University of St Andrews

Centre for Dynamic Macroeconomic Analysis

Endogenous Infrastructure Development and Spatial Takeoff

Alex Trew

CDMA Working Paper Series No. 1601
12 Jan 2016 (revised 17 Jan 2019)
JEL Classification: H54; O11; O18; O33; N13; N93; R12.
Keywords: Industrial revolution, growth, transport, spatial development.

Endogenous Infrastructure Development and Spatial Takeoff in the First Industrial Revolution*

Alex Trew†
University of St. Andrews

January 17, 2019

Abstract

This paper develops a model in which the evolution of the transport sector occurs alongside the growth in trade and output of agricultural and manufacturing firms. Simulation output captures aspects of the historical record of England and Wales over 1710–1881. A number of counterfactuals demonstrate the role that the timing and spatial distribution of infrastructure development plays in determining the timing of takeoff. There can be a role for policy in accelerating takeoff through improving infrastructure, but the spatial distribution of that improvement matters.

JEL Classifications: H54; N13; N93; O11; O18; O33; R12.
Keywords: Industrial revolution, growth, transport, spatial development.

*I am grateful to Leigh Shaw-Taylor for permission to work with the data produced by the Cambridge Group. I thank two anonymous referees for comments that have greatly improved this paper as well as Dan Bogart, Mark Casson, Radek Stefanski and Scott Swisher for helpful discussion as well as seminar participants at Bristol, Cambridge, St. Andrews, Warwick and the 2017 DYNARE conference in Tokyo. I am grateful for support from the Institute for New Economic Thinking grant #INO15-00025.

†School of Economics and Finance, University of St. Andrews, St. Andrews, Fife, UK, KY16 9AJ. Tel: +44(0)1334 461950. Email: alex.trew@st-andrews.ac.uk. Web: `http://www.st-andrews.ac.uk/~awt2`.

1 Introduction

This paper studies infrastructure development in the context of a growth model where trade occurs across a continuum of space (after Desmet and Rossi-Hansberg, 2014) and where the costs of trading across that space are a function of the endogenous supply of infrastructure at each location. By explicitly modelling the geography and the endogenous determination of trade costs, we can capture the bi-directional interaction between the evolution of trade costs and the location, concentration and growth of firms in different sectors. Such interaction can happen via two channels. First, the growth of employment in transport and distribution services can be a response to the growth of, and the spatial concentration in, the agricultural and industrial sectors. Second, the demand for new infrastructure investment can emerge in different locations as some regions become wealthy or as industrial hotspots emerge far from large markets for first-geography reasons. In turn, the employment of labour in transport services means less labour is available for industry and agriculture; this may lower firm scale and delay the onset of investment in innovation. Further, the lower transport costs that follow infrastructure improvements stimulate further changes to the economic geography. With the model in hand, we match quantitative aspects of the macroeconomic and spatial history of England and Wales over the period of the industrial revolution, from 1710 to 1881. We then use counterfactual treatments of the model to consider the dynamic impact on growth of policies that change the timing and spatial distribution of infrastructural development.

Investment in transport infrastructure is often central to efforts at stimulating development in low-income countries (World Bank, 2015). We know that significant change in economic geography is a feature of development (Desmet and Henderson, 2015) and that falling transport costs have an impact on economic activity (Redding and Turner, 2015). We have only a limited understanding, however, of how transport costs, economic geography and economic growth interact in an economy going through a transition to high growth. How does the quantity of labour employed in the tertiary sector, the quantity of infrastructure investment, its timing relative to transition to high growth, and its geographical focus, affect long-run macroeconomic outcomes such as structural transformation and the emergence of sustained high growth? The answers to these questions should inform policies directed toward infrastructure.

Modern infrastructure investment, even if financed in partnership with the private sector, is generally organized and planned by the State. The experience of England and Wales during the industrial revolution, in contrast, was of infrastructure development that was largely driven locally by private enterprise. The industrial revolution was accompanied by a revolution in transport infrastructure that occupied an increasing proportion of the workforce and that vastly reduced the cost and increased the speed of transporting goods (Bogart, 2014). That transport revolution was a response to, and stimulant of, the reshaping of the economic geography of England and Wales which occurred during the 18th and 19th centuries. England and Wales thus provides a useful benchmark environment in which we can understand the development of transport infrastructure in a laissez faire economy. With a model of that infrastructure development in hand, we can conduct counterfactual analysis to explore the consequences of an alternative infrastructure policy, such as ones similar to those seen today. While we may be hesitant to compare early industrial England and Wales to the modern world, there are Continental European examples in this historical period where centralized infrastructure policy significantly affected the timing and spatial distribution of transport developments (Smith, 1990). Among others, Szostak (1991) has made the case that those policies also delayed growth. Although there are significant differences today, such as the rise of commuting, modern infrastructure is subject to many of the same underlying changes in demand that emerge endogenously as spatial economic change occurs. By studying whether historical policies that centralise infrastructure development had real consequences, we can thus begin to learn about how infrastructure policy and growth are fundamentally related.

We develop a model that captures the evolution of the transport sector in England and Wales over the period 1710–1881. To do so, we introduce endogenous transportation costs to the framework of spatial development in Desmet and Rossi-Hansberg (2014). In particular, we make the cost of transporting goods through a location endogenous to the supply of infrastructure at that location. Infrastructure supply is part stock of fixed infrastructure and part transportation service produced by local carriers. Landowner-carriers[1] hire transportation labour to facilitate the wholesaling and distribution of goods through their land. Landowners may also lease land to infrastructure companies who improve the infrastructure on their land. As in Desmet and Rossi-Hansberg (2014), firms can invest in developing improved production methods at a location and, as in Trew (2014), non-homothetic preferences mean that consumption shifts toward manufactured goods as incomes grow. The evolution of the transport sector occurs simultaneously with those firm decisions that drive the aggregate takeoff in growth rates and the structural transformation of the economy.

We initialize the model using data for the occupational structure of England and Wales at 1710. We then run the model for 171 periods and track its predictions against various macroeconomic variables, such as average growth rate and overall structural transformation, as well as against the spatial distribution of employment in each sector, the endogenous decline in transport costs and the spatial distribution of infrastructure improvements. An industrial takeoff in the North of England, and the specialization of the South in agriculture, means that a greater quantity of output is traded over greater average distances. The demand for transportation improvements emerges locally as a response to the demand for inter-regional trade. Since transport costs also affect the scale of pro-

[1]While agents in the model own a diversified portfolio of all land, we refer to 'local landowners' as agents in each location who manage the local land and receive a share of the total rent.

duction, incentives to innovate and the emergence of agglomerations, there is a feedback from that infrastructural development to the speed of the takeoff in growth.

By modelling infrastructural development as an endogenous process, and by matching that model to the historical experience during the industrial revolution, we can ask how policies which depart from the experience in England and Wales may have affected the timing and speed of takeoff. We find that the timing and spatial distribution of infrastructure improvements can matter. In particular, we show that exogenously higher transport costs can bring forward the date of takeoff as it increases the agglomeration forces that make it more likely firms overcome the fixed costs of innovation. However, early infrastructure investment can accelerate industrial takeoff since it also releases labour out of providing transportation services. Large gains from early infrastructure are only realized if the infrastructure investment is focused on those locations that have a local demand for it. If an equivalent infrastructure improvement is implemented in a spatially-uniform way, the impact on the timing of takeoff is severely muted.

1.1 Related Literature

The paper builds on a number of different strands in the literature. First, we relate to the literature on urban economics, development and the impact of trade costs. Desmet and Henderson (2015) surveys the literature on the relationship between economic development and the changing geographical organization of economic activity. As they show, one particular avenue for research is in modelling the interplay between macroeconomic outcomes, such as growth, and the spatial distribution of economic activity. For this paper, technological progress occurs as firms attain scale in cities; those agglomerations are themselves a function of a transportation network that evolves as the economy grows. While there is a recent literature[2] that makes endogenous the costs of transportation, or the transport network itself, these are limited to static models and, mostly, a discrete number of large spatial units. Here, we study trade across a continuum of space and adopt the spatial development framework of Desmet and Rossi-Hansberg (2014). A further contribution is in understanding the role of transport costs on economic activity. Redding and Turner (2015) surveys the literature that looks to identify the consequences of exogenous variations in infrastructure supply on the spatial distribution of economic activity. These include structural approaches such as Donaldson (2015) and Allen and Arkolakis (2014). Those papers incorporate the general equilibrium effect of improving market access in one location on activity in other locations. Such models tend, again, to be static and typically only model one sector (agriculture). Nagy (2015) models spatial development over time in two sectors and obtains results on the impact of infrastructure on growth. Here we model structural transformation and evolution of infrastructure as it

[2] See, for example, Kleinert and Spies (2011), Asturias and Petty (2013) and Swisher (2015).

occurs over one-dimensional space and time. Fajgelbaum and Schaal (2017) studies the optimal design of infrastructure networks, but in a static model.

Second, the paper relates to the literature on growth, structural transformation and transport costs. As shown in Desmet and Rossi-Hansberg (2012), one way to generate endogenous growth is to model firms that innovate based on it being tied the use of land as an excludable factor of production. This paper uses that framework and extends Trew (2014) to incorporate labour employed in the transport sector. It thus contributes to the literature on structural transformation (see Herrendorf et al., 2014). A number of papers, such as Adamopoulos (2011), Herrendorf et al. (2012) and Gollin and Rogerson (2014), show how exogenous transport costs affect labour allocations in a static setup with exogenous growth. For this paper, transport costs and growth evolve endogenously over time.

Third, we contribute to the literature on the industrial revolution. Shaw-Taylor and Wrigley (2014) have recently shown that the truly dynamic part of the economy over the course of the 19th century, the period in which per capita growth took off (Crafts and Harley, 1992), was the tertiary sector. A large portion of that tertiary sector growth constituted a *transport revolution* that was projected by largely private enterprise which faced a need to reduce the costs of inter-regional trade (Bogart, 2014). Szostak (1991) is among those that have made a case for the importance of transportation Britain's early industrial lead over France. Recent work by Desmet et al. (2015) investigates how lower transport costs increased spatial competition, reduced the power of guilds and led to industrialization. This paper contributes to that debate by simultaneously modelling the role of transportation services and growth. By conducting counterfactual exercises, we can ask whether industrial takeoff may indeed have been delayed as a result of different transport policies.

1.2 Outline

In Section 2 we briefly discuss England and Wales during the industrial revolution as well as evidence on the nature of the transport revolution during that time. We develop the full model of endogenous infrastructure in two steps. First, in Section 3 we make labour in transport and distribution endogenous to output and trade across space but keep the infrastructure stock fixed. In Section 4 we present simulation results which demonstrate that this model can match the aggregate and spatial development of England and Wales over the period 1710–1881. Second, in Section 5 we extend the model to allow endogenous investments in infrastructure stock, comparing the model to the data for England and Wales in Section 6. Section 7 offers some concluding remarks.

2 Occupational change, growth and transportation in England and Wales, 1710–1881

We focus on the industrial revolution in England and Wales for two main reasons. First, infrastructure development in England and Wales during the 18th and 19th centuries was relatively decentralized (see Trew, 2010). Today, infrastructure projects are carefully planned by central governments (even where financed privately). For 18th and 19th century England and Wales, infrastructure development was characterized by local projection (i.e., not organized into a national system) and local finance. While the contrast to modern policies is clear, we can, as discussed below, draw a more direct historical contrast in the policies toward infrastructure between Britain and France. Second, we have exceptionally detailed data at high spatial resolution for England and Wales over the period when it was the first nation to enter a sustained period of high growth and the scene of the most rapid revolution in the speed and cost of transportation to that date. This data permits an understanding of the changing economic geography and spatial takeoff that underpinned the industrial revolution. We can thus ask whether and how the development of infrastructure interacted with the spatial takeoff.

2.1 Occupational structure and growth

For occupational information we use the data described in Shaw-Taylor et al. (2010a) at three dates: 1710, 1817 and 1881. The data is observed at the level of 624 registrations districts covering nearly all of England and Wales. For 1881, the available census records provide a complete picture of the local occupational structure. For 1817, there was no such occupational census. Shaw-Taylor et al. describe the process of constructing a 'quasi-census' based on information in parish baptismal records. The data for 1710 uses the observations of baptismal records in 1,062 of the 11,102 ancient parishes in England at that time. Trew (2014) uses this sample to construct predicted values for the number of adult males in each major occupation (primary, secondary, tertiary and textiles) at the level of the registration district.

Occupations are categorised using the Wrigley (2010) Primary-Secondary-Tertiary (PST) system. Primary occupations are mainly agriculture but also forestry, fisheries and mining (the latter of which we remove).[3] Secondary occupations are manufacturing, processing and construction. The tertiary sector is composed of four groups: Transportation and communication; sellers; dealers; and professionals. The distributive sector (i.e., the tertiary sector less professions) makes up over half the total tertiary employment

[3] In the Wrigley classification, the Primary sector includes mining. We remove mining since it behaves quite differently from agriculture in terms of the aggregate trend and the spatial sorting behaviour. When we refer to primary in this paper, we mean primary less mining.

through the period of study. We use the tertiary sector without professionals.[4] Figure 1 depicts the aggregate structural transformation at the PST level, with professions and mining omitted, alongside the trend in per capita growth of output over the same period from Mokyr (2004).

Figure 1: Occupational structure and growth

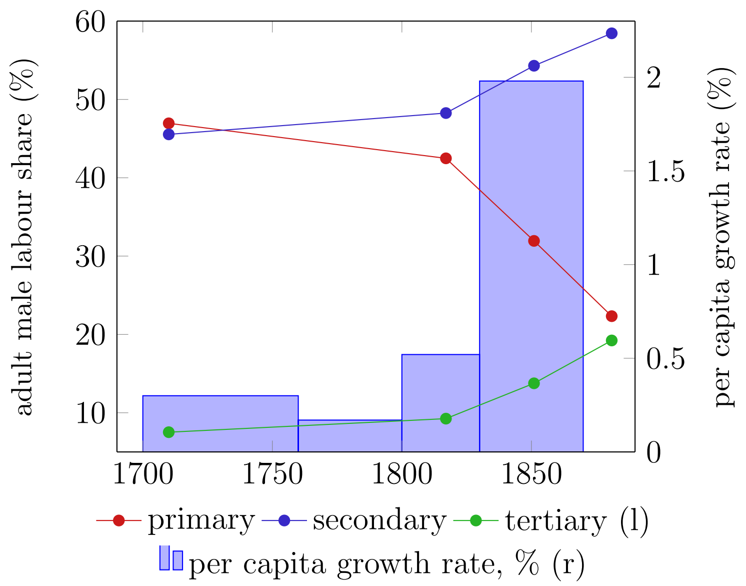


Figure 1 captures the surprising aspect of the Shaw-Taylor et al. (2010b) findings (see also Shaw-Taylor and Wrigley, 2014). The secondary sector is stable and growing over the course of the 18th and 19th centuries, from 45% in 1710 to 48% in 1817 and finally 58% in 1881.[5] The decline in the primary sector starting early in the 18th century accelerates after 1817 but this is not reflected entirely by a growth in the relative importance of industry. Instead, the decline in agriculture is accompanied by the rapid growth of the tertiary sector. Most strikingly, the accelerated shift of labour out of the primary sector and the rapid increase in the tertiary sector is coincident with the takeoff in per capita growth. At a macroeconomic level, this suggests that there is something missing in the typical understanding of growth as simply a result of a shift of resources from agriculture to industry alone. One way of understanding this is bound up in space, as Shaw-Taylor et al. (2010b) conjecture:

> It may be that the majority of tertiary growth in the 19th century was required simply to move the greatly increased output of primary and secondary goods longer average distances around the country. If this is the case then the rise of the tertiary sector was caused, at least in part, by the marked expansion in

[4]The trend for tertiary including professions and services over the 19th century is much like that shown in Figure 1. When we refer to tertiary employment in this paper, we mean tertiary less professionals.

[5]These are shares of employment with mining excluded and with professions removed from tertiary.

the productivity of other sectors and in that sense heralds the onset of modern economic growth.

Shaw-Taylor et al. (2010b), p.21.

That is, tertiary employment growth can be a function of changes in the size and spatial concentration of the primary and secondary sectors.

2.2 The spatial transformation

Economic activity is unevenly distributed across space and that unevenness can change over time and in different ways in different sectors. This is a form of spatial structural transformation. If each sector changed uniformly at all points in space, or if each sector each grew only at one point in space, then the complexity added by modelling that space is not necessarily important for developing a model of what drives growth.

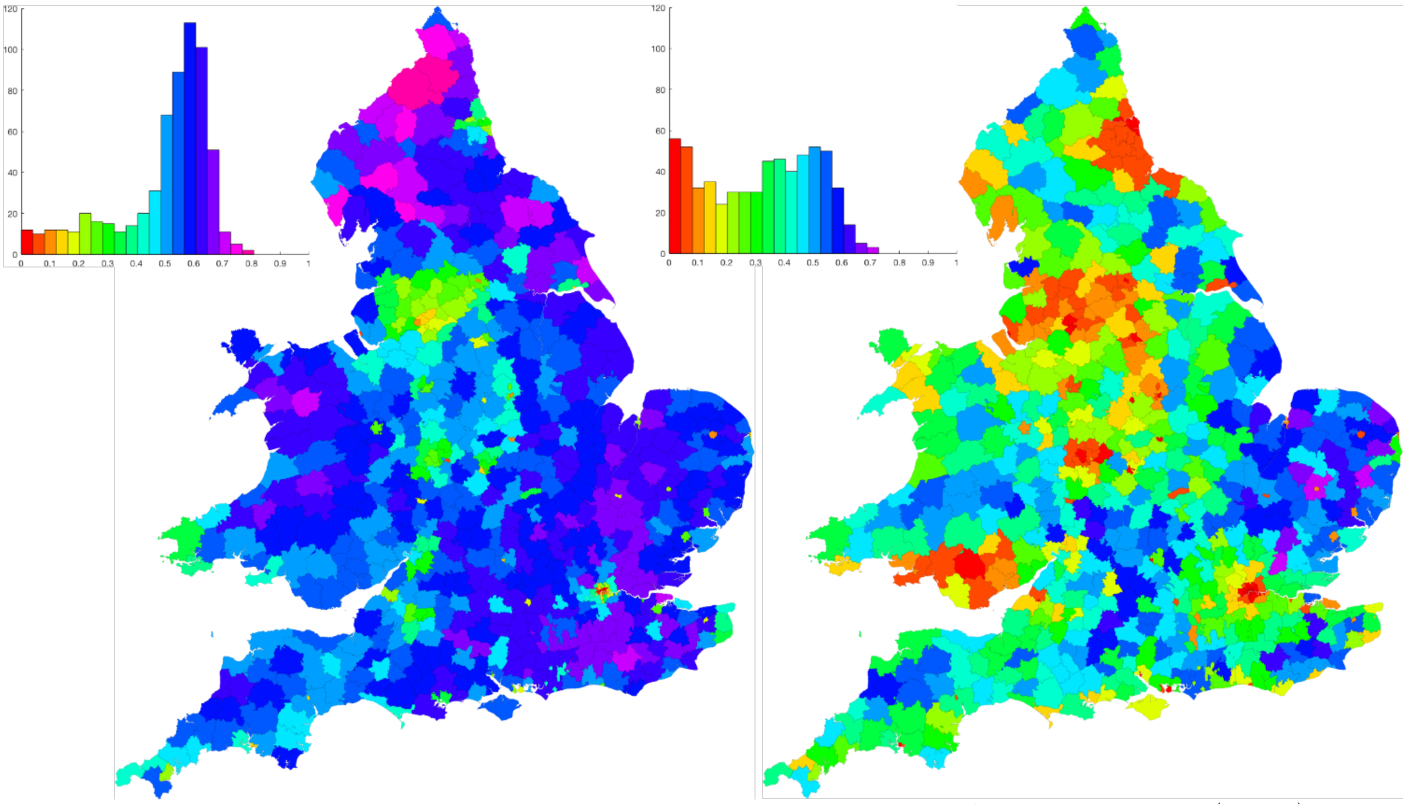
Since we have information at the level of the registration district, we can consider whether there was such a spatial structural transformation. Figures 2–4 use the Shaw-Taylor et al. (2010b) data to map the registration district shares of adult male employment in primary, secondary and tertiary occupations, respectively (colors represent share levels[6]). The primary sector becomes more spatially specialized in the South and East of England. The spatial distribution of the secondary sector is relatively stable over the period. The small industrial hotspots are visible in the North and Midlands; these become slightly larger over the period. As will be seen below, a significant characteristic of change in the secondary sector results from movement of population to the industrial regions. Since, in local terms, those regions at the heart of the industrial revolution were already highly industrialized at 1710, Figure 3 masks this important feature of spatial change. The most striking change is that shown in Figure 4. In particular, there is little clear change to the spatial distribution of tertiary employment. Most areas outside the London area have less than 10% of workers employed in the tertiary sector at 1710. By 1881, no registration district has less than 5% employed and the modal share is in the 10-15% range. The growth in the local importance of the tertiary sector occurs uniformly across all areas of the country, whether they were initially agricultural or industrial. The growth in shares of tertiary employment thus means that in absolute terms, tertiary employment grows along with the growth in employment in other sectors.

The spatial structural transformation in England and Wales over the period 1710 thus takes the form of a primary sector that declines in the North and grows in the South, a secondary sector that maintains its Northern hotspot and a tertiary sector that grows in relative importance at most points in space. A model of this spatial detail, and of how

[6] The histograms at the North-West of each panel are the count of registration districts (y-axis) against the registration district share in that sector (x-axis).

different sectors change spatially over time, could be necessary to understand the drivers of takeoff in the first Industrial Revolution.

Figure 2: Primary Employment Shares, 1710 (l) and 1881 (r)



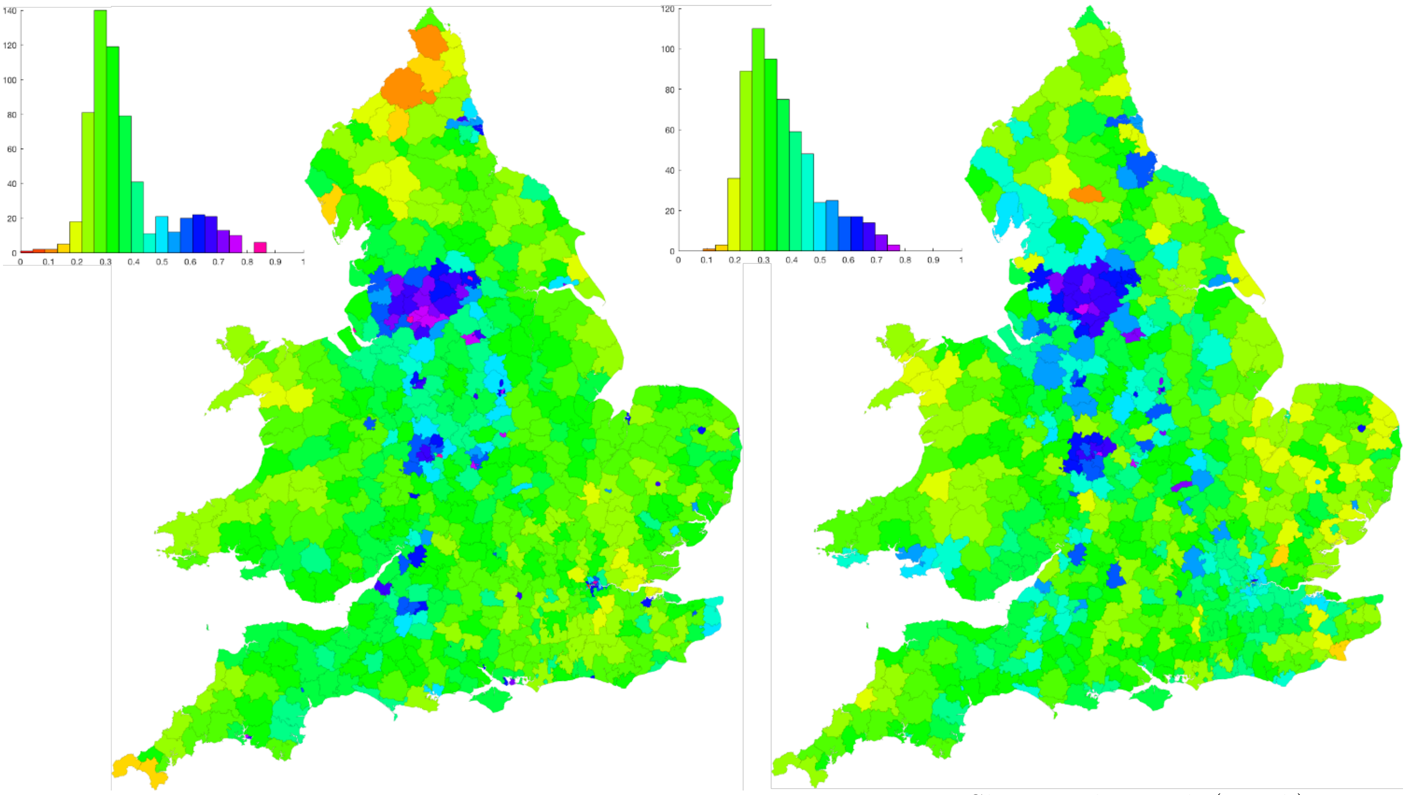
Shaw-Taylor et al. (2010b)

2.3 The transport revolution

Alongside the changes to the number of workers employed in the tertiary sector is significant investment in infrastructure capital and declines in the costs of transportation. The nature of the transport revolution in Britain is discussed in Bagwell (1974) and the excellent survey in Bogart (2014). For Bagwell (p.15), "It was the rapidly increasing volume of inter-regional exchange that made imperative the introduction of more sophisticated forms of goods transport." Those improvements took the form of new technologies such as better paving and river improvements, as well as better modes of transport. For the purposes of this paper, we treat the improvements within a mode and across modes of transport as one continual improvement in transport infrastructure.

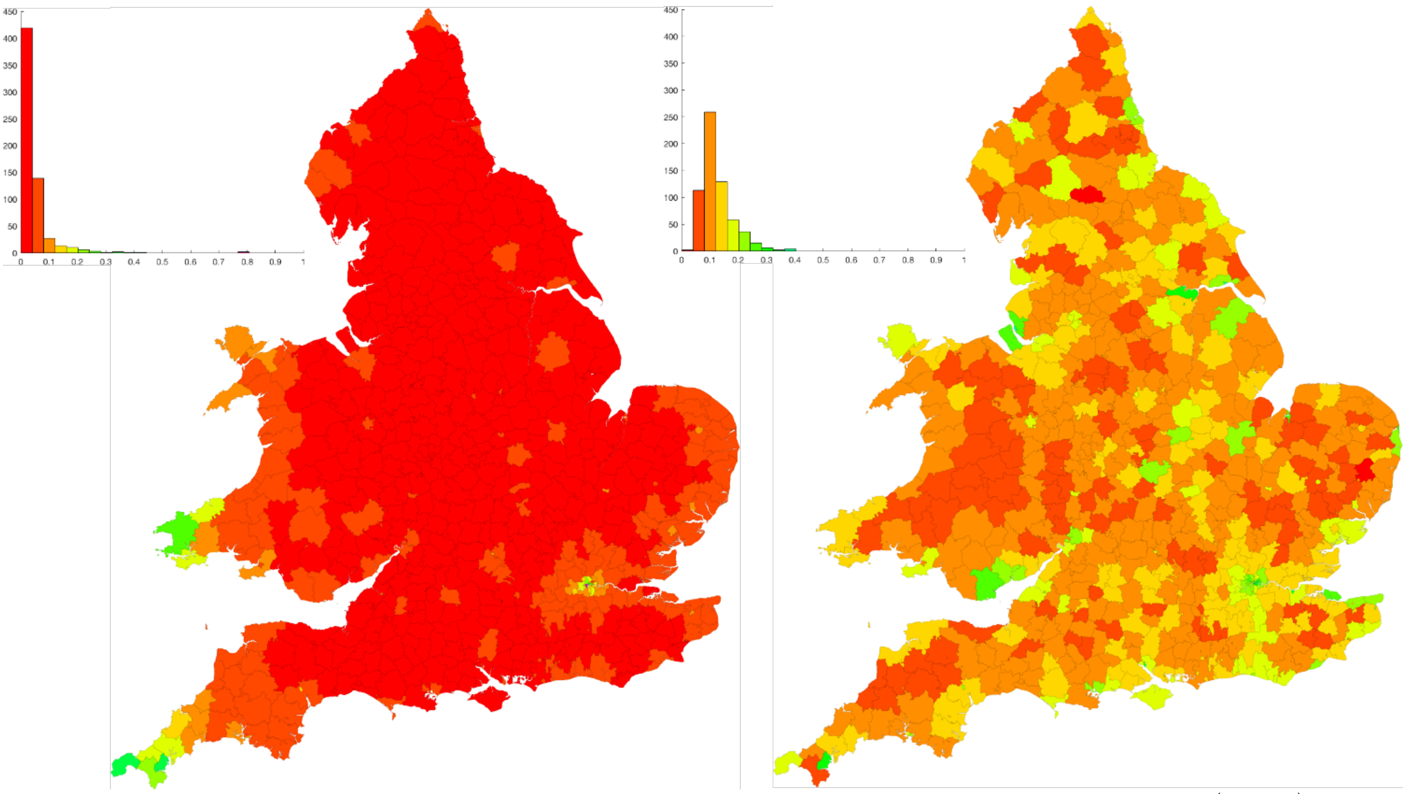
The principal means of improving inland transportation in the early 18th century was the introduction of privately managed roads in the form of turnpike trusts. These trusts, while not profit-making, were permitted by Act of Parliament to charge tolls for passage along the road. As Buchanan (1986) and Bogart (2007) describe, the trusts were run by those local to the infrastructure; they were often formed in response to local demand. A

Figure 3: Secondary Employment Shares, 1710 (l) and 1881 (r)



Shaw-Taylor et al. (2010b)

Figure 4: Tertiary Employment Shares, 1710 (l) and 1881 (r)



Shaw-Taylor et al. (2010b)

distinction is also apparent between the role of the funds raised for the initial turnpike improvement and the levying of tolls for subsequent current spending on maintenance.[7]

The second major mode of transport took the form of extending navigable rivers and the construction of canals, particularly from the early 19th century on. For canals, Parliament authorized the formation of joint-stock companies that could raise the capital required for the expected construction costs. Using data on subscriptions to canal company stock, Ward (1974) shows that, as with turnpikes, the canals were organized and financed by those local to the infrastructure. By consequence, as Turnbull (1987) finds, those early canal improvements again served local markets in a way that was not initially connected to a national system. As Hadfield (1981) describes, canal companies did not typically also provide carrier services to those wishing to move freight (until 1845 they were not permitted to do so without an Act of Parliament). Into the 19th century, the construction of railways became the means of improvement to inland transportation. Once more, as Hunt (1935), Broadbridge (1955) and Shea (2012) document, the railways were often projected for local purposes on the back of local subscriptions to the railway companies. Broadbridge (1955) finds that 'local interests' in financing railways remained dominant until the latter half of the 19th century.

The formation of an infrastructure company was thus the vehicle to overcome the fixed costs of an improvement with the aim to increase the value of local land and the profits of local businesses. Subsequent tolls were typically for maintenance of the infrastructure and for carriage of goods on that infrastructure which was, particularly for roads and waterways, typically carried out by separate entities. Given the uncertainties of the new technologies and unknown local demand, the opportunities for speculation by remote investors were limited and, where investment mania occurred, schemes began to reserve shares for landowners on the route and for the inhabitants of local towns (Pollins, 1954). In many cases, the companies were explicitly promoted as being in the local public interest.[8] An infrastructure company was thus a concern which improved the local economy first, and provided returns to the company shareholder second.

One further aspect of the history worth noting is the nature of dispute resolution when alternative proposed routes were in conflict. Since an Act of Parliament was needed to approve the formation of a trust or joint stock company, competing interests benefited from submitting a single proposal to Parliament with support from those Members of Parliament that owned the land along the route. Parliament was typically not the venue for conflict; as Harris (2000, p.135) notes, it "served only as the arena and set the procedural rules." As such, many non-controversial projects were approved in Parliament with little

[7] As Buchanan (1986, p.228) notes, the Acts of Parliament allowed the "raising of funds in a financial market and the levying of tolls on road users... in general the former provided the long-term capital of the Trust (invested in new and improved roads with their attendant Parliamentary and legal costs) and the latter its current revenue (spent on road repairs, administration, and interest payments)".

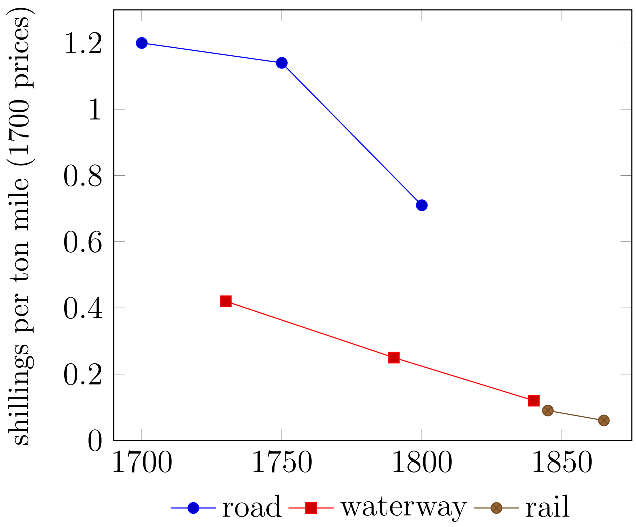
[8] The motto of the Trent & Mersey Canal, completed in 1777, was *Pro Patriam Populumque Fluit* – 'it flows for country and people' (Hadfield, 1981).

conflict. In some instances, however, neighboring cities and counties disagreed on the route to be taken by a canal or railway that would connect them. A good example of the resolution of competing projects is the Leeds and Liverpool Canal, which was authorized in 1770 (see Clarke, 1994). Manufacturers in Yorkshire wanted access to Atlantic trade and a supply of lime for improving agricultural land, while those in Lancashire needed access to the coal inland. Committees in both Yorkshire and Lancashire were formed and as many as five different routes were proposed. A solution was sought by the employment of an arbitrator who chose a route based on the interests of both counties. As Clarke (1994, p.68) concludes, "Following the solution of various disagreements... progress of the Bill through Parliament was swift." The resolution of competing railway plans was often less straightforward, with much debate within Parliament. As Bagwell (1974, p.101) describes, the main consequence was "a golden harvest" for the solicitors representing each interest to Parliament. The sum spent on securing Parliamentary authorization was limited to only two per cent of total outlay on railway capital, at the end of which Parliament would reconcile competing interests and authorize a single company.

The pattern of local projection and local finance of transport improvements reveals just how endogenous the transport network was to regional economic developments. Since industrial developments occurred in a small number of hotspots, thus did infrastructure improve alongside those local developments. Those producers wishing to transport their goods faced two costs: First, the infrastructure company charged a toll on the users of the turnpike, canal or railway. Second, the carrier (the owner of the wagon or boat) charged a fee based on the quantity and type of goods to be carried as well as the distance to be travelled. The overall fall in transport costs over the period was dramatic. Bogart (2014) collates data (see Figure 5) that demonstrates there was a 95% reduction in real shillings per ton mile from 1700 to 1865 across all forms of transport (road, waterway, rail). The fall in freight cost between waterways in 1730 to rail in 1865 is 86%. It should be noted that this is a naïve estimate of the decline in transport costs. The declining freight charges for a growing canal and railway network may be partially offset by the increasing need to move goods across different parts of an increasingly multi-modal network. Recent evidence in Alvarez-Palau (2018) on the costs of trans-shipment across roads, coasts and canals up to 1830 suggests that 15% of freight costs can be attributed to trans-shipment. As the railways later emerged, adding a dimension to trans-shipment costs, real freight charges would be even higher by the end of our period.

To put British infrastructural development into an historical contrast, we can compare it with the situation in France and Germany. The Becquet plan of 1822 envisioned a public-private partnership: A centrally planned waterway system paid for by private capital. A group of civil engineers, the Corps des Ponts et Chaussées, was charged with setting and enforcing the regulations for a waterway network of sufficient quality. As Lévy-Leboyer (1978) notes, the centralized nature of infrastructural development in France

Figure 5: Freight costs (Bogart, 2014)



extended beyond canals and covered also railways. At their introduction, there was also uncertainty over the role of railways in the context of the canal plan (Smith, 1990). Even once Napoleon III began to promote the private finance of a dominant railway infrastructure, private plans were still subject to the layout, location and specifications dictated by the Corps. Milward and Saul (1973, p. 336) argue that government "beset railway building with so many safeguards as to delay its flourishing by a full decade." In contrast to France, disunity of German states meant that no such co-ordinated plan could be developed. As Smith (1990) notes in regard to Germany, a liberal railroad law emerged in 1838 "left companies the intiative to propose routes as well as engineering design" (p.671).

3 Model with exogenous fixed infrastructure

We briefly outline the structure of the model before developing it in detail. Trew (2014) introduces non-homethetic preferences into the spatial development framework of Desmet and Rossi-Hansberg (2014, henceforth DR-H). In this set-up, we have two final goods: Agricultural and manufacturing. As income grows, the marginal utility from consuming agricultural (manufacturing) goods declines (increases) and so consumption (and labour demand by firms) shifts towards manufactured goods. Firms, arranged along an interval of space, hire labour and rent land; they can also invest in a chance to improve their production technology.[9] Perfectly mobile workers can move in advance of productivity

[9] To establish the importance of modelling with a continuum of locations, we introduce in Appendix Section B a simple two-region version of this model. We can draw some conclusions from simply comparing the first period (year 1710) of the calibrated two-region model to the first period of the calibrated continuum model. A calibrated two-region model does not have trade in equilibrium with the original

realisation to the location that yields the greatest expected utility. Subsequent to innovation and production, there is spatial diffusion of productivity before the start of next period.

Since workers consume at the location of their employment, each point in space may export one type of final good and import another. Such trade is costly and, in DR-H and Trew (2014), this cost is assumed to be a fixed iceberg cost. This present study is of the development of transport infrastructure so we make additional departures from Trew (2014) as motivated by two facts. First, at an aggregate level, Section 2 showed that a substantial portion of labour is employed in the tertiary sector (that is, without professionals: transport, distribution and wholesaling). Second, significant resources are expended on improving the stock of infrastructure at a location by, for example, improving roads or constructing canals. Transport costs at a location are determined by both the labour employed in transport and wholesaling and by the stock of transport infrastructure at a location. Moreover, both the labour employed and the stock of infrastructure varies significantly across space and changes over time.

In this section, we make transport costs endogenous to the infrastructure supply at each location. Infrastructure supply is part stock of fixed infrastructure, which we initially take to be exogenous, and part labour used to facilitate the transportation of goods through a location. Conditions on the equilibrium relationship between infrastructure stock, transport labour and transport costs for equilibrium are identified. In particular, while equilibrium trade determines the demand for labour in transport, that labour in transport is also drawing labour away from the production of final goods. As such, we need to ensure that an equilibrium in the goods and labour markets exists. We consider the quantitative performance of the model against the data in Section 4. In Section 5, we make the supply of fixed infrastructure over connected intervals endogenous to investments by infrastructure companies in those intervals.

3.1 Preferences

Agents earn a wage $w(\ell, t)$ from supplying labour to firms which are ordered at points ℓ on the closed interval of land $[0, 1]$, as in DR-H. Agents also hold a diversified portfolio of all land and so receive an equal share of all rental income, $R(t)/\bar{L}$ where $R(t)$ is the aggregate land rent and $\bar{L}$ is the fixed total labour supply. There is no storage good. A consumer solves the following optimization problem,

$$\max_{\{c_A(\ell,t), c_M(\ell,t)\}_0^\infty} E \sum_{t=0}^{\infty} \beta^t U\left(c_A(\ell, t), c_M(\ell, t)\right) \tag{1}$$

transport cost, misses aggregate labour shares and cannot explain a role for energy prices. There is, in this instance, merit in treating space as an interval of many locations.

$$\text{s.t.} \quad w(\ell,t) + \frac{R(t)}{\bar{L}} = p_A(\ell,t)c_A(\ell,t) + p_M(\ell,t)c_M(\ell,t), \quad \forall(\ell,t),$$

where $c_A(\ell,t)$ and $c_M(\ell,t)$ is consumption of agricultural and manufacturing goods, respectively, $p_A(\ell,t)$ and $p_A(\ell,t)$ are prices of agricultural and manufacturing goods, respectively, and $U(\cdot)$ is the instantaneous utility function that takes the following Stone-Geary form, as in Trew (2014),

$$U(c_A(\ell,t), c_M(\ell,t)) = (c_A(\ell,t) - \gamma)^{\eta}(c_M(\ell,t))^{1-\eta} \tag{2}$$

where $\eta \in (0,1)$ and $\gamma > 0$ captures a subsistence requirement in agriculture. Given free mobility of labour, equilibrium prices, wages and rental income equalize utility to $\bar{u}$ across all locations at a given point in time.

3.2 Firms and innovation

A firm at location ℓ can produce either agricultural or manufacturing goods using labour $L_i(\ell,t)$ and productive land[10] (which is normalized to one),

$$A(\ell,t) = Z_A(\ell,t)L_A(\ell,t)^{\alpha}, \tag{3}$$

$$M(\ell,t) = Z_M(\ell,t)L_M(\ell,t)^{\mu}, \tag{4}$$

where $\alpha < \mu$ captures agricultural production that is more land intensive and where $Z_A(\ell,t)$ and $Z_M(\ell,t)$ are the location-dependent productivity levels in each sector.

Prior to hiring labour and making a bid for land, firms may expend resources on obtaining a draw for better technology at their location. In particular, a firm buys a probability ϕ of innovating at sector-dependent cost $\psi_i(\phi)$. If a firm is successful in innovating, it draws a $\hat{z}$ from a Pareto distribution with minimum value 1 and Pareto parameter a_i.[11] Successful innovation yields a production technology of $\hat{z}Z_i(\ell,t)$. We let the expected draw be greater in manufacturing than in agriculture (i.e., $a_A > a_M$). The expected technology for a firm that spends resources on a chance at innovation is thus,

$$E\left(Z_i(\ell,t)|Z_i\right) = \left(\frac{\phi}{a_i - 1} + 1\right) Z_i. \tag{5}$$

Firms that attempt to innovate may offer a greater rent to landowners if the expected gains from innovating outweigh the costs. However, at the end of each period (after production and consumption), technology in each sector is spatially diffused with decay

[10] Each location can hold transport infrastructure on unproductive land.

[11] We assume that innovation draws are spatially correlated – firms arbitrarily close receive the same innovation otherwise an infinite number of draws combined with continuous diffusion would lead to infinite productivity.

δ, that is,

$$Z_i(\ell, t) = \max_{r \in [0,1]} e^{-\delta|\ell - r|} Z_i(r, t-1). \tag{6}$$

Since labour is perfectly mobile, and since technology diffuses at the end of each period, DR-H show that despite the persistence of the new technologies that emerge from innovation, the advantage to innovators dissipates leaving the firm problem as a maximization of current-period profits. That is, a firm chooses ϕ_i to solve,

$$\max_{\phi_i} p_i(\ell, t) \left(\frac{\phi}{a_i - 1} + 1 \right) Z_i \hat{L}_i(\ell, t)^{\imath} - w(\ell, t)\hat{L}_i(\ell, t) - \hat{R}_i(\ell, t) - \psi(\phi(\ell, t)), \tag{7}$$

There are fixed and marginal costs to obtaining a chance of innovation. As in Trew (2014), we let these costs vary by sector and, in particular, we account for the feature of industrial growth that innovative manufacturing technologies were more energy-intensive than agricultural ones. The cost of drawing a probability of innovation ϕ in sector i is,

$$\psi_i(\phi) = \psi_{1,i} + \Gamma_i \xi(\ell) + \psi_{2,i} \left(\frac{1}{1 - \phi} \right) \qquad \text{if} \qquad \phi > 0, \tag{8}$$

where $\xi(\ell)$ is the energy cost at location ℓ,[12] $\psi_{1,i} > 0$ is the fixed cost parameter, $\psi_{2,i} > 0$ is the marginal cost parameter and $\Gamma_i = 1$ if $i = M$ and 0 otherwise.[13] Conditional on the expected net gain being positive, firms choose the ϕ that maximize the expected increase in net profits, that is, the optimal investment probability is,

$$\phi_i^*(\ell, t) = 1 - \left(\frac{\psi_{2,i}(a_i - 1)}{p_i(\ell, t) Z_i(\ell, t) \hat{L}_i(\ell, t)^{\imath}} \right)^{1/2}. \tag{9}$$

As can be seen from equation (9), there is a scale effect present in the intensity of innovation: Higher output at a location is accompanied by a higher optimal innovation probability.

3.3 Transport costs and tertiary labor

Transport infrastructure at a location is composed of a fixed infrastructure (roads, canals, railway tracks) and a transportation service (which would include carriage services, rolling stock, and so on). We assume that the fixed infrastructure is free to access and exogenously fixed until Section 5. In this section, we introduce the role of a landowner-carrier that

[12] We choose this form of energy costs as an alternative to introducing a second secondary production technology since we do not have data to initialize the model with seperate energy-intensive secondary and traditional secondary sectors. In practice, once a firm begins to find it optimal to draw an innovation it continues to do so in each period thereafter (thus generating the sustained growth observed in simulations below).

[13] As in DR-H, in simulations we make $\psi(\cdot)$ proportional to wages to ensure that the cost of innovation keeps pace with the growing economy.

provides transportation services. The carrier charges a toll for the carrying of freight and the construction of rolling stock to run on the fixed infrastructure. We assume that the rolling stock fully depreciates each period. While full depreciation is a simplification, the difference between the longevity of the fixed infrastructure and rolling stock is clear in England and Wales, as elsewhere, where new trains can run on Victorian tracks and where new lorries drive on decades-old roads.

Both the fixed infrastructure and the transportation service determine the cost of transporting goods through a location. The transport cost is made up of a physical cost, $\bar{\kappa}(\ell, t)$, and a toll, $\tilde{\kappa}(\ell, t)$, charged by landowner-carriers. The total transport cost is,

$$\kappa(\ell, t) = \bar{\kappa}(\ell, t) + \tilde{\kappa}(\ell, t). \tag{10}$$

The physical cost results directly from the level of the local fixed infrastructure stock, $T(\ell, t)$. Improving fixed infrastructure reduces the cost of transporting goods by, for example, increasing the speed of travel or by increasing the quality of transportation (reducing spoilage, spillage, and so on). This physical cost is akin to a standard 'iceberg' cost and is lost to the economy. However, the toll is charged by a landowner to fund the production of the transportation service and is paid to transportation workers who then use that income to consume goods. Since the toll $\tilde{\kappa}(\ell, t)$ is not lost to the economy in a normal iceberg-sense, the price of a good i being shipped from location s to location r is thus a function of the physical cost alone,

$$p_i(r, t) = \exp\left\{\int_s^r \bar{\kappa}(\ell, t)\mathrm{d}\ell\right\} p_i(s, t). \tag{11}$$

That is, the price of good shipped from location s to location r takes account of the accumulated melt of the shipped good, $\int_s^r \bar{\kappa}(\ell, t)\mathrm{d}\ell$. In DR-H, the iceberg cost is assumed to be fixed and labor is used only in the production of consumption goods, so $\bar{\kappa}(\ell, t) = \kappa$ and equation (11) reduces to $p_i(r, t) = e^{\kappa|r-s|}p_i(s, t)$.

In this paper, both parts of the transport cost are endogenous. We take the physical transport cost $\bar{\kappa}(\ell, t)$ to be a decreasing function of the fixed infrastructure stock, $T(\ell, t)$,

$$\bar{\kappa}(\ell, t) = \varkappa e^{-T(\ell,t)}, \tag{12}$$

where $\varkappa > 0$. In other words, a better infrastructure stock means faster (or more secure) transport and less melt of goods. Since infrastructure is built on unproductive land at each location, a higher $T(\ell, t)$ does not reduce the land available at that location for firms.[14] We keep $T(\ell, t)$ fixed until Section 5.

[14] At peak, there were 35,684km of turnpikes in 1838, 9,069 km of navigable waterway in 1848 and 31,824 km of railway in 1920. At an average width of 20m for each mode of transport, this makes 1,531km^2 used by these transport infrastructures. The area of England and Wales is 151,174km^2. This suggests that transport infrastructure used around 1% of total land, which is consistent with estimates

The toll is determined by a landowner who produces a transportation service to transport goods using a CES production technology Y_T that is a function of an economy-wide transport efficiency, $Z_T(t)$, transport labour hired, $L_T(\ell,t)$, and the local fixed infrastructure,

$$Y_T(\ell,t) = Z_T(t)\left[\zeta L_T(\ell,t)^r + (1-\zeta)T(\ell,t)^r\right]^{\frac{1}{r}}. \tag{13}$$

where $Z_T(t)$ is exogenous but can grow over time. Production is characterized by a constant elasticity of substitution between L_T and T and we assume that transport labour and fixed infrastructure are substitutes, $r \in (0,1)$. An individual landowner takes the fixed infrastructure as given, choosing labour $L_T(\ell,t)$ to maximize its return,

$$\pi_T(\ell,t) = \tilde{\kappa}(\ell,t)Y_T(\ell,t) - L_T(\ell,t)w(\ell,t). \tag{14}$$

For a given toll, the optimal choice of transport labour, $\hat{L}_T(\ell,t)$, taking Z_T, T, and w as exogenous, is,

$$\hat{L}_T(\ell,t) = \left\{\left[\left(\frac{w(\ell,t)}{\tilde{\kappa}(\ell,t)Z_T(t)\zeta}\right)^{\frac{r}{1-r}} - \zeta\right]/(1-\zeta)\right\}^{-\frac{1}{r}} T(\ell,t) \tag{15}$$

Local labour employed in transport is, ceteris paribus, increasing in transport productivity, infrastructure stock and the toll, decreasing in the local wage rate.

The demand for transportation services, $D(\ell,t)$, is the sum of output produced at a location (which requires wholesaling and distribution) and that traded through the location (which requires transportation). The trade flow arriving at a location is taken as given by its individual landowner. Since firms use a fixed amount of land and choose labour and technology investment optimally subject to wages and prices, the output of a given sector at a location is also invariant to the rent charged. As such, an individual landowner takes this demand for transportation services to be exogenous. The toll that satisfies demand, using $Y_T(\ell,t) = D(\ell,t)$ and equations (13) and (15), is,

$$\hat{\kappa}(\ell,t) = \frac{w(\ell,t)}{Z_T(t)\zeta}\left\{\zeta(1-\zeta)\left[\left(\frac{D(\ell,t)}{Z_T(t)T(\ell,t)}\right)^r - 1\right]^{-1} + \zeta\right\}^{\frac{r-1}{r}}. \tag{16}$$

Equation (16) is the minimum toll a landowner must charge to hire the transport labour required to produce the output to meet the demand at their location. The minimum toll is increasing in the demand for transportation services and so, by equation (15), is the local demand for transport labour. For a given level of demand, the toll is decreasing in Z_T and T, increasing in w.

Since land is modelled as an interval, we assume that all locations between two points r and s must be traversed. Any landowner between r and s could thus choose to set

of the land used by modern road networks of around 1.5-2% (Rodrigue, 2017).

$\tilde{\kappa}(\ell,t) = 1$ to seize all trade flow for local consumption. In a two-dimensional reality, a landowner that did so would find that trade flows around their location by an indirect route over land or by shipping by sea along the coast. For the purposes of this model, we suppose that at each location there is an indirect shipping technology that is identical to equation (13) but has efficiency $\breve{Z}_T(t) = (1-\varepsilon)Z_T(t)$. We may think of this lower efficiency as resulting from the loading of goods onto and off of vessels that would otherwise would go through a location. By (16), this alternative implies a higher minimum toll at each location and so ε captures the technological cost of shipping indirectly. This limits a landowners rent from being the monopoly provider of transportation services at their location. For simplicity we let this ε become negligibly small. Lemma 1 establishes the choice of toll and conditions for it to be bounded.

Lemma 1 *(i) The toll, $\tilde{\kappa}(\ell,t)$, approaches equation* (16) *as $\varepsilon \to 0$ in the presence of indirect shipping. (ii) $\hat{\kappa}(\ell,t) \in (0,1)$ if $D(\ell,t) > Z_T(t)T(\ell,t)$ and $w(\ell,t) < Z_T(t)\zeta^{1/r}$ for all t and ℓ.*

Proof. (i) Indirect shipping around location ℓ with efficiency $\breve{Z}_T(t)$ implies a minimum toll $\breve{\kappa}(\ell,t) > \hat{\kappa}(\ell,t)$ at each location. If a landowner charges more than the alternative shipping then consumption at each location is constrained to that produced locally. Since each location specializes in one sector, worker utility would be zero at any location not engaging in trade. As a result, a landowner sets $\tilde{\kappa}(\ell,t) = \breve{\kappa}(\ell,t)$ and as $\lim_{\varepsilon\to 0} \breve{\kappa}(\ell,t) = \hat{\kappa}(\ell,t)$. (ii) That $\hat{\kappa}(\ell,t) > 0$ if $D(\ell,t) > Z_T(t)T(\ell,t)$ follows from equation (13); if demand for transportation services can be satisfied without transport labour then the toll is zero. If $D(\ell,t)/[Z_T(t)T(\ell,t)]$ grows over time, (16) makes clear that $\lim_{t\to\infty} \hat{\kappa}(\ell,t) = w(\ell,t)/[Z_T(t)\zeta^{1/r}]$. ■

With the toll determined by (16), transport labour is simply,

$$\hat{L}_T = \left\{ \left[\left(\frac{D(\ell,t)}{Z_T(\ell,t)T(\ell,t)} \right)^r - 1 \right] /\zeta \right\}^{1/r} T(\ell,t). \tag{17}$$

Equations (15) and (16) capture two forms of congestion since higher transport demand generates a higher toll and greater tertiary labour. A higher toll lessens the increase in trade and the greater transport labour increases local consumption of the final goods. As such, we require that an increase in output at a location is not fully absorbed by a higher toll or by higher local consumption of the extra tertiary labour. The same is true of an increase in trade that results from higher output. The requirement that congestion is not overwhelming places a slightly stronger restriction on the level of demand relative to non-labour transport input, as the following Lemma shows.

Lemma 2 *A sufficient condition for a limited congestion effect on the toll $\frac{\partial \hat{\kappa}}{\partial D} < 1$ is $D(\ell,t)/[Z_T(t)T(\ell,t)] > 2^{1/r}$ for all t and ℓ.*

Proof. From equation (16),

$$\frac{\partial\hat{\kappa}(\ell,t)}{\partial D(\ell,t)} = \overbrace{\frac{w(\ell,t)}{Z_T(t)\zeta}(1-r)(\zeta(1-\zeta))}^{(a)}\overbrace{\left\{\zeta(1-\zeta)\left[\left(\frac{D(\ell,t)}{Z_T(t)T(\ell,t)}\right)^r - 1\right]^{-1} + \zeta\right\}^{1/r}}^{(b)} \times$$
$$\times \underbrace{\left[\left(\frac{D(\ell,t)}{Z_T(t)T(\ell,t)}\right)^r - 1\right]^{-2}}_{(c)} \underbrace{\left(\frac{D(\ell,t)}{Z_T(t)T(\ell,t)}\right)^{r-1} (Z_T(t)T(\ell,t))^{-1}}_{(d)} \tag{18}$$

Equation (18) is the product of four positive parts each in the unit interval. Since $\hat{\kappa}(\ell,t) > 0$, $w(\ell,t) < Z_T(t)\zeta^{1/r} < Z_T(t)\zeta$ and by assumption $r \in (0,1)$ and $\zeta \in (0,1)$ so we know that part (a) is in $(0,1)$. Since $\zeta \in (0,1)$, (b) is a convex combination of two parts less than or equal to one if $D(\ell,t)/[Z_T(t)T(\ell,t)] > 2^{1/r}$. Part (c) is less than one by the same argument on $D(\ell,t)/[Z_T(t)T(\ell,t)] > 2^{1/r}$. Part (d) is in $(0,1)$ since each bracket is greater than one. Since $\frac{\partial\hat{\kappa}}{\partial D} < 1$, by equation (15) $\frac{\partial L_T}{\partial D}$ is bounded ■

While tertiary labour is determined by economic activity it also determines economic activity via its effect on the labour supply remaining for production of the final good. For labour markets to clear, we need the sum of agricultural, manufacturing labour and tertiary labour to equal the total labour supply.

3.4 Equilibrium in land, labour and goods

Landowners rent land to the firm that offers the highest rental payment,

$$R(\ell,t) = \max\left\{\hat{R}_A(\ell,t), \hat{R}_M(\ell,t)\right\}. \tag{19}$$

where $\hat{R}_i(\ell,t)$ is the maximum land bid that a firm in sector i can make at location ℓ, conditional on optimal labour and innovation decisions.

We let $\theta_i(\ell,t) = 1$ if firm $i \in \{A, M\}$ is producing at (ℓ,t). Following Rossi-Hansberg (2005), $H_i(\ell,t)$ is the stock of excess supply of good i between locations 0 and ℓ. This $H_i(\ell,t)$ is defined by $H_i(0,t) = 0$ and the following partial differential equation, where $x_i(\ell,t)$ is net output of firm $i \in \{A, M\}$,

$$\frac{\partial H_i(\ell,t)}{\partial \ell} = \theta_i x_i(\ell,t) - c_i(\ell,t)\left(\sum_i \theta_i(\ell,t)\hat{L}_i(\ell,t) + \hat{L}_T(\ell,t)\right) - \bar{\kappa}(\ell,t)|H_i(\ell,t)|, \tag{20}$$

where, again, $\bar{\kappa}$ appears since it is that portion of goods traded which is lost to the economy. Clearing in each traded good i requires $H_i(1,t) = 0$. Equilibrium in the labour market requires that the sum of labour in transport and in each sector is equal to $\bar{L}$,

$$\int_0^1 \hat{L}_T(\ell,t) + \sum_i \theta_i(\ell,t)\hat{L}_i(\ell,t)\mathrm{d}\ell = \bar{L} \tag{21}$$

Let $\tilde{L}_G(t)$, denote the total labour that goes to production of agricultural and manufacturing goods at time t and $\tilde{L}_T(t) = \int_0^1 \hat{L}_T(\ell,t)\mathrm{d}\ell$.

Lemma 3 *There is a $\tilde{L}_T$ such that: i) The labour in production $\tilde{L}_G$ generates equilibrium outputs and trade flows associated with that $\tilde{L}_T$ via equation* (15)*; and, ii) market clearing equation* (21) *holds.*

Proof. Let $\varphi(\tilde{L}_G)$ be the total tertiary labour implied by a total productive labour supply of $\tilde{L}_G = L_A + L_M$, that is, $\varphi(\tilde{L}_G) = \max\left\{\int_0^1 \tilde{L}_T(\ell,t)\mathrm{d}\ell \middle| \tilde{L}_G, 0\right\}$ using equation (15). Clearly, $\varphi(0) = 0$. A higher $\tilde{L}_G$ weakly increases $Y(\ell,t)$ and $|H(\ell,t)|$ for all $\ell \in [0,1]$ and so $\varphi' > 0$ by equation (16). Total labour demand is $\Gamma(\tilde{L}_G) = \varphi(\tilde{L}_G) + \tilde{L}_G$ which is thus increasing in $\tilde{L}_G$. Total labour supply is fixed at $\bar{L}$ and so labour market clearing requires $\Gamma(\tilde{L}_G) = \bar{L}$. Since $\Gamma(0) = 0$ and $\Gamma' > 0$, there is a $\tilde{L}_G^* > 0$ at which $\Gamma(\tilde{L}_G^*) = \bar{L}$.■

Having established that an equilibrium can exist in which tertiary labour is endogenous to output and trade across space, we proceed to simulate the model to consider its performance in matching the industrial revolution.

