources and population could now move away more easily to seek better utilizations. As a result, the intensified agglomeration effects in large towns and diminished dispersion effect in small towns resulting from improved transport and increased market access likely led to different population movement patterns between these towns.

Figure 8: marginal effect of market access based on model with interaction between market access and initial population size, excluding 20 largest towns in 1830



The aforementioned point becomes even clearer when we consider the marginal effect of market access on population growth in the regression model excluding the 20 largest towns in 1830. As shown in figure 8, once we leave out the largest towns, which are most likely to benefit from the intensified agglomeration effect as a result of improved transport connections, the marginal effect of market access is significantly negative for the rest of the towns.

Another way to test our hypothesis concerning the heterogeneous effects of market access is to analyse the effect of market access on occupational structure. It is well established in the literature that agglomeration was stronger in the secondary sector. Hence, we can hypothesise that towns that lost population or experienced less growth because of improved market access may also have lost employment in the secondary sector. If the heterogeneous effects of market access are indeed driven by the changing balance between agglomeration and dispersion effects, we would expect the secondary sector share of employment to be negatively related to market access, especially for smaller towns. Unfortunately, we cannot directly observe each town's occupational structure, and instead use census employment data for the consistent unit each town is located within as a proxy.[11] In Table 7, we apply our estimation specifications to the new dependent variable – difference in the natural logarithm of the secondary sector share of employment between 1851 and 1911. Though not precisely estimated, the sign of the coefficient is consistent with our expectation. On average, market access also had a negative effect on the

[11] We used transitive closure method to generate a set of consistent units across three datasets: 1801 to 1911 parish population, 1817 baptism registers, 1851 to 1911 ICeM data. Detailed description on the consistent units can be found in a separate data note.

secondary sector share of employment. Moreover, as expected, our preferred specification in column (5), with initial population conditions and an interaction term, shows that market access also had heterogeneous effects with regard to secondary employment: market access had a positive effect on secondary employment in large urban centres, while smaller towns lost secondary employment in the light of improved market access.

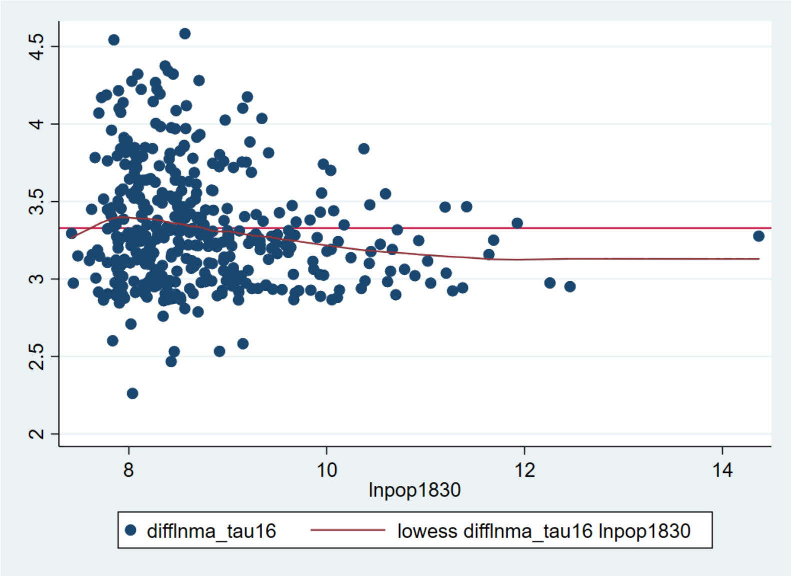
Table 7: effect of market access on secondary employment

	(1)	(2)	(3)	(4)	(5)
diff lnMA	-0.0697	-0.0925*	-0.0798	-0.0771	-0.696*
	(0.0554)	(0.0411)	(0.0484)	(0.0524)	(0.331)
ln secondary share 1851				-0.0359	-0.0580
				(0.125)	(0.120)
lnpop1830				-0.00613	-0.228
				(0.0148)	(0.129)
diff lnMA*lnpop1830					0.0686
					(0.0380)
Constant	0.360	1.242	1.060*	0.996	2.762*
	(0.612)	(0.762)	(0.519)	(0.674)	(1.207)
Observations	413	413	413	413	413
R-squared	0.108	0.207	0.258	0.259	0.270
FE	Yes	Yes	Yes	Yes	Yes
First Nature	No	Yes	Yes	Yes	Yes
Second Nature	No	No	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
The dependent variable is the natural log difference in the secondary sector share of employment between 1911 and 1851. The coefficient measures the elasticity of secondary employment with respect to market access. Specifications include town fixed effects, year fixed effects, nine region-specific time trends, and year specific cubic polynomials in latitude and longitude. Standard errors are clustered on region

So far, we have discussed improvements in transport networks as the driver behind changes in market access. However, based on the formula for market access, we should not ignore the fact that a location's market access can improve due to other locations' faster rate of population growth. Holding transport improvement fixed at the average level across towns, it is mechanically possible for towns with smaller population sizes to have larger market access than towns with larger population sizes. This point is visually demonstrated by a scatterplot of changes in market access by towns' initial population size in 1830 (figure 9). The horizontal line shows the average of the difference in the logarithm of market access between 1830 and 1911, which is c. 3.32. The curved line shows the lowess smoothing. It is apparent that towns with smaller initial population sizes had higher increases in market access on average. To separate the effects arising from other locations' faster population growth from those arising from a town's own improvements in transport connections, we calculate an alternative measure of market access by holding every town's population fixed at its 1830 level. We run our models again using this alternative measure of market access as a robustness check. The results remain similar.