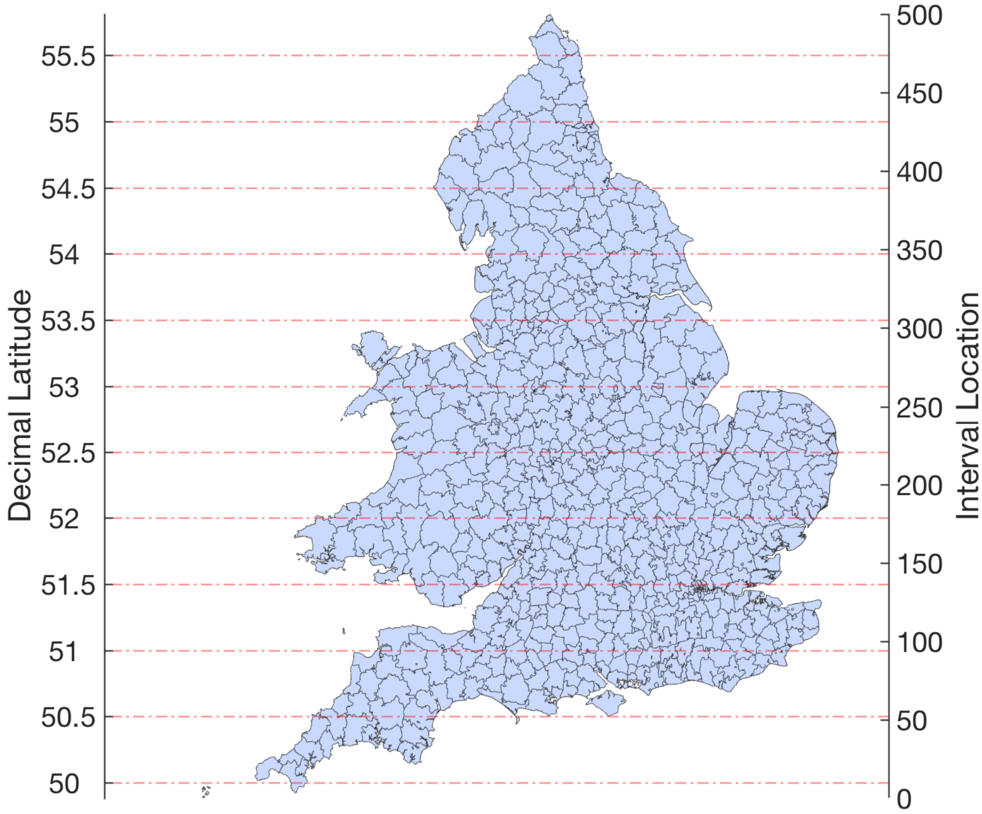
4 Quantitative analysis

The model presented in Section 3 can be simulated and compared to quantitative evidence. We do so using occupational information for England and Wales over the period 1710–1881. A particular advantage of using data for England and Wales is that, as described below, we can map 2-dimensional data into a North-South interval that captures many of the distinct spatial features of the country. We use the initial distribution of labour in agriculture and manufacturing to initialize the spatial distribution of productivity levels and, having parameterized other parts of the model using available evidence, we run the simulation for 171 periods and compare its predictions against the evidence. Before turning to the endogenous infrastructure stock, we can consider the extent to which the model with endogenous tertiary labour matches the aggregate and spatial features of structural transformation, as well as macroeconomic variables such as per capita growth, average relative prices and average land rents.

4.1 Application to England and Wales

In order to initialize the model, and to compare its output to the evidence, we need to map the 2-dimensional occupational and infrastructural information into the 1-dimensional interval of the model. A benefit of using England and Wales for this purpose is that it is of a roughly North-South orientation. To transform the 2-dimensional map of England and Wales we, first, sum occupations along the East-West axis at each point along the North-South axis and then, second, scale each East-West sum by the inverse of the East-West distance. Doing so makes the North-South interval invariant to the East-West size of the country. However, since the South is both more agricultural and broader in the East-West dimension, this scaling generates a slightly lower primary share on aggregate. We thus make a third adjustment and fit the aggregate share to that in the original data by uniformly adjusting employment at all locations. The North-South orientation is then the interval we use to connect with the model (see Figure 6). In simulations, we work with 500 discrete and equally-sized 'parishes' that make up the whole, so the interval in the figures are mapped into this $[0, 500]$ interval for comparison to simulation output. Also shown in Figure 6 are the decimal latitudes that we use to refer to locations in the data and in simulation output.

Figure 6: England and Wales Interval

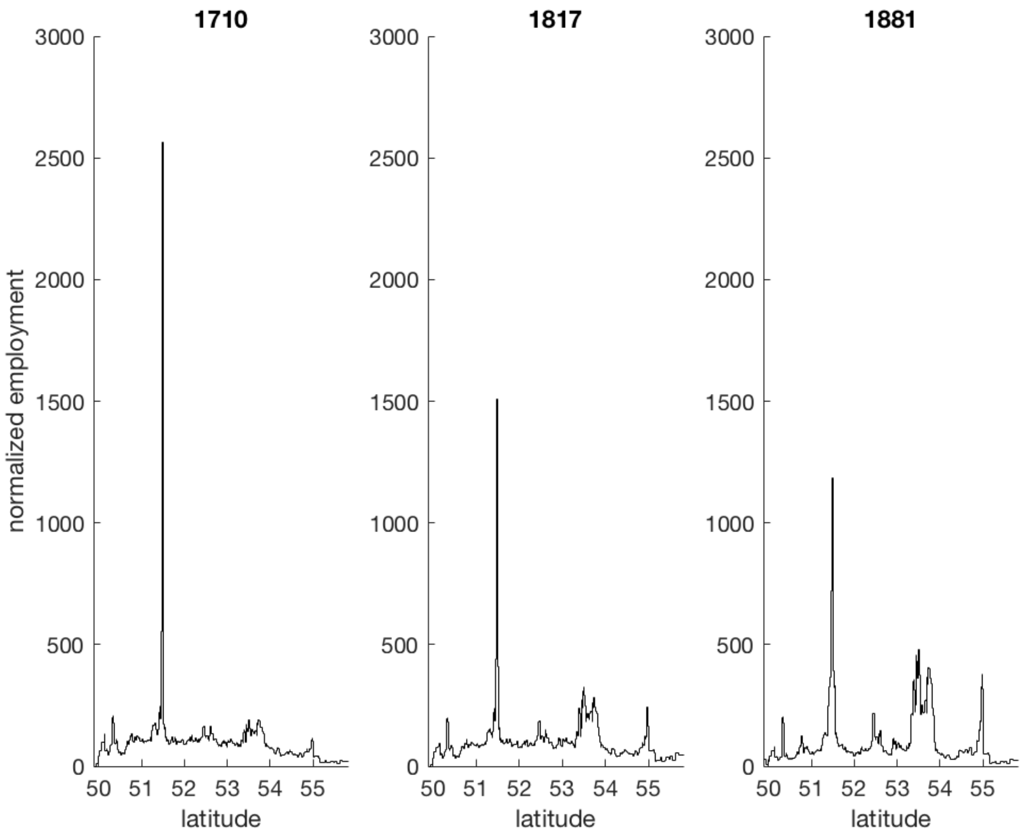


Despite this simplification of the data, many parts of the occupational geography are recognizable in the interval representation. Figure 7 depicts normalized total employment mapped into the interval at three dates.[15] Evident in the figure around latitude 51.5° is

[15]For comparison, Appendix Figure 20 gives the interval distribution of total employed without making

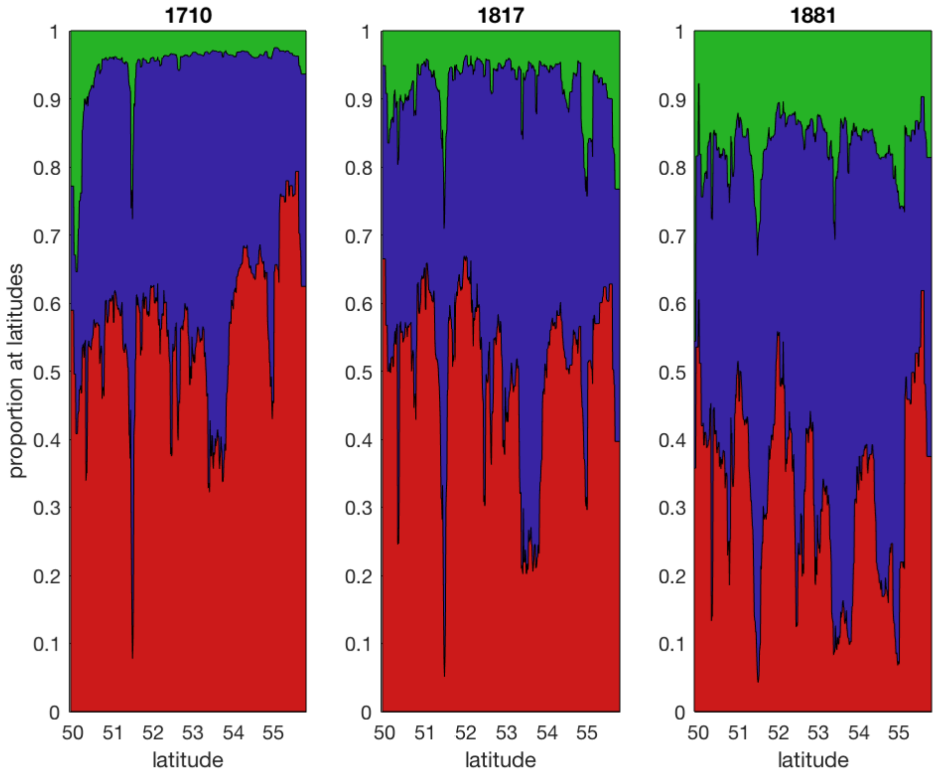
the city of London. Also visible is the growth in the labour employed in the North of England around 53.5° and the relative decline in dominance of London. The interval permits an understanding of both the geographical distance from London to the hotspot in the North and of the magnitude of the shift of employment from the South to the North. One way to look at the occupational structure using the interval representation is to consider the occupational structure at each location. Figure 8 shows this at three dates. While the secondary sector remains relatively stable over the period, the primary sector becomes increasing spatially concentrated. This is a result of the growth of the tertiary sector that occurs most rapidly after 1817 and, strikingly, occurs all over the country. This points to the particular role of the tertiary sector in transport and distribution of the increased output of the primary and secondary sectors; it grows in importance both within *and between* cities. It also shows that the features seen in the 2-dimensional maps (Figures 2–4) are replicated in the 1-dimensional interval representation.

Figure 7: Interval distribution of total employed (normalized)



an East-West adjustment. As can be seen, employment in the middle of the interval appears slightly higher where there is a large East-West breadth while that toward 53.5° appears lower as the country's breadth tapers.

Figure 8: Sector Proportions (bottom to top: primary, secondary, tertiary)



4.2 Parameterization

To parameterize the initial spatial distribution of productivity in primary and secondary sectors, we use the 1710 spatial distribution of employment in each sector. In particular, we invert the production functions, equations (3)-(4), to obtain an expression for local productivity as a function of labour, prices and wages at each location. We then use observations for labour employed along with prices and wages which solve the model. Since the initial spatial diffusion of technologies in the model creates a jump in productivity levels (both on aggregate and spatially), we report model output starting at $t = 1$ after a $t = 0$ diffusion of technology in each sector. (see Trew, 2014, for more detail).

The baseline parameterization is given in Table 1. We select ζ and r to fit the initial aggregate level and spatial distribution of tertiary labour in the model to that observed in the data. Setting $r < 1$ captures a substitutability between labour and infrastructure stock in the production of transportation services. The preference parameters η and γ are chosen to match the initial share of labour in agriculture and manufacturing as well as the extent of the shift out of agriculture over the period. We use Valentinyi and Herrendorf (2008) to pin-down the production parameters α and μ; as described above, $\mu > \alpha$ since agricultural production is relatively more land-intensive. The Pareto parameters, a_M and a_A, are chosen to match the historical record on growth rates. As Allen (2004) documents, there was a long, slow growth agricultural productivity in the period leading up to the industrial revolution; a_A captures a near-doubling of productivity every 150 years. The manufacturing innovation parameter, a_M, generates a long-run manufacturing growth rate of 2% as in Crafts and Harley (1992). The fixed and marginal costs for innovation are chosen to, first, begin at $t = 1$ (year 1710) with some agricultural innovation and, second, to pin down the timing of the takeoff of manufacturing innovation in the baseline simulation. The diffusion decay parameter, δ, affects the pace of takeoff and so we choose is to match the evidence in Crafts and Harley (1992).

Table 1: Parameterisation for baseline model

β	0.95	Standard discount factor.
$Z_A(\ell, 0)$ $Z_M(\ell, 0)$	See text	Data in Shaw-Taylor et al. (2010b) and own working.
$\varkappa$	0.3	High initial transport cost.
$T(\ell, 0)$	See text	
ζ	0.85	To match aggregate initial tertiary labour in Shaw-Taylor et al. (2010b).
r	0.7	
$Z_T(0)$	1.50	
γ_T	1.02	Estimate of TFP growth from Bogart (2014).
x^T/N	1.15	Construction costs in Pollins (1952) and nominal GDP in Mitchell (1988).
η	0.175	To match aggregate employment shares over
γ	0.05	1710-1860 in Shaw-Taylor et al. (2010b).
α	0.59	Firm-level employment share for agriculture in
μ	0.67	Valentinyi and Herrendorf (2008).
a_M	70	Long-run growth of 2%, Heston et al. (2011).
a_A	305	Slow early agricultural growth rate, Allen (2004).
δ	15	To match speed of takeoff in Crafts and Harley (1992).
$\psi_{1,A}$	0.3976	A takeoff in agriculture at $t = 1$.
$\psi_{1,M}$	1.1355	Timing of manufacturing takeoff as in Crafts and
$\psi_{2,A}, \psi_{2,M}$	0.002	Harley (1992).
$\xi(\ell)$	See text	Coal prices from Allen (2009), Clark and Jacks (2007).
$\bar{L}$	100	Normalised total labor supply.

The initial transport parameter is set at $\varkappa = 0.3$. This means an initial physical transport cost of 0.11 (compared with $\kappa = 0.008$ in DR-H) to capture the large costs to transporting goods in the early 18th century. We set $T(\ell, t)$ to be fixed for all t at the initial distribution of access to transport in 1710.[16] For energy prices, we use data in Clark and Jacks (2007) and Allen (2009) on the relative price of coal at different locations

[16]At 1710 there is some variation in access to early turnpikes and navigable rivers. The calculation of this is described in Section 6.1. In practice, relative to the later emergence of canals and railway, this implies a distribution of underlying fixed infrastructure that is close to being spatially uniform.

in 1700. In counterfactual exercises below we vary the initial distribution of the coal price and interact it with developments in transport costs.

4.3 Simulation output

Results from using the baseline calibration are reported in Figures 9-15. In each figure, the thick line is a mean average of 100 simulations with shading to represent a confidence interval of two standard deviations around the mean. Since the model incorporates a continuum of firms, the randomness of innovation realizations should disappear on aggregate in a single run. However, as in DR-H, we model a finite number of discrete intervals that each receive the same innovation realization. That the confidence intervals are relatively tight to the mean model output suggests that the chosen number of discrete intervals is enough to ensure that separate runs of the model do not generate widely different outcomes.

Figure 9 gives the performance of the model in terms of a number of macroeconomic variables. The takeoff in growth matches the timing and magnitude of that reported in Mokyr (2004). The relative price of manufactured goods captures the decline observed in the data (calculated using the method in Yang and Zhu, 2009). The path of real wages and average land rent is consistent with the data in Clark (2002) although with perhaps too much increase in real wages and too little increase in land rents.

In terms of the aggregate structural transformation, Figure 10 demonstrates some success in matching the data described in Section 2. Employment in the secondary sector exhibits a slow, steady increase over the whole period. This is despite that sector generating the increases in productivity that underpin the takeoff in growth. The more rapid decline of the primary sector after 1800 is accompanied by a larger acceleration in the employment of the tertiary sector. The driver of this change is the increased output and trade referred to by Shaw-Taylor et al. (2010b) which in the model causes a greater demand for transportation services. Absent the endogenous tertiary sector, the model would not explain the slow change in the relative importance of the secondary sector alongside the fast change in the relative importance of the primary. As is clear, the growth in the relative importance of tertiary labour outstrips that in the data and a consequence is that the primary sector falls too much and the secondary sector grows too little. Since infrastructure is fixed in this version of the model, landowners can respond to the increased demand for transport services only by hiring more labour. As we will see in Section 5, when we permit investments in fixed infrastructure, some of the pressure on tertiary labour can be relieved by investment in infrastructure improvements.

The benefit of using a model that incorporates continuous space is that it can be compared against data that offers a high spatial resolution. To that end, Figure 11 plots the model implications for the distribution of primary employment against the data at

Figure 9: Baseline Simulation: Macroeconomic Variables
(Average of 100 simulations $\pm$ 2s.d.)

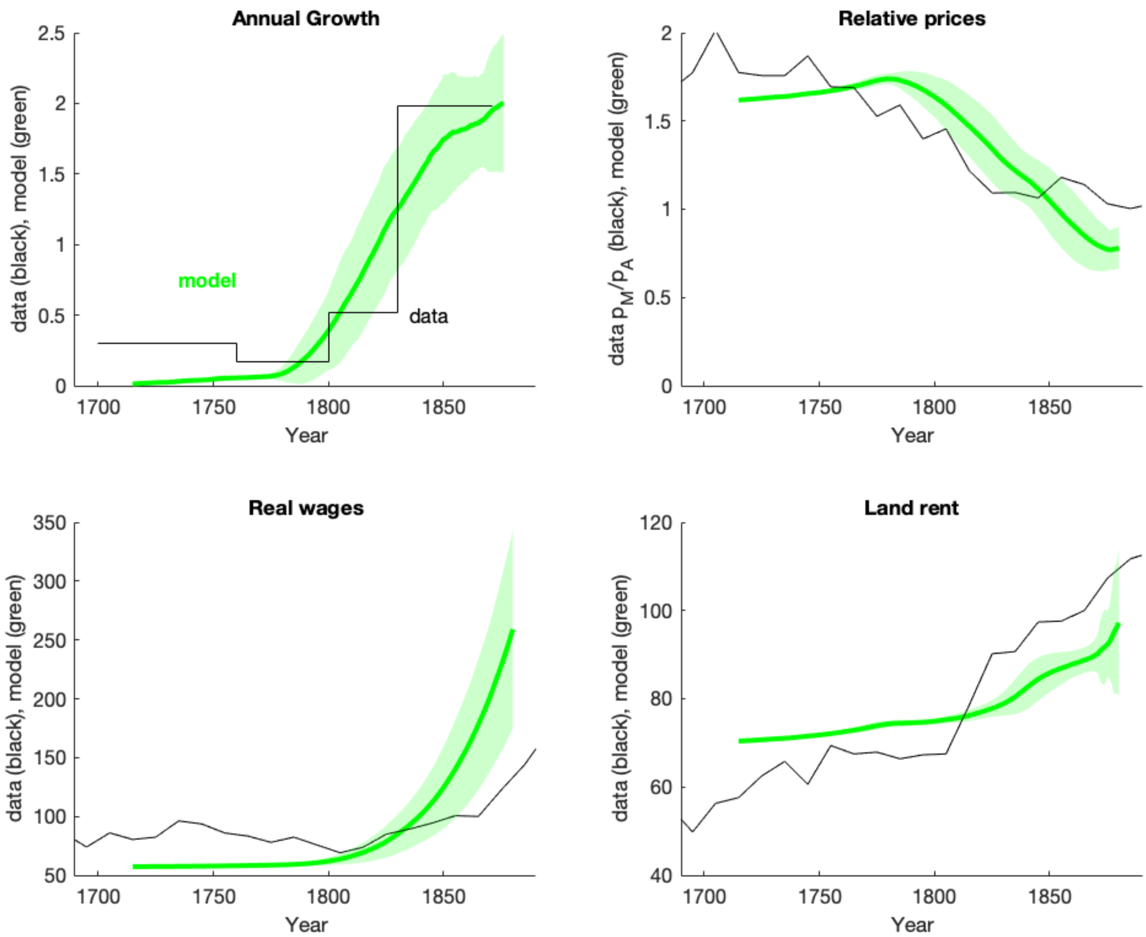
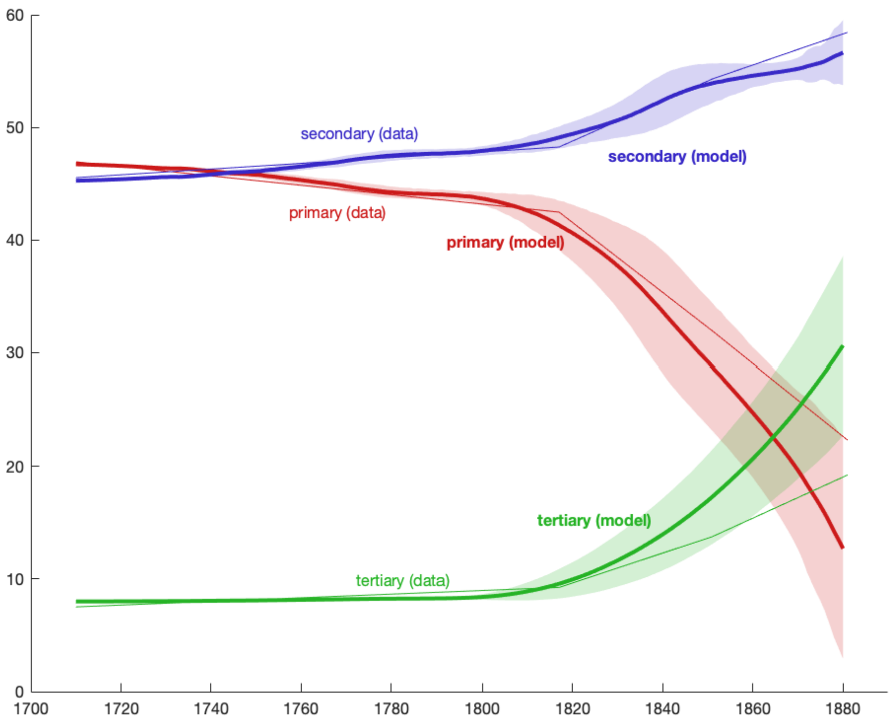


Figure 10: Baseline Simulation: Structural Transformation (% labor share)
(Average of 100 simulations ± 2s.d.)

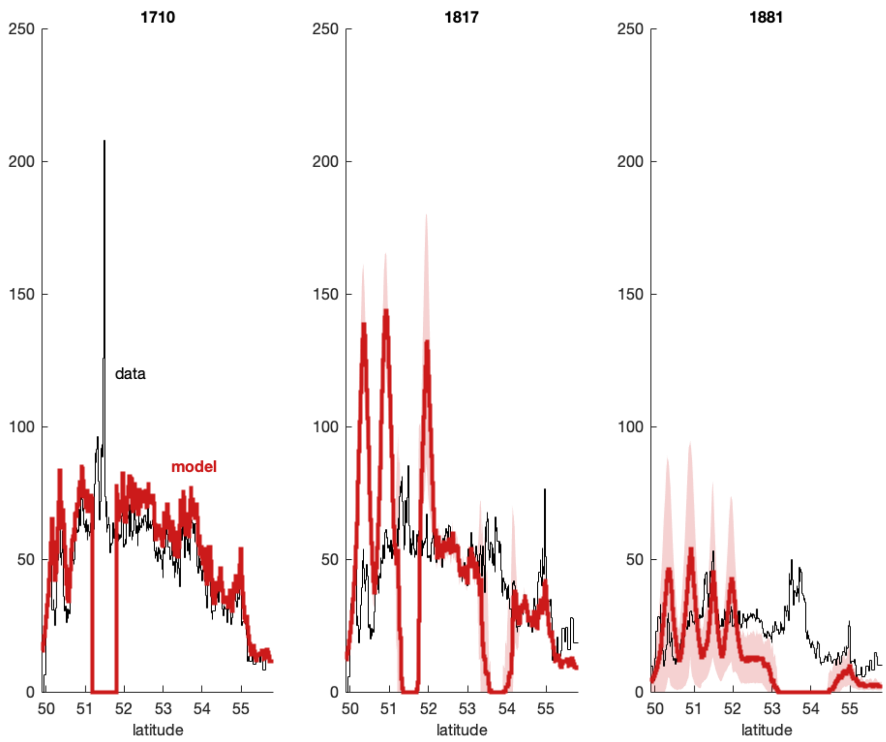


three dates (1710, 1817 and 1881).[17] Productivity growth in the agricultural 'south' of the interval (over the latitude range of roughly 50.5° to 52°) occurs endogenously as firms there are initially larger and can amortize the fixed costs of innovation. This causes primary output to specialize in the south in a way that mirrors the data, although less clearly in 1881, and which is consistent with the literature (Allen, 2004). The localized agricultural innovations cause labour to be more concentrated than in the data. The eventual decline in the southern primary workforce in response to the shift toward manufactured goods matches that in the data. However, the complete primary-secondary specialization in each location means that the model misses some primary employment in the South at 1710 and some primary employment in the North at 1881. Figure 8 places the data in Figure 11 in context – while there are areas of high primary employment shares in the North, they decline in relative terms as the secondary sector grows. For example, the high primary employment level in 1881 around latitude 53.5° is only around one sixth of the level of secondary employment at the same location. The model captures the relatively high share of primary employment in the South and, as we see below, the relatively high share of secondary employment in the North. Moreover, the specialization is not critical

[17]The discrepancy between 1710 and 1817 at around 51.5 in Figure 11 is a result of the reconstruction of the 1710 data using the roughly 10% of the 11,102 ancient parishes for which we have occupational information at that date. As described in Trew (2015), occupational information for registration districts in 1710 is estimated using a parish-level regression model to predict the missing parish information before aggregating up to the 624 registrations districts. Given the large number of residents of London, the primary sector there is likely over-represented at 1710.

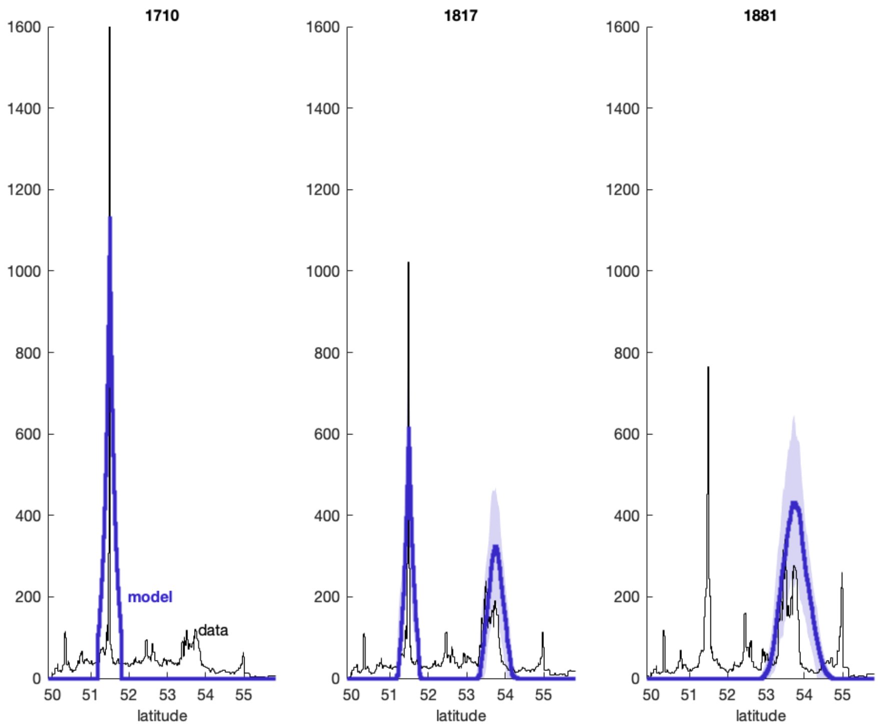
for the ability of the model to capture the industrial takeoff – the model captures a slow agricultural revolution in the South in spite of missing some primary employment in London at 1710.

Figure 11: Baseline Simulation: Primary Employment
(Average of 100 simulations ± 2s.d.)



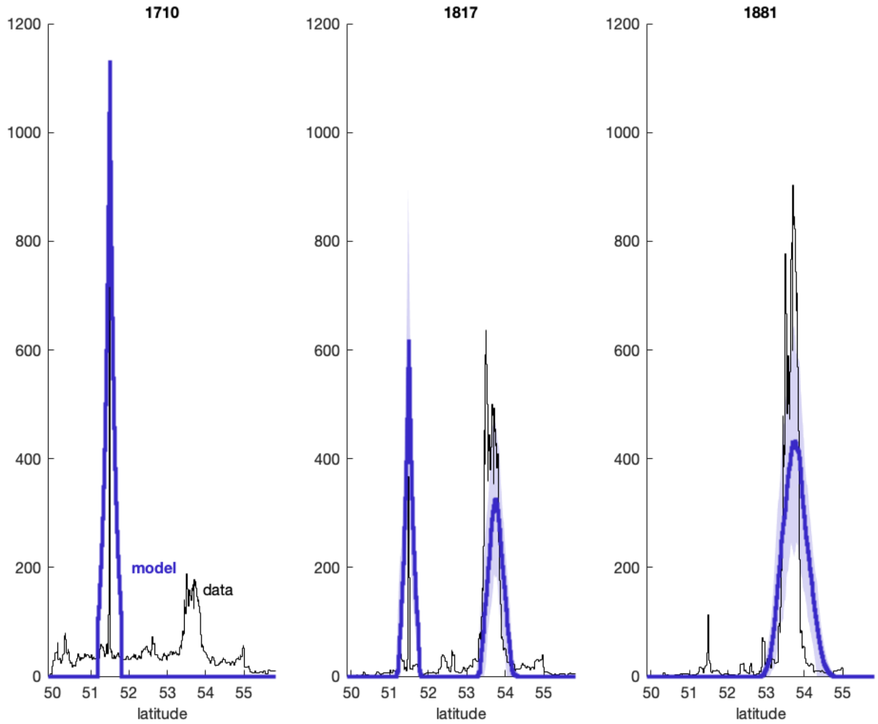
As the agricultural revolution proceeds in the south, consumption demand, and so labour employed, shifts towards manufacturing firms. That raises the optimal size of manufacturing firms and makes them more likely to be able to offer land rents in excess of those offered by agricultural firms. The growing scale of manufacturing firms also means they may amortize the fixed cost of innovation over a larger output, while because of the scale-effect in equation (9) the innovation intensity increases. At $t = 72$ (year 1782), manufacturing firms at latitude 53.7 find it optimal to invest in innovation. Thereafter, the innovation-driven industrial revolution proceeds; aggregate growth increases and the shift of consumption out of agriculture accelerates. As Figure 12 shows, labour in the secondary sector moves toward a northern hotspot. The location of this industrial hotspot matches the data, though the size of the takeoff in the model is in excess of that in the data. The secondary employment in the model also predicts that London (around latitude 51.4°), declines more than in the data. In reality, of course, London continued to be an important centre of activity throughout the eighteenth and nineteenth centuries. However, there are a number of things not in the model that make London different. Following its founding by the Romans, London grew into the administrative and legal capital of the country, eventually forming a large part of the first financial revolution following the Glorious Revolution in 1688. Its principal sources of employment throughout this period were not those sectors that underpinned the industrial revolution. As the model is of only one sector, a better comparison is to the single sector, textiles, that during this period drove industrial growth. As Figure 13 shows, the model captures the location, timing and magnitude (once rescaled) of the takeoff in textiles employment. Moreover, we can consider whether the model would still be able to capture the endogenous emergence of a Northern hotspot if London were arbitrarily forced to persist. What triggers the rise of the industrial hotspot in the North is a steady increase in the relative price of manufactured goods that results from the slow agricultural revolution in the South. With an arbitrary 'capital employment' in the secondary sector added to the model around 51.4°, we would still see this change in relative prices so long as the agricultural revolution in the South occurs.

Figure 12: Baseline Simulation: Secondary Employment
(Average of 100 simulations ± 2s.d.)



Finally, we can consider the spatial fit of the model to the tertiary employment data. As Figure 14 shows, the model does well in matching much of the local tertiary employment, particularly in the initial period and in the industrial hotspot that emerges. The model does less well in others (such as, again, in London). The model prediction for the transport labour in the industrial region is greater than that in the data, again since infrastructure supply is fixed in this version of the model. One of the most salient aspects

Figure 13: Baseline Simulation: Textiles Employment (rescaled)
(Average of 100 simulations ± 2s.d.)



of the dynamism of the tertiary sector was shown in Figures 4 and 8, that the growth of the tertiary sector grew as a proportion of local employment in a highly uniform way. As Figure 15 shows, the model is able to explain the uniform upward shift in the share of tertiary employment across the interval.[18] This is quite distinct from the model implications for the primary and secondary sectors which mirror the data in concentrating in one region. The model over-estimates the share in the middle of the interval. This partly results from the assumption that goods are traded wholly across the land. As a result, the model predicts that the accumulated traded good peaks around 53.1°. In reality, some trade would take place in coastal shipping which would connect points in the South with the North without inducing transport labour demand in the Midlands. In being able to match the data with the tertiary sector explained as a function of output and trade in the other sectors, the findings are consistent with the hypothesis presented in Shaw-Taylor et al. (2010b).

[18] Appendix Figures 21-22 report similar figures with shares for the primary and secondary sectors. Since in the model each location specializes in the production of one good, the fit of the shares to the data is not as clear as in the tertiary employment. What Figures 21-22 show is that the ability of the model to fit other aspects of the data, such as the local levels of primary and secondary employment, as well as aggregate variables, is principally down to the movement of population across space, rather than changes in local employment shares.

Figure 14: Baseline Simulation: Tertiary Employment
(Average of 100 simulations ± 2s.d.)

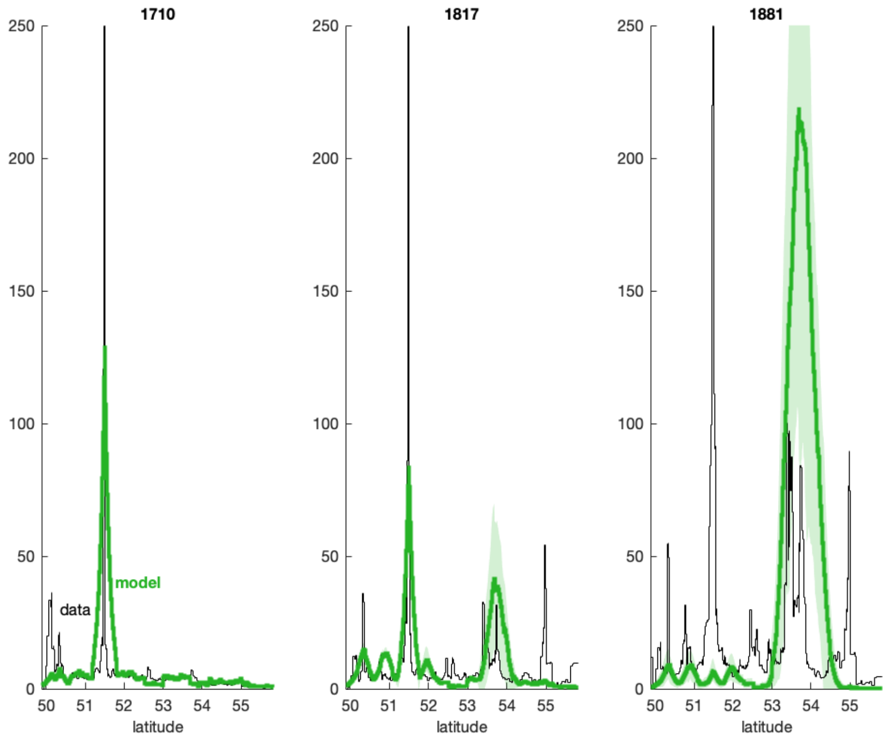
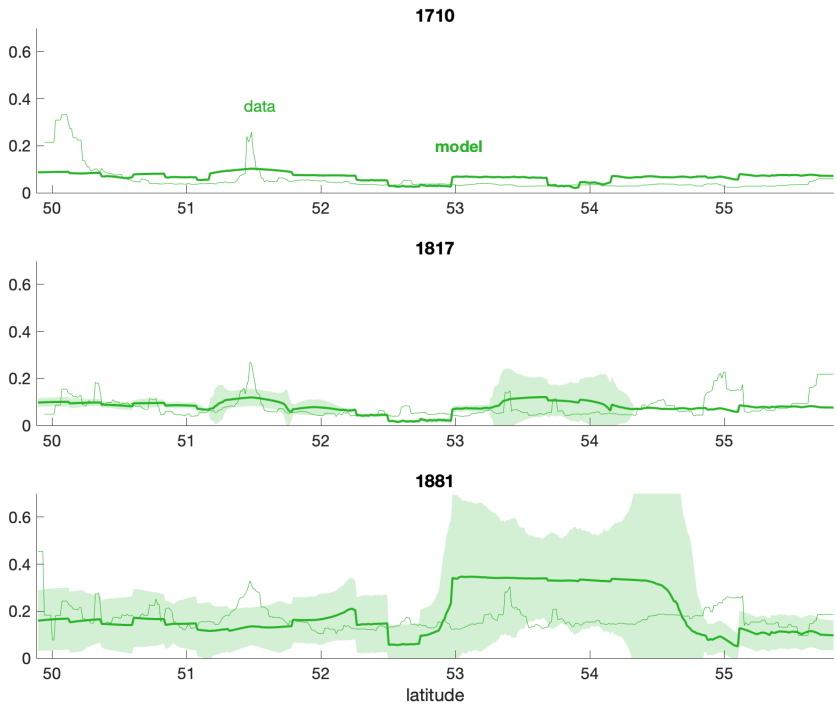


Figure 15: Baseline Simulation: Tertiary Shares
(Average of 100 simulations ± 2s.d.)



4.4 Counterfactual policies

Before moving to the model with endogenous infrastructure investment, we can consider the impact of policies that might be used to stimulate industrial takeoff. There are two main mechanisms at work in the impact of these policies. First, since tertiary labour is not directly productive, policies that release labour into the primary and secondary sectors can bring forward the point of industrial takeoff since they can increase the size of firms and make it more likely that a firm can amortize the fixed costs of innovation. Second, a policy can, directly or indirectly, affect transport costs which can have a consequence for the spatial concentration of firms. We present in Table 2 the results of four counterfactual exercises and the baseline from Section 4.3. The table reports the year of the first instance of manufacturing innovation (the year of takeoff), the decimal latitude of that first innovation, the average transport cost in period 1 and the average labour share in period 1.

Table 2: Baseline and Counterfactuals (averages of 100 runs)

	takeoff year	takeoff latitude	years to 1.5%	initial av. κ	initial $L_T\%$
Baseline	1781.7	53.7°	59.3	0.20	8.00%
1. Better transport technology	1761.7	53.7°	58.3	0.19	6.15%
2. Transport labour subsidy	1781.6	53.7°	59.4	0.19	8.00%
3. Higher transport costs	1770.8	53.7°	57.2	0.24	8.00%
4. Lower transport costs	1798.8	53.7°	62.2	0.19	8.00%

A first intervention is to exogenously improve the technology used by landowners to transport goods. We run the baseline simulation with a 25% improvement to $Z_T(\ell, 0) = 1.875$. An improved transport technology significantly reduces the labour required to supply a given level of transportation demand and so reduces the toll charged for a given demand. Demand for those transportation services increases as more labour is used in manufacturing and agriculture but, by Lemma 2, this does not induce an net increase in transport labour. The consequence of releasing labour into the non-transport sectors is that firms are larger and innovation happens in more places. Takeoff happens in year 1761.7 on average, twenty years earlier than without the intervention (counterfactual 1 in Table 2).

The second counterfactual reduces the cost of transport labour by placing a subsidy of 25% on the wages paid by landowners. Since the lower wages do not directly affect the non-transport sectors, the demand for transportation services is not significantly affected and the share of labour in the transport sector does not change. The consequence of the lower wages is that the tolls charged by landowners decrease marginally but this has no consequence for the average takeoff time. Recall that the toll enters the model in a different way to the physical transport cost, which is lost to the economy.

A more stark demonstration of the impact of transport costs is shown in the third and fourth counterfactuals in which the fixed infrastructure parameter $\varkappa$ in equation (12) is varied by increasing it to $\varkappa = 0.375$ in counterfactual 3 and decreasing to $\varkappa = 0.275$ in counterfactual 4. A higher transport cost generates an average takeoff year that is nearly 10 years earlier than in the baseline because it encourages concentration, particularly in the agriculture sector, that makes it more likely the firms will reach the scale required to overcome the fixed costs of innovation. That faster growth in agriculture means labor moves into manufacturing more quickly, making the date at which manufacturing firms begin to innovate happen earlier. By the same logic, counterfactual 4 shows that a lower transport cost pushes back the date of takeoff considerably. This finding relates with Proposition 3 in DR-H, in which an exogenous increase in transport costs raises aggregate productivity. In this case, a reduction in transport costs reduces the strength of agglomeration, reducing firm size and slightly delaying the year at which manufacturing firms can overcome the fixed costs of innovation. As we will see in Section 6.4, when we conduct counterfactual exercises on the model with endogenous infrastructure supply, the spatial distribution of a change in transport costs and release of transport labour can matter to the impact on takeoff.

Each of these counterfactuals point to the importance of scale for the takeoff in manufacturing. An economy characterized by less labour used in transport, or by more concentrated cities, is more likely to have firms of the size sufficient to overcome the fixed costs of innovation. None of these counterfactual exercises move the location of takeoff (to one decimal point).

4.5 Counterfactual energy prices

The factors that determine the location of takeoff include the distance from the largest market (London), the local initial productivity in manufacturing and the local price of energy. There are three areas of relatively high manufacturing productivity at 1710 that can be seen in Figure 7 – the large agglomeration of London (51.4°), the smaller North (54°) and the Midlands (52.8°) which is smaller again. The local price of energy is determined by both deposits of coal and the cost of transporting it (see Clark and Jacks (2007) and Allen (2009)). In 1710, the price of coal in London was over six times that in the North and over seven times that in Newcastle (55°). The counterfactual exercises in the previous Section show that the location of takeoff does not change when modify the cost of transportation when it only affects the cost of transporting goods. In this subsection, we explore the sensitivity of takeoff location to changes in the spatial distribution of coal prices.

We first impose a uniform drop in coal prices across all locations, keeping other parameters as in the baseline calibration in Table 1. A 56.6% drop in coal prices across all locations is required for the location of industrial takeoff to move. With this decline, the location of the industrial revolution (that location at which manufacturing first begins to innovate) shifts to 51.4° (London) and occurs much earlier than in the baseline simulation (at year 1710). Next, we introduce counterfactuals that change the spatial distribution of coal prices, rather than simply reducing them everywhere. Holding the price in London constant, we ask how much coal prices need to drop in other potential industrial locations for the industrial revolution to happen there. Newcastle is the source of initially very cheap energy. Based on the baseline calibration, even setting the price of energy to zero in Newcastle is not sufficient to induce an industrial takeoff there first. Coal was in reality close to free at times in this region – since it lay so close to the surface of the earth in that region, 'sea coal' could simply be picked up from the beaches of County Durham. Newcastle has low initial manufacturing productivity and is a great distance from London; since the fixed costs to innovation are still positive when energy costs are zero, firms in Newcastle never see a positive gain from investing in innovation. Surprisingly, the same is also true of the Midlands. Even with energy costs at zero in the Midlands, the first innovation in manufacturing still happens in the initially more productive North with its relatively cheap (but not free) energy. In other words, it is not alone the price of coal that matters; the initial advantage that the North had in manufacturing productivity is crucial to explain its place as the geographical cradle of the Industrial Revolution.[19]

5 Model with endogenous fixed infrastructure

Transport costs at a location are a function of the stock of fixed infrastructure, $T(\ell, t)$. An improvement in infrastructure (an increase in $T(\ell, t)$) may be an innovation to an existing mode of transport (such as a better system of locks or a different railway gauge) or else a new mode (such as the steam-powered railway). Such improvements directly reduce the cost of transportation by equation (12), but they also reduce the transport labour required for a given trade demand and so can reduce tolls, by equation (16).

There are a number of problems in making an infrastructure improvement endogenous. First, since the infrastructure stock is persistent, the infrastructure investment decision may be a dynamic one. Second, investment must take place over a connected interval of locations, since individual landowners are of measure zero. Third, the success of an infrastructure investment in one location may depend on the presence (or absence) of an infrastructure investment in another location.

We use the historical record of fixed infrastructure development in 18th and 19th century England and Wales, described in Section 2.3, to motivate a number of simplifying

[19] This outcome is partly the result of the adjustment for the breadth of the country in the transformation to the one-dimensional interval. Compared to the North, there is a larger total quantity of secondary labour employed at the latitude of the Midlands (as can be seen in larger total population in Figure 20). When we use unadjusted data and invert the model to obtain productivities, this would be interpreted as a higher productivity of the secondary sector there. When we transform the data as described in Section 4.1, the Midlands has less labour in the secondary sector and thus has lower productivity.

assumptions that lend tractability to our model. First, we distinguish between the upfront costs raised to cover the construction of an infrastructure improvement and the ongoing costs of maintenance and carriage. Second, since we saw that infrastructure investment took place when joint-stock companies, formed by Act of Parliament, raised finance locally, we incorporate local finance in a simple way by dividing the interval of the economy into a number of 'counties'. Finally, we incorporate 'Parliamentary co-ordination' as part of our solution concept where equilibrium investments are not unique.

5.1 Infrastructure investment

Let $\mathcal{L} \subset [0,1]$ be a connected interval of locations with measure $m(\mathcal{L}) > 0$. An infrastructure investment over interval $\mathcal{L}$ at time t permanently increases $T(\ell, t)$ for all $\ell \in \mathcal{L}$ by an amount $\delta(t) > 0$ at a cost $x^T \cdot m(\mathcal{L})$, where $x^T > 0$ is a constant.[20]

A group of connected landowners along such an interval is thus required for an infrastructure improvement. Since individual landowners are of measure zero, the potential for free-riding means that equilibrium investment in infrastructure is zero. We suppose that an 'infrastructure company' can be established at cost $c > 0$. An infrastructure company can lease land from a group of landowners for one period[21] and then collects rent from firms using the leased land to produce agricultural and manufactured goods. In return for the lease, the infrastructure company pays a dividend to landowners and invests in infrastructure.

Infrastructure companies are competitive in the sense that any landowner in an interval can propose to form a company by offering a prospectus which defines rents to be collected from firms, the infrastructure investment and the dividend payment. Landowners select that prospectus which offers the highest dividend.

We simplify the spatial nature of infrastructure investment by supposing that an infrastructure company can be formed only at the level of a 'county'. A county is one of $N \in \mathbb{Z}^+$ connected intervals of equal measure that span the complete unit interval of locations in the economy. In particular, county $i \in [1, \ldots, N]$ is the connected interval $\Delta(i) = \left[\frac{i-1}{N}, \frac{i}{N}\right]$. The cost to improving infrastructure in a county is thus x^T/N. Only if the company leases all land in the county can the infrastructure improvement take place.

The optimization problem of an infrastructure company in county i at time t is to

[20] The infrastructure gain $\delta(t)$ grows over time as transport technologies improve. Bogart (2014) calculates that TFP growth in the transport sector was around 2% over the period of study. As such, we let $\delta(t) = (1.02)^t \delta(0)$ where $\delta(0)$ is an initial level of transport technology. That the cost x^T is fixed means that we abstract from any interdependence in the cost of projecting infrastructure at different locations; see Swisher (2015).

[21] The restricted duration of the land lease is clearly a simplification of reality. However, the historical evidence discussed above supports the idea that initial capital was raised primarily for construction costs; we are here compressing the raising of funds and construction of infrastructure into one year in the model. Moreover, as we shall see, despite restricting the investment decision to a one-period problem, the scale of total infrastructure investment in the model is quantitatively close to that found on aggregate in the data.

maximize its profit, $\pi^I(\ell, t)$,

$$\max_{R^F,\mathcal{D},\mathcal{I}} \pi^I(i,t) = \int_{\ell\in\Delta(i)} R^F(\ell,t)\mathrm{d}\ell - \mathcal{D}(i,t) - \mathcal{I}(i,t)\left[\frac{x^T}{N} + c\right], \tag{22}$$

where $R^F(\ell, t)$ is the rent collected from a firm at location ℓ, conditional on infrastructure investment, $\mathcal{D}(i,t)$ is the dividend paid to landowners and $\mathcal{I}(i,t)$ is an indicator function equal to 1 if there is an infrastructure investment in county i at time t, and 0 otherwise.

We proceed as follows: First, we note that an infrastructure company's problem involves maximizing current period dividends only. Second, we show that if an infrastructure investment increases total rent in a county (net of costs), then the company can commit to building infrastructure. Finally, we establish the nature of economy-wide (i.e., multi-county) equilibria and specify the selection criteria in the case of multiple equilibria.

Remark 1 *If landowners in a county lease their land, it is to the company whose prospectus offers the highest dividend. Since fixed infrastructure is persistent, a company could in principle take account of a future higher stream of rental income and pay a higher stream of future dividends. However, each lease is for one period only and there is free access to the improved infrastructure in subsequent periods. The bidding to form a company in period* $t+1$ *is the same for all potential companies. With free entry, landowners thus bid to form an infrastructure company by maximizing* $\mathcal{D}(i,t)$ *only.*

Lemma 4 *Maximizing current dividends means that infrastructure companies make zero net profit, and that individual firms pay the maximum land rent conditional on infrastructure investment. An infrastructure investment is optimal for the infrastructure company when it increases total current rent, net of costs, within a county.*

Proof. The prospectus that maximises the dividend to the landowners forms the infrastructure company which leases the land. Since there is free entry to offering a prospectus, companies make zero profit in equilibrium and the dividend paid to landowners in a county i is,

$$\mathcal{D}(i,t) = \int_{\ell\in\Delta(i)} R^F(\ell,t)\mathrm{d}\ell - \mathcal{I}(i,t)\left[\frac{x^T}{N} + c\right]. \tag{23}$$

The dividend paid by the infrastructure company is determined by the rent collected from firms and the choice over infrastructure investment. Suppose that a prospectus in county i proposes a total dividend $\mathcal{D}_1(i,t)$ based on $R_1^F(\ell,t)$ rents paid by firms. If higher rent $R_2^F(\ell,t)$ could be collected by a company then, by equation (23), $\mathcal{D}_2(i,t) > \mathcal{D}_1(i,t)$. Landowners choose the prospectus that delivers the highest dividend, so land is allocated by the infrastructure company to the firms that can pay the highest rent at each location. The problem for producer-firms is thus identical to that in Section 3.

Let $\hat{R}^F(i,t)$ be the maximum sum of firm rents collected by a company that selects $\mathcal{I}(i,t) = 1$ and $\tilde{R}^F(i,t)$ be the rents if $\mathcal{I}(i,t) = 0$. If $\tilde{R}^F(i,t) > \hat{R}^F(i,t) - \frac{x^T}{N} - c$

then, by equation (23), dividends are maximised where infrastructure investment does not take place. For the infrastructure company to be able to commit to an infrastructure investment, it must be that it increases total current rent net of costs within the county. ■

Such an infrastructure company does not always exist in each county in each period, but can exist where the increase in the rents paid by firms is sufficient to amortize the fixed costs of infrastructure.

Finally, the profitability of an infrastructure investment in one county can depend on the presence (or absence) of investment in another county. We focus on pure-strategy Nash equilibria where each county's binary infrastructure investment is an optimal response to each other county's investments. We first define equilibrium infrastructure investment. Let each county's investment decision be $s_i \equiv \mathcal{I}(i, t) \in \{0, 1\}$, where we drop the time subscript for notational convenience. The payoff to each county, $u_i(s_i|s_{-i})$, is the net rent collected from firms, conditional on the strategies s_{-i} of other counties. The following defines a Nash equilibrium in this context and specifies the selection criteria in the case where we have multiple, or no, such equilibria.

Definition 1 *An equilibrium in infrastructure investment is a strategy profile s^* such that for each i and all $s_i \in \{0, 1\}$, $u_i(s_i^*|s_{-i}^*) \geq u_i(s_i|s_{-i}^*)$. Where s^* is not unique, Parliamentary co-ordination exists to select that profile $s^{**} \in s^*$ which maximises total net rent in the economy. Where s^* is empty, Parliamentary co-ordination imposes $s_i = 0$ for all i.*

This equilibrium concept uses the historical evidence, described in Section 2.3, on the role of Parliament to motivate the reconciliation of alternative proposals where there are multiple Nash equilibria. Since we do not allow for mixed strategies, it may be that a Nash equilibrium will not exist. In this case, we impose that no investment takes place in any county. Appendix C details the computation of equilibrium. Calculating the Nash equilibria can be computationally intensive since individual county investments must be calculated in all permutations of other county investments. That is, for an interval split into N counties, there are up to $N \cdot 2^{N-1}$ simulations to run for each period of each simulation. We select $N = 10$ in the quantitative analysis below and limit the number of simulations to 20 instead of the previous 100.[22]

6 Quantitative analysis

We use the model as extended in Section 5 to capture three aspects of the spatial takeoff of England and Wales during the 18th and 19th centuries in addition to those macroe-

[22]While there are optimizations that reduce the number of repeated simulations required, even with $N = 10$ the full model with endogenous infrastructure takes between seven and eight hours on a server with 32 cores.

conomic and spatial changes matched in Section 4. Specifically: The overall decline in transport costs, the timing of the transport revolution, and, the spatial distribution of the investments in infrastructure.

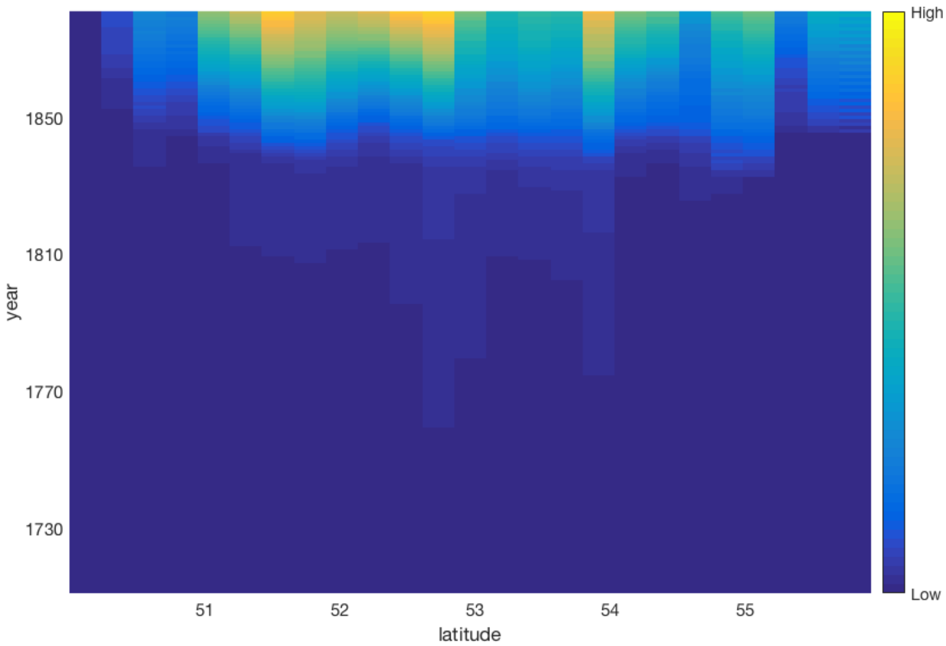
6.1 Interval representation of the transport revolution

We first map data on the spatial distribution of transport infrastructure into the one-dimensional interval of the model. We consider the three major modes of inland transportation (road, waterway and rail) over the period 1710-1881. We use the turnpike network at 1830 and waterways (navigable rivers and canals) at 1670 and 1817 and a dynamic GIS of the railway network over the period from 1830. To map the two-dimensional information into a one-dimensional interval we sum mileage along the East-West axis and adjust for width, as with the transformation of employment. To interpolate turnpikes and waterways at decadal intervals over the period we use information on road and canal acts per county for the period 1701 to 1830 (Jackman, 1962; Appendix 13).

In order to combine the three modes into one map of infrastructure development, we weight each mode to form a combined picture of the spatial infrastructure development. Compared to a turnpike, a canal project would have been more directly useful in haulage (see Bagwell, 1974; Turnbull, 1987; Bogart, 2014). Though measures of the volume of goods traded by road are not available (see Bagwell, p.58), Bogart (2014) suggests a horse-drawn wagon may carry as much as seven tons by 1800; a horse-drawn barge could be loaded with ten times that much. We thus weight a canal mile as equivalent to ten turnpike miles. Railways, like canals, can handle heavy loads and, with the introduction of steam power, do so at much greater speeds than horse-drawn barges. We can thus weight railways against canals based on their relative speeds. Bogart (2014) puts the speed of transporting passengers on 18th-century stagecoaches at around one eighth of 19th century trains. Our focus is on the transport of output (i.e., freight), for which the relative advantage of rail may be smaller. Taylor (1951) presents evidence for early 19th-century U.S. that puts the speed of transporting freight on roads and canals at around one fifth that on rail (2 vs. 10-12 m.p.h). Our measure of combined transport development thus weights a canal mile as having ten times the contribution of a turnpike mile and a railway mile as having five times the contribution of a canal mile.

Figure 16 presents the combined spatial infrastructure data mapped into twenty five equally-sized counties. Some improvements occur in the 18th century in a somewhat disjointed fashion, separately around 52.8° and 54° (that is, in the Midlands and in the North). A merging into a more connected infrastructure network begins later into the 19th century with the arrival of the railway which, by 1850, represents the historically significant change in supply of fixed infrastructure. While this period has more infrastructure in all locations (except the most Southern county), clear in the Figure is the persistent concentration in infrastructure improvement around London (51.4°), the Midlands (52.8°) and the North (54°). This one-dimensional representation of a patchwork development mirrors that discussed in the literature (see Section 2).

Figure 16: Infrastructure level (data)



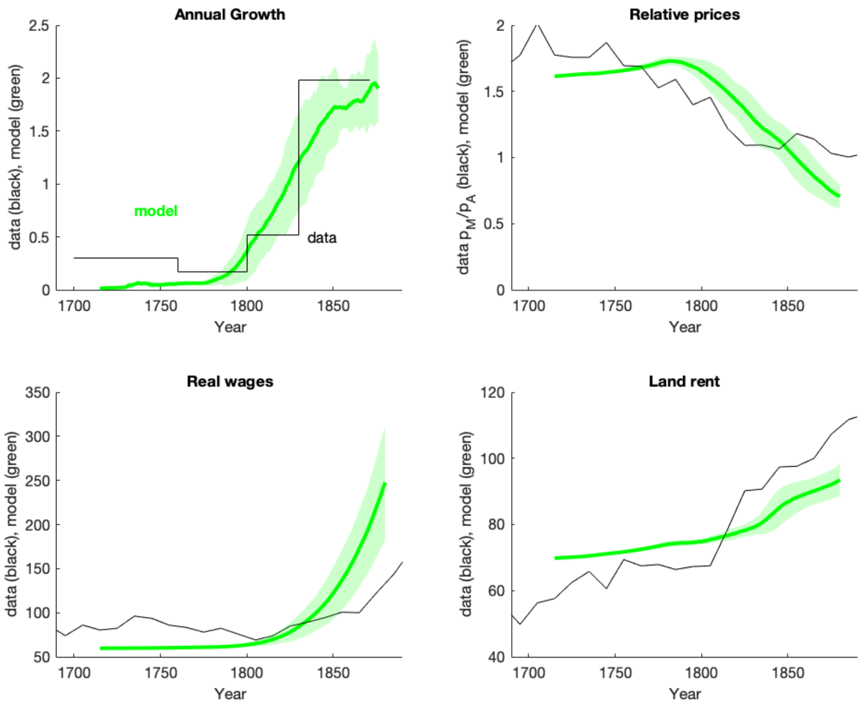
6.2 Parameterization

All baseline parameters remain as in Table 1. Again, we initialise the model at 1710 with a distribution of infrastructure development that reflects the initial distribution of access to transport. There are two additional parameters to pin-down: The rate of growth of technology in the transport sector and the fixed costs of investment in physical infrastructures. For technology growth, we use the estimate in Bogart (2014) of approximately 2% per year. The fixed costs of constructing transport infrastructure were significant, with large quantities of skilled and unskilled labour required in addition to the materials, purchase of land, legal fees and fees for agents to navigate the passage of required Acts of Parliament. Such costs applied just as much to turnpikes as to canals and railways (see Jackman (1962, p. 236) and Bagwell (1974, p.99). To parameterize the fixed costs of infrastructure improvement we use data in Pollins (1952) on nominal construction costs for early railways with data in Mitchell (1988) for nominal GDP. The average railway construction cost is 0.39% of nominal GDP. Using the baseline calibration, this equates to a per-county fixed cost of $x^T/N = 1.15$.

6.3 Simulation output

We run the full simulation with endogenous infrastructure 20 times and present results with a confidence interval as before. The year of takeoff is on average 1789, four years later than the baseline without endogenous infrastructure. The fit against macroeconomic variables remains close to the data, as can be seen in Figure 17. Figure 18 shows that the pattern of aggregate structural transformation in the model is now improved. As output and trade increase, so do the gains from investing in infrastructures that save on tertiary labour. Without endogenous infrastructure investments, the model predicts too much growth in the relative importance of the tertiary sector and too little decline in the primary sector. Where infrastructure investment can be improved, the structural transformation from the primary towards the tertiary sector is more modest.

Figure 17: Macroeconomic Variables
(Average of 20 simulations ± 2s.d.)



We report in Appendix Figures 23-28 the fit of the spatial distribution of the level of employment in each sector, as well as the shares. There is some improvement in the fit of the tertiary shares in the model. The most interesting additional component of the simulation output concerns the timing and intensity of the endogenous infrastructure development. Figure 19 depicts the infrastructure level in the model as it occurs over space and time in the 10 counties of the model.

As can be seen in Figure 19, the model generates infrastructure development that is consistent with the data in many respects. The acceleration in the supply of infrastructure

Figure 18: Labour shares
(Average of 20 simulations ± 2s.d.)

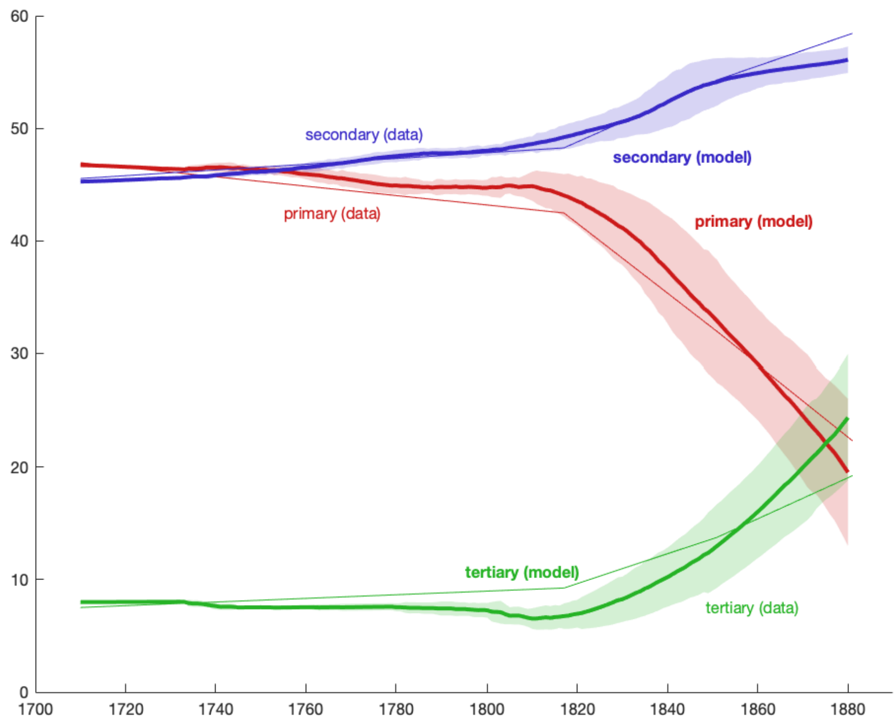
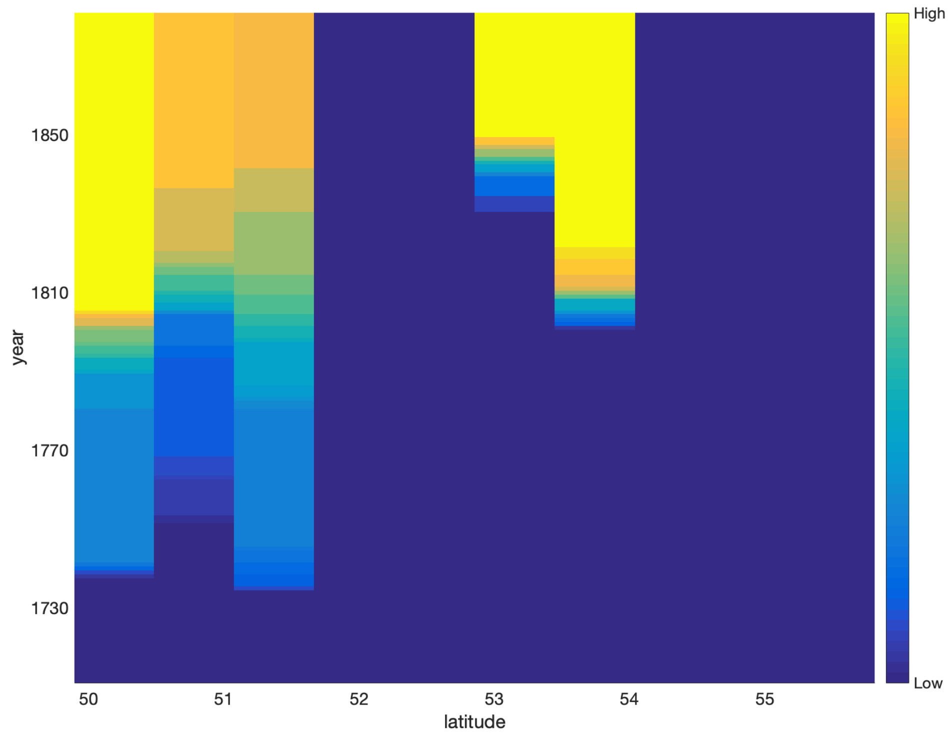


Figure 19: Infrastructure level
(Average of 20 simulations)



begins early in some locations and accelerates around 1850 across a broad geography. The disjointed nature of infrastructure development seen in the data is also replicated in the model, in particular capturing the industrial hotspot around 53.8°. The omission of the infrastructure investment around 52.5° is a result of the model not capturing the (relatively small) secondary concentration that emerges there.

One stark difference is the model prediction for significant infrastructure investment close to 50°. This points to a limitation of the mapping into the one-dimensional interval – in the model, the county close to 50° is directly adjacent to London; in reality, that portion of England (which is South Devon and Cornwall) is, as clear in Figure 6, actually very distant from it. The model thus predicts significant opportunities to gain from infrastructure in that location which in reality do not exist. An additional aspect which does not quit fit is the breadth of the improvements seen after at the end of the time period. We use the data in Bogart (2014) on TFP growth in transport to parameterize the model. This steady increase in technology is most likely not accurate; the later arrival of steam and railway technologies would have accelerated the attractiveness of improvements in infrastructure into the 18th century.

The model predicts an average decline in total transport costs over the simulation period of 19.3% which is significantly less than that found in the data (86%). When we weight the transport costs decline by trade volume (that is, the how much actual output and trade is lost to transport costs), the decline is 50.0%. If we imposed an acceleration in the growth rate of infrastructure technology discussed above, the decline in transport costs by the end of the sample period would be greater. Moreover, when we add the costs of trans-shipment, as in Alvarez-Palau (2018), the real decline in freight costs will be lower than 86%.

The fit of the model to the location and timing of the spatial takeoff is improved, since now the infrastructure investment around the Northern hotspot means that the extent of tertiary labour growth seen in Figure 14 is no longer as far away from the data.

Finally, since the model imposes that infrastructure investments take account of only a one-period gain, we may be concerned that the model significantly under-predicts total investment. We can compare the aggregate spend on infrastructure with that in the data. As Bogart (2014) summarizes, turnpike and canal investment was around £27 million by 1820, which is 7.3% of 1820 nominal GDP (based on the output estimates in Broadberry et al., 2015). Railway investment by 1870 totalled £232 million, which is 22.6% of 1870 GDP. In the model, total spend over the whole period on infrastructure is on average 5.5% of 1870 output. That is, the model captures the scale of the early investments in infrastructure, but not the later improvements. As with the lower simulated decline in transport costs, this can be partly explained by the constant growth in transport TFP that we take from Bogart (2014); the model neglects an unmeasured acceleration of transportation technology improvements in the nineteenth century. In addition, the spatial unit for infrastructure improvement, the 'county', is fixed over time in the model. As a result, we also miss the growing geographical scale of the projects and the increasing levels of finance required to fund them. As the geographical scale of individual infrastructure improvements grew, the problems of co-ordination and free-riding increased, the local financing model became strained and so a greater role for the public sector emerged. We would thus expect the model to somewhat under-predict the level of later investments.

6.4 Counterfactual infrastructure policies

As described in Section 2, the development of transport infrastructure in England and Wales during the industrial revolution was largely driven by private and local finance in the context of a laissez faire regulator. The model of endogenous infrastructure development introduced above captures aspects of that history. Alternative policy treatments of infrastructure development exist, however, both in history (as noted in France) and in modern infrastructure development. We use our model to ask in this Section whether counterfactual infrastructure plans may have brought forward or delayed the advent of the industrial takeoff.

Specifically, we consider counterfactuals that modify the timing and spatial distribution of the infrastructural development generated endogenously by the model. Our baseline is the average outturn of endogenous infrastructure development depicted in Figure 19. Table 3 reports statistics based on 100 runs of the baseline and each counterfactual. We report the year and latitude of takeoff (that is, the year and latitude in which manufacturing firms first begin to innovate), the number of years until 1.5% growth, the initial average transport cost and the initial labour share in the transport sector.

Table 3: Infrastructure Baseline and Counterfactuals (averages of 100 runs)

	takeoff year	takeoff latitude	years to 1.5%	initial av. κ	initial L_T%
Baseline	1782.7	53.7°	57.3	0.20	8.00%
1. Early (50)	1777.6	53.7°	53.7	0.17	7.40%
2. Early (100)	1769.2	53.7°	59.8	0.13	6.25%
3. Late (50)	1782.5	53.7°	59.5	0.20	8.00%
4. Late (100)	1782.0	53.7°	57.0	0.20	8.00%
5. Uniform baseline	1783.8	53.7°	62.2	0.20	8.00%
6. Uniform early (50)	1787.3	53.7°	–	0.17	7.49%

The first counterfactuals bring forward the baseline infrastructure development by 50 and 100 years (counterfactuals 1 and 2, respectively). In the previous counterfactual exercises in Section 4.4, we found that exogenous variations in the transport cost were inversely related with the year of takeoff. This was consistent with Proposition 3 in DR-H, which found that higher transport costs raise productivity. In the counterfactuals here, we find that earlier infrastructure development accelerates takeoff. Where infrastructure development is brought forward by 50 years, takeoff occurs five years earlier; where development occurs 100 years early, takeoff is accelerated by over thirteen years. In each case, bringing forward the infrastructure development means that transport costs are on average lower at the initial date. As the previous counterfactuals showed, this would tend to delay the takeoff since it can reduce agglomeration incentives and set back the date at which manufacturing firms are large enough to meet the fixed costs of innovation. However, since transport labour is here endogenous to the level of infrastructure supply, bringing forward the advances in infrastructure also releases labour into manufacturing and agriculture, making firms larger and bringing forward the point at which manufacturing firms can amortize the fixed costs of innovation. This suggests a tension between the role of infrastructure in hastening takeoff by reducing the labour required in transport, and its impact on delaying takeoff by reducing transport costs. Based on the spatial pattern of infrastructure development in Figure 19, it is clear that the gain from the lower labour in transport dominates the loss from the lower transport costs.

Counterfactuals 3 and 4 delay the infrastructure rollout by 50 and 100 years, respectively. Given the timing of endogenous development depicted in Figure 19, this delay is equivalent to there being no infrastructure development at all prior to takeoff (i.e., equivalent to the baseline in Section 4.3). The initial transport cost and initial labour share are unchanged and the takeoff date is not significantly different from the baseline.

One finding that is consistent with the previous counterfactuals is the lack of any impact on the location of takeoff. The high initial productivity in the North of England, together with the already low energy costs there, make it the location of takeoff in all counterfactuals. We showed in Section 4.5 that substantial changes in energy prices can result in a shift of the location of the first innovation by manufacturing firms to London. This occurs after energy prices fall there by over 56%. Given the locations of coal fields, only a significant, nationally-interconnected improvement to inland transportation would generate this decline. Coastal shipping along the Eastern coast of England did emerge to transport coal from Newcastle to London (Davis, 1962), but it did so at slow speeds and high costs (the high initial coal price difference between Newcastle and London is reflects the coastal shipping technology at 1710). The large costs of building infrastructure meant that such a national project did not emerge endogenously and the industrial hotspot proceeded in the North.

None of the counterfactuals thus far consider the role of the *spatial heterogeneity* of the infrastructure improvements. We explore this by taking take the same total growth in the level of infrastructure, but apply it broadly rather than locally within a county. In particular, we use the same initial distribution of transport costs and introduce the endogenous improvements in Figure 19 in a uniform way by averaging each period's total infrastructure investment across all locations. Counterfactual 5 shows that making the endogenous improvements uniform does not affect the timing of development. Since the initial distribution of transport costs and initial labour used in transportation is unchanged by making the later development of infrastructure uniform, and since that later infrastructure development is partly a result of that takeoff, then spreading out the infrastructure development without changing its timing has no clear impact on the year of the takeoff.

What does make a difference is to bring forward the uniformly distributed infrastructure improvements. Counterfactuals 1 and 2 showed that bringing forward the spatially-heterogeneous improvements in infrastructure can substantially accelerate the timing of the industrial takeoff. In those cases, the impact of lower labour in transport dominated the effect of the decline in agglomeration forces. Counterfactual 6 introduces infrastructure development in a uniform way but brought forward by 50 years. This is similar to counterfactual 4 in Table 2, where lower transport costs delayed takeoff. As can be seen in Table 3, bringing forward uniform development of infrastructure by 50 years means a takeoff over four years later than the baseline. The pace of acceleration in growth is also substantially slower, not reaching 1.5% on average in the period simulated (up to 1881). These outcomes are clearly different to bringing forward the infrastructure development with the pattern defined by Figure 19, as Table 3 shows. The initial average transport costs and the initial share of labour in transportation are each lower than the baseline, as in counterfactuals 1 and 2, but the spatial distribution of those changes are different.

The spatial distribution of the infrastructure gains thus clearly matters to whether the acceleration induced by lower transport labour dominates the delay caused by lower transport costs. Where the infrastructure improvements are localized, those particular areas benefit the most in terms of lower transport labour demand while the cost of transportation outwith those areas remains high. Where the better infrastructure is uniformly distributed, in contrast, transport costs are lowered in all locations and this has the effect of significantly muting the gains from releasing labour out of the provision of transportation services.

This set of counterfactual results suggest that only policies targeted at localized infrastructure improvements may generate significant returns in terms of inducing earlier innovation. Bringing forward localized infrastructure development can significantly accelerate the date of takeoff. Bringing forward uniform improvements does nothing to accelerate, and may delay, the onset of innovation despite costing the same in terms of infrastructure improvements. The problem for such a localized policy is identifying those areas where infrastructure should be improved, potentially before the greatly increased demand for transportation services even emerges. The one clear implication is that the gains from a controlled, uniform improvement to infrastructure most likely do not outweigh their high costs.

7 Concluding Remarks

We have developed a model of endogenous infrastructure development in the context of a spatial growth framework and initialized it to England and Wales in 1710. The model can explain aspects of the evolution of the economy over the subsequent 171 years. In particular, we capture the structural transformation across three sectors, the acceleration in growth and the changing spatial distribution of activity in agriculture, industry and transportation. Where the early infrastructure investments were made at a local level, the model captures the magnitude of total investment. As the scale of individual infrastructure projects grew toward the end of the period, the model under-predicts total investment. Finally, we have seen how changes in the timing and spatial distribution of infrastructural development can have an effect on the timing of the acceleration in growth that accompanies industrialization.

We leave a number of extensions to be considered in future work. First, the Shaw-Taylor et al. (2010a) data shows that there was specialization within industrial subsectors as well as between agriculture and industry. A model of multiple secondary subsectors that each specialize spatially may improve the fit of the model. Second, one aspect of the private evolution of infrastructure during the industrial revolution in England and Wales is that it grew in scale over time. A model that captures the expanding scale of finance and growing scope of infrastructure may better explain the timing of the infrastructure development seen in the data. Related to that is the changing industrial organization of infrastructure companies, from small local concerns to national private companies to eventual public ownership. Third, by the late 20th century much of the railway network was considered to be inefficiently organized and in over-supply. In the 1960s, the government implemented dramatic reductions in station numbers and mileage. The long-run consequence of the laissez faire regulation of the early infrastructure development has yet to be understood. Fourth, for tractability, infrastructure investments were assumed to take place within one period. Incorporating a more realistic endogenous investment that can spread the fixed costs over multiple periods may be important. Finally, while we have shown that there is a role for policy in accelerating development, we have not addressed the question of how an infrastructure policy that is optimal for growth could be designed.

References

Adamopoulos, T. (2011). 'Transportation costs, agricultural productivity and cross-country income differences'. *International Economic Review*, 52(2):489–521.

Allen, R. C. (2004). 'Agriculture during the industrial revolution, 1700–1850'. Chapter 1 in Floud, R. and Johnson, P. (eds.). *The Cambridge Economic History of Modern Britain: Volume 1, Industrialisation 1700-1860.* Cambridge University Press.

Allen, R. C. (2009). *The British Industrial Revolution in Global Perspective*. Cambridge University Press.

Allen, T. and Arkolakis, C. (2014). 'Trade and the topography of the spatial economy'. *Quarterly Journal of Economics*, 129(3):1085–1140.

Alvarez-Palau, E. J. (2018). 'Multi-modal Models of the Transport Network in England and Wales, 1680-1830-1911'. European Social Science History Conference.

Asturias, J. and Petty, S. (2013). 'Endogenous Transportation Costs'.

Bagwell, P. S. (1974). *The Transport Revolution from 1770*. B. T. Batsford Ltd. Manchester University Press.

Bogart, D. (2007). 'Neighbors, Networks, and the Development of Transport Systems: Explaining the Diffusion of Turnpike Trusts in Eighteenth-Century England'. *Journal of Urban Economics*, 61:238–362.

Bogart, D. (2014). 'The Transport Revolution in Industrializing Britain: A Survey'. Chapter 13 in Floud, R., Humphries, J. and Johnson, P. (eds.). *The Cambridge Economic History of Modern Britain: Volume 1, Industrialisation 1700-1860*. Cambridge University Press.

Broadberry, S., Campbell, B. M. S., Klein, A., Overton, M., and van Leeuwen, B. (2015). *British Economic Growth, 1270–1870*. Cambridge University Press.

Broadbridge, S. (1955). 'The Early Capital Market: The Lancashire and Yorkshire Railway'. *The Economic History Review*, 8(2):200–12.

Buchanan, B. J. (1986). 'The Evolution of the English Turnpike Trusts: Lessons from a Case Study'. *The Economic History Review*, 39(2):223–43.

Clark, G. (2002). 'The Agricultural Revolution and the Industrial Revolution, England 1500–1912'. Available at http://www.econ.ucdavis.edu/faculty/gclark/data.html.

Clark, G. and Jacks, D. (2007). 'Coal and the Industrial Revolution, 1700–1869'. *European Review of Economic History*, 11(1):39–72.

Clarke, M. (1994). *The Leeds & Liverpool Canal*. Carnegie Publishing.

Crafts, N. F. R. and Harley, C. K. (1992). 'Output Growth and the British Industrial Revolution: A Restatement of the Crafts-Harley View'. *The Economic History Review*, 45(4):703–30.

Davis, R. (1962). *The Rise of the English Shipping Industry*. MacMillan.

Desmet, K., Greif, A., and Parente, S. L. (2015). 'Spatial Competition, Innovation and Institutions: The Industrial Revolution and the Great Divergence'. mimeo.

Desmet, K. and Henderson, J. V. (2015). 'The Geography of Development Within Countries'. Chapter 22 in Duranton, G., J.V. Henderson and W. Strange. *Handbook of Regional and Urban Economics*, Elsevier.

Desmet, K. and Rossi-Hansberg, E. (2012). 'Innovation in Space'. *American Economic Review: Papers and Proceedings*, 102(3):447–452.

Desmet, K. and Rossi-Hansberg, E. (2014). 'Spatial Development'. *American Economic Review*, 104(4):1211–43.

Donaldson, D. (2015). 'Railroads of the Raj: Estimating the Impact of Transportation Infrastructure'. *American Economic Review*, forthcoming.

Fajgelbaum, P. and Schaal, E. (2017). 'Optimal Transport Networks in Spatial Equilibrium'. NBER Working Paper Series No. 23200.

Gollin, D. and Rogerson, R. (2014). 'Productivity, Transport Costs, and Subsistence Agriculture'. *Journal of Development Economics*, 107:38–48.

Hadfield, C. (1981). *The Canal Age.* David & Charles, 2nd edition.

Harris, R. (2000). *Industrializing English Law: Entrepreneurship and Business Organization, 1720-1844.* Cambridge University Press.

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). 'Growth and Structural Transformation'. chapter in. *Handbook of Economic Growth.*

Herrendorf, B., Schmitz, J., and Teixeira, A. (2012). 'The Role of Transportation in U.S. Economic Development'. *International Economic Review*, 53(3):693–715.

Heston, A., Summers, R., and Aten, B. (2011). 'Penn World Table Version 7.0'. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hunt, B. C. (1935). 'The Joint-Stock Company in England, 1830-1844'. *Journal of Political Economy*, 43(3):331–64.

Jackman, W. T. (1962). *The Development of Transportation in Modern England.* Frank Cass & Co.

Kleinert, J. and Spies, J. (2011). 'Endogenous Transport Costs in International Trade'. IAW Discussion Paper No.74.

Lévy-Leboyer, M. (1978). 'Capital Investment and Economic Growth in France, 1820–1930', in Mathias, P. and Postan, M. M. (eds.). *The Cambridge History of Europe. Volume VII(1).* Cambridge University Press.

Milward, A. S. and Saul, S. B. (1973). *The Economic Development of Europe, 1780-1870.* George Allen & Unwin Ltd.

Mitchell, B. R. (1988). *British Historical Statistics.* Cambridge University Press.

Mokyr, J. (2004). 'Accounting for the industrial revolution'. Chapter 1 in Floud, R. and Johnson, P. (eds.). *The Cambridge Economic History of Modern Britain: Volume 1, Industrialisation 1700-1860.* Cambridge University Press.

Nagy, D. K. (2015). 'City location and economic development'. mimeo.

Pollins, H. (1952). 'A Note on Railway Constructional Costs 1825–1850'. *Economica*, 19(76):395–407.

Pollins, H. (1954). 'The Marketing of Railway Shares in the First Half of the Nineteenth Century'. *The Economic History Review*, 7(2):230–9.

Redding, S. J. and Turner, M. (2015). 'Transportation Costs and the Spatial Organization of Economic Activity'. Chapter 20 in Henderson, J. V. and Strange, W. (eds.). *Handbook of Regional and Urban Economics.*

Rodrigue, J. (2017). *The Geograhpy of Transport Systems.* Routledge.

Rossi-Hansberg, E. (2005). 'A Spatial Theory of Trade'. *American Economic Review*, 95:1464–91.

Shaw-Taylor, L., Davies, R. S., Kitson, P. M., Newton, G., Satchell, A. E. M., and Wrigley, E. A. (2010a). 'The occupational structure of England and Wales c.1817-1881'. Work in Progress. For details of the wider project see http://www.geog.cam.ac.uk/research/projects/occupations/.

Shaw-Taylor, L. and Wrigley, E. A. (2014). 'Occupational structural and population change'. Chapter 2 in Floud, R., Humphries, J. and Johnson, P. (eds.). *The Cambridge Economic History of Modern Britain: Volume 1, Industrialisation 1700-1860.* Cambridge University Press.

Shaw-Taylor, L., Wrigley, E. A., Davies, R. S., Kitson, P. M., Newton, G., and Satchell, A. E. M. (2010b). 'The occupational structure of England c.1710 to c.1871'. Work in Progress. For details of the wider project see http://www.geog.cam.ac.uk/research/projects/occupations/.

Shea, G. S. (2012). 'The Handbook of 18th and Early 19th Century British Corporate Finance'. University of St. Andrews, mimeo.

Smith, Jr., C. O. S. (1990). 'The Longest Run: Public Engineers and Planning in France'. *The American Historical Review*, 95(3):657–92.

Swisher, IV, S. N. (2015). 'Reassessing Railroads and Growth: Accounting for Transport Network Endogeneity'. mimeo.

Szostak, R. (1991). *The Role of Transportation in the Industrial Revolution.* Montreal.

Taylor, G. R. (1951). *The Transportation Revolution, 1815–60.* Routledge.

Trew, A. (2010). 'Infrastructure Finance and Industrial Takeoff in England'. *Journal of Money, Credit and Banking*, 42(6):985–1010.

Trew, A. (2014). 'Spatial Takeoff in the First Industrial Revolution'. *Review of Economic Dynamics*, 17:707–25.

Turnbull, G. (1987). 'Canals, Coal and Regional Growth during the Industrial Revolution'. *The Economic History Review*, 40(4):537–60.

Valentinyi, A. and Herrendorf, B. (2008). 'Measuring factor income shares at the sectoral level'. *Review of Economic Dynamics*, 11:820–35.

Ward, J. R. (1974). *The Finance of Canal Building in Eighteenth Century England.* Oxford University Press.

World Bank (2011). *Transformation Through Infrastructure: Infrastructure Strategy Update FY2012–15*. World Bank Group.

Wrigley, E. A. (2010). 'The PST System of Classifying Occupations'. Paper 1 at www.hpss.geog.cam.ac.uk/research/projects/occupations/britain19c/papers/.

Yang, D. T. and Zhu, X. (2013). 'Modernization of Agriculture and Long-Term Growth'. *Journal of Monetary Economics*, 60(3):367–82.

A Additional figures

Figure 20: Interval distribution of total employed (normalized, unweighted)

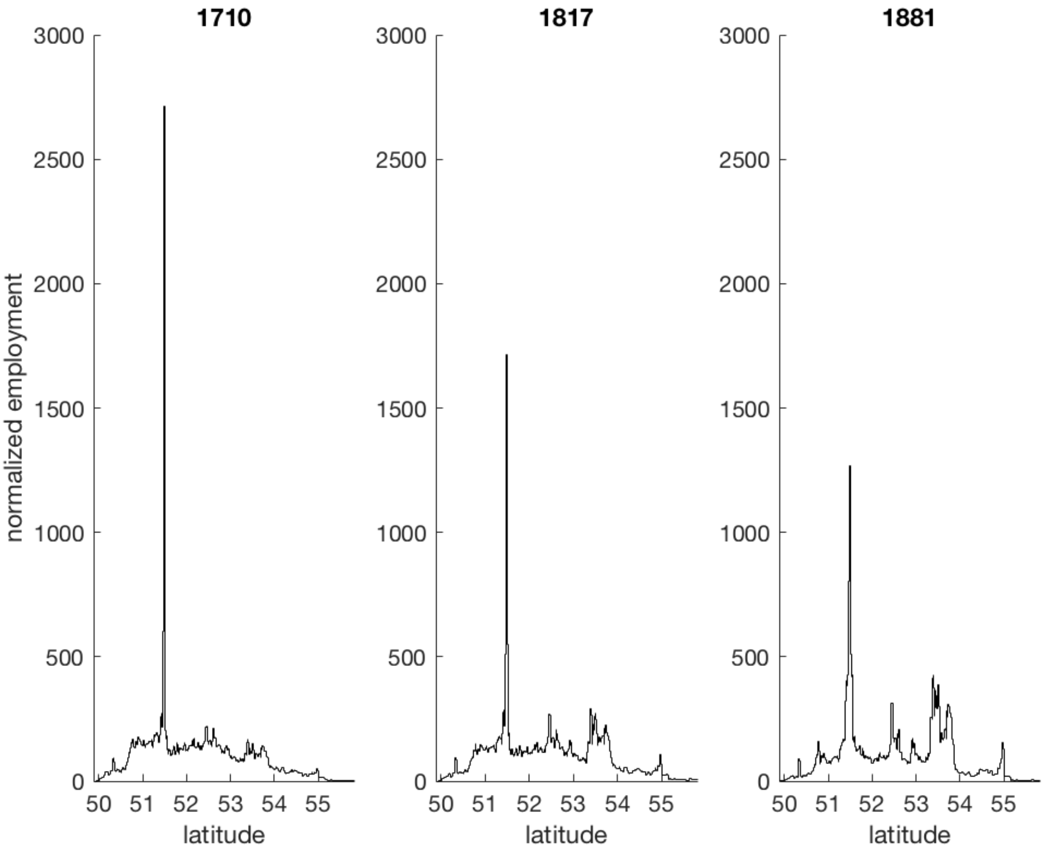


Figure 21: Baseline Simulation: Primary Shares
(Average of 100 simulations $\pm$ 2s.d.)

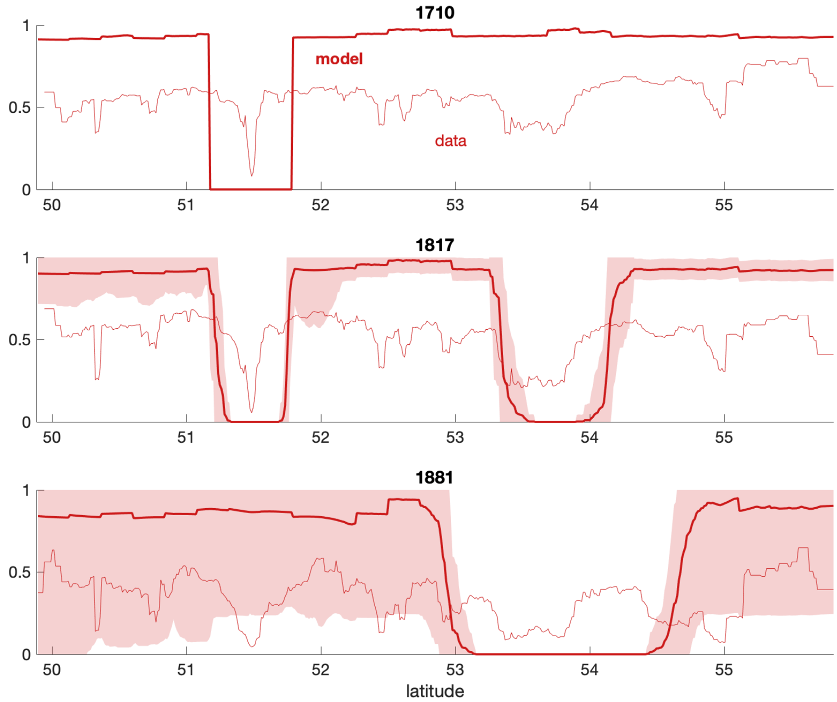


Figure 22: Baseline Simulation: Secondary Shares
(Average of 100 simulations $\pm$ 2s.d.)

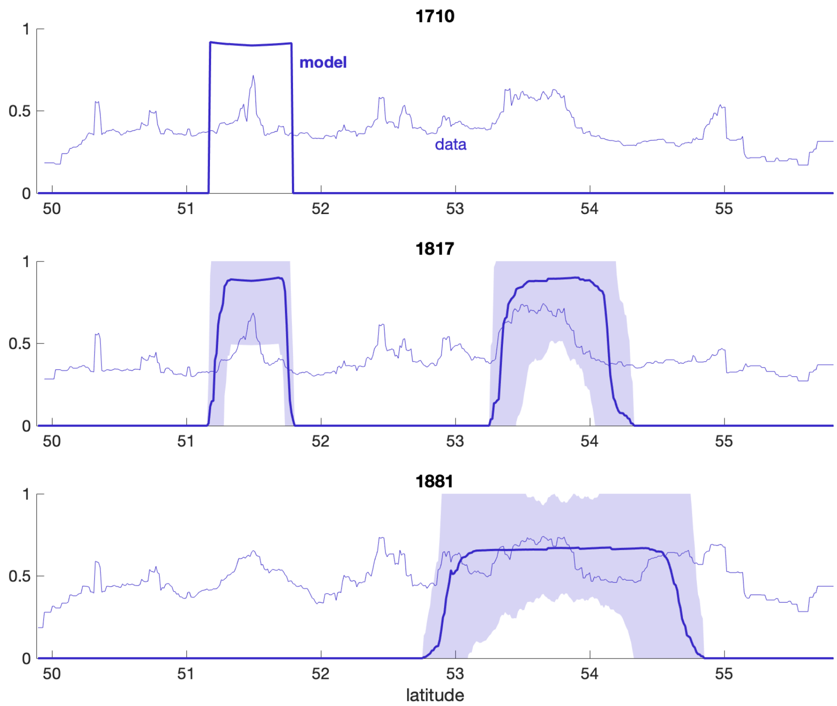


Figure 23: Endogenous Infrastructure: Primary Employment
(Average of 20 simulations ± 2s.d.)

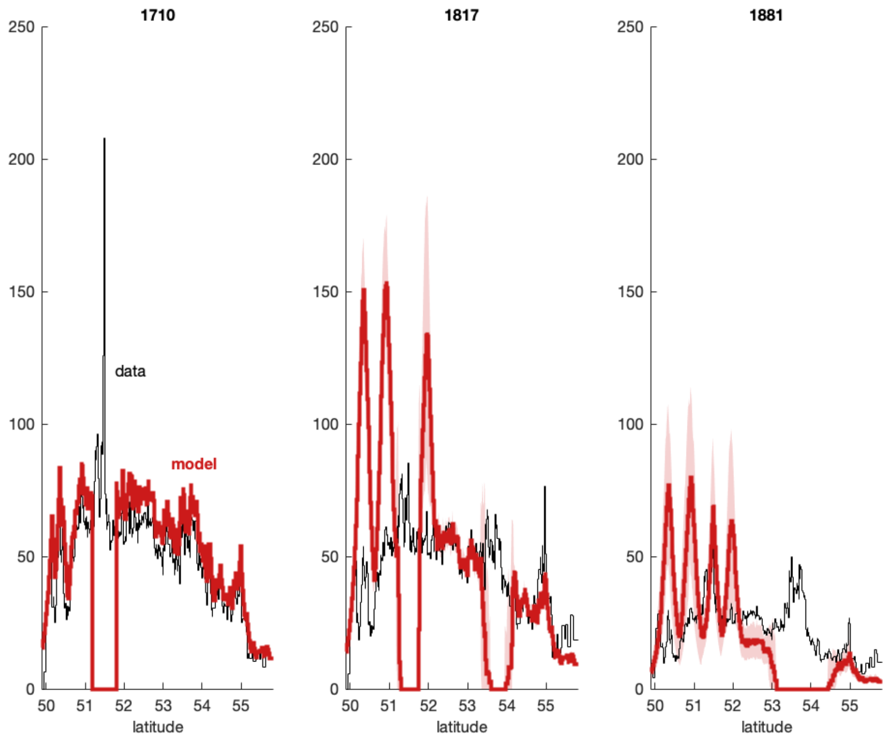


Figure 24: Endogenous Infrastructure: Secondary Employment
(Average of 20 simulations ± 2s.d.)

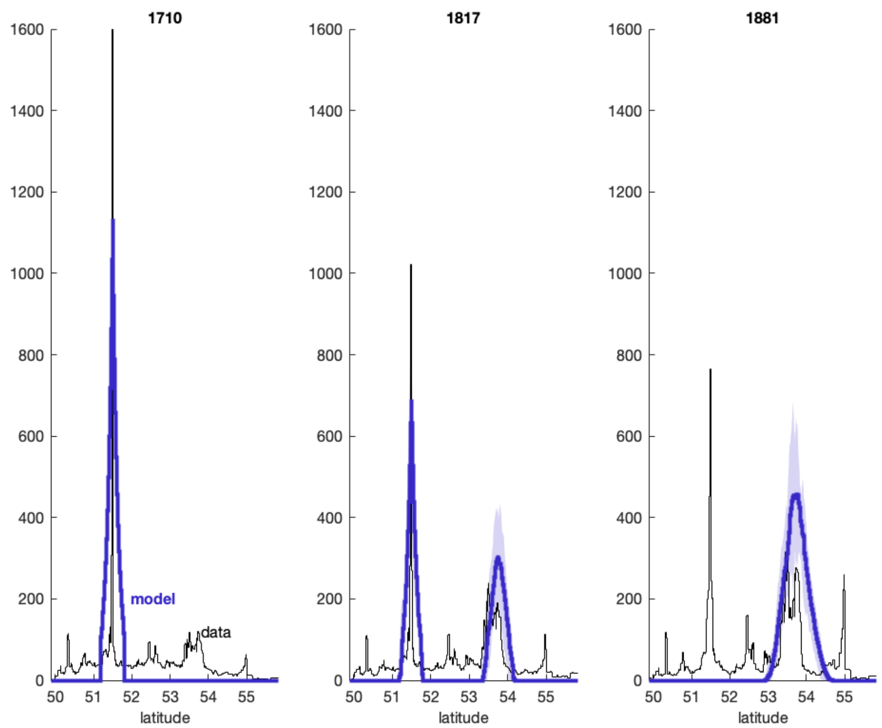


Figure 25: Endogenous Infrastructure: Tertiary Employment
(Average of 20 simulations ± 2s.d.)

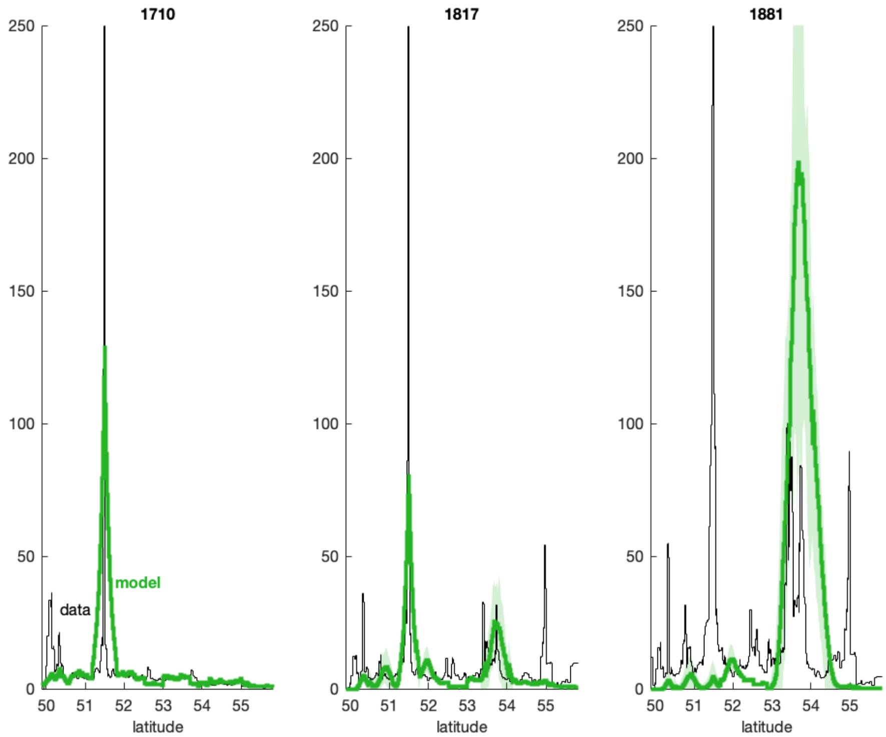


Figure 26: Endogenous Infrastructure: Primary Shares
(Average of 20 simulations ± 2s.d.)

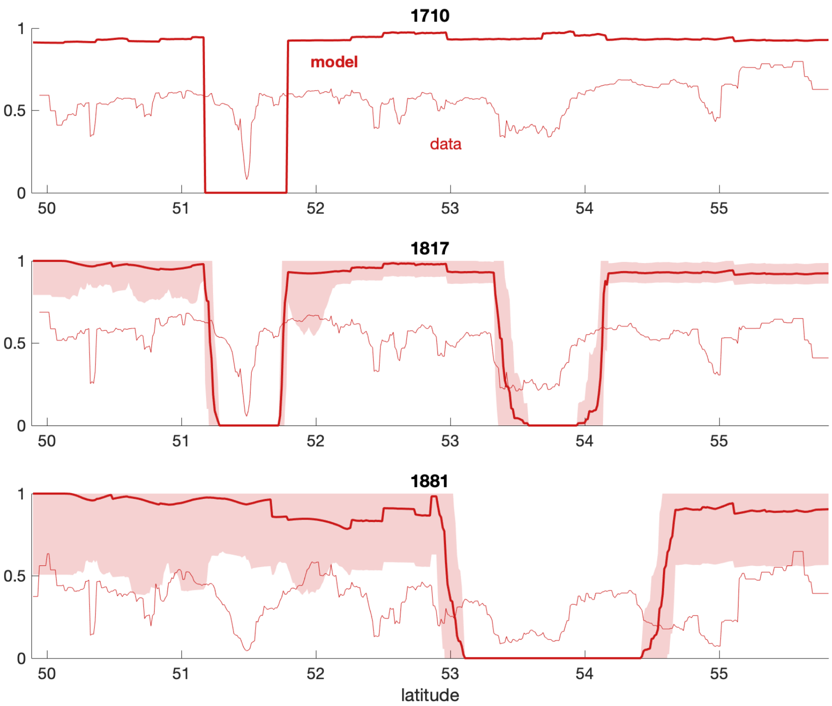


Figure 27: Endogenous Infrastructure: Secondary Shares
(Average of 20 simulations ± 2s.d.)

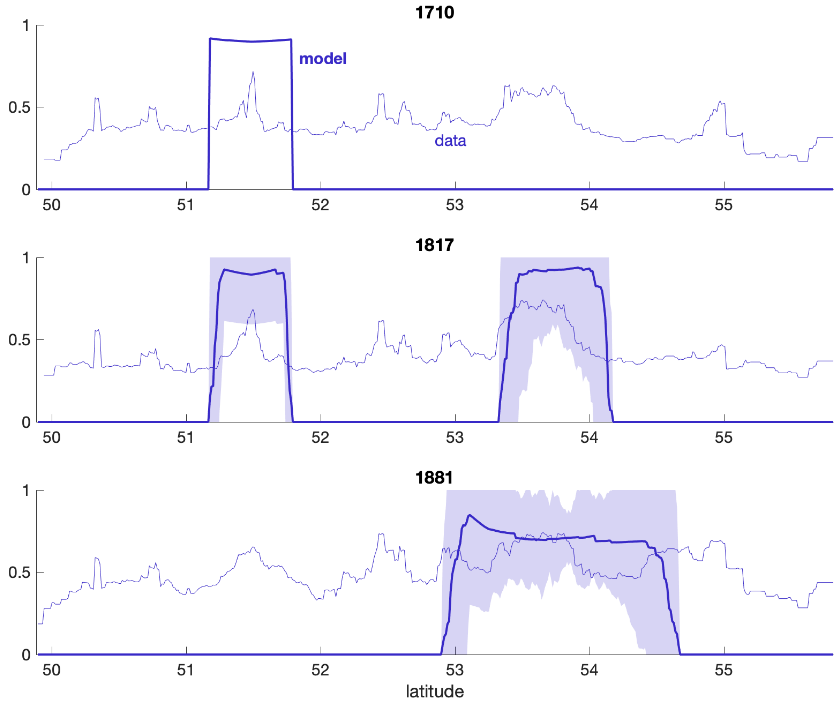