Figure 9: scatterplot of diff lnMA against lnpop1830



7. Conclusion

This article takes a market access approach to study the effects of the introduction of steam-powered transport on urban population change in England and Wales and between 1830 and 1911. Based on transport costs calculated from our multimodal transport model with unprecedented detail, our baseline model produces the unexpected result that, on average, improved market access had a negative effect on urban population growth. We hypothesize that this unexpected result was the average outcome of strong heterogeneous effects of market access according to initial population size. Indeed, our preferred specification shows that while market access had a significantly positive effect on population growth in large towns, it had a significantly negative effect on small towns.

Our results highlight the effects of transport development on urban population change in a context that was considerably different from that in previous studies. The unique historical context of post-1830 England and Wales was charactered by the existing agglomeration centres and the, by and large, spatially homogeneous improvements in transport. This implies that, given the dense network available, small towns were no longer shielded from the competition from larger regional and national centres. They may have either lost population to the latter, or lost their appeal as destinations for rural migrants relative to the latter. Meanwhile, the existing larger agglomerations could draw resources and population more easily from other locations.

Hence, consistent with the stylized facts of urbanization in England and Wales during this period, our study shows that, despite its revolutionary nature, steam-powered transport did not completely change the urban hierarchy, as the transport development in the previous century did. Instead, it led to greater growth rates of the already-larger urban centres, which reinforced the urban hierarchy and agglomerations that we observed for the beginning of this period.

Reference

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.

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

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

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

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

Banerjee, A., Duflo, E. and Qian, N., 'On the road: access to transportation infrastructure and economic growth in China', NBER working paper (2012).

Bagwell, P. S. The Transport Revolution from 1770. London, 1974

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

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

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

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

Bogart, D., 'Turnpike trusts and property income: new evidence on the effects of transport improvements and legislation in eighteenth century England', Economic History Review 62 (2009), 128-152

Bogart, D. "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

Buxton-Dunn, Oliver and Alvrez-Palau, Eduard and Bogart, Dan and Shaw-Taylor, Leigh (2020). *Historical light aids to navigation 1514-1911.* [Data Collection]. Colchester, Essex: UK Data Service. 10.5255/UKDA-SN-854172

Caruana-Galizia, Paul and Martí-Henneberg, Jordi (2013) *European regional railways and real income, 1870–1910: a preliminary report.* Scandinavian Economic History Review, 61 (2). pp. 167-196

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

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

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

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

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

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

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

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

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

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

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

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

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

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.

Freeman, M. J. and Aldcroft, D. H., Transport in Victorian Britain. Manchester University Press, 1988

Flinn, M., The history of British coal industry: 1700-1830. Oxford University Press, 1984

Fogel, R. Railways and American Economic Growth. Baltimore: Johns Hopkins Press (1979).

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

Kotavaara, O., Antikainen, H. and Rusanen, J. (2011) Population Change and Accessibility by Road and Rail Networks: GIS and Statistical Approach to Finland 1970-2007. Journal of Transport Geography, 19, 926-935

Koopmans, C. and Rietveld, P., 'An accessibility approach to railways and municipal population growth, 1840-1930', Journal of Transport Geography (2012), 98-104.

Law, C. M., "The growth of urban population in England and Wales, 1801-1911." Transactions of the Institute of British Geographers (1967): 125-143.

Leunig, Timothy. "Time is money: a re-assessment of the passenger social savings from Victorian British railways." The Journal of Economic History 66.3 (2006): 635-673.

Mineur, C. and Queyroi, F. 'Revisiting the structuring effect of transportation infrastructure: An empirical approach with the French railway network from 1860 to 1910', Historical Methods (2018), 65-81.

Rosevear, A., Satchell, M., Bogart, D., Shaw Taylor, L., Aidt, T. and Leon, G., 'Turnpike roads of England and Wales, 1667-1892', 2017. http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html
Satchell, M. and Shaw-Taylor, L., 'Exposed coalfields of England and Wales' 2013. http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html

Satchell, M., Shaw-Taylor, L., Wrigley E.A., '1830 England and Wales navigable waterways shapefile' (2017). http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html

Satchell, M., Potter, E., Shaw-Taylor, L., Bogart, D., 'Candidate Towns of England and Wales, c.1563-1911', 2017.A description of the dataset can be found in M. Satchell, 'Candidate Towns of England and Wales, c.1563-1911 GIS shapefile' http://www.geog.cam.ac.uk/research/projects/occupations/datasets/documentation.html

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.

Shaw-Taylor, Leigh, and Xuesheng You. "The development of the railway network in Britain 1825-19111." 2020.

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.

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

Wrigley, E. A., People, cities and wealth: the transformation of traditional society. Oxford: Basil Blackwell, 1987.
Waller, P. J., Town, city, and nations: England 1850-1914. Oxford University Press, 1983.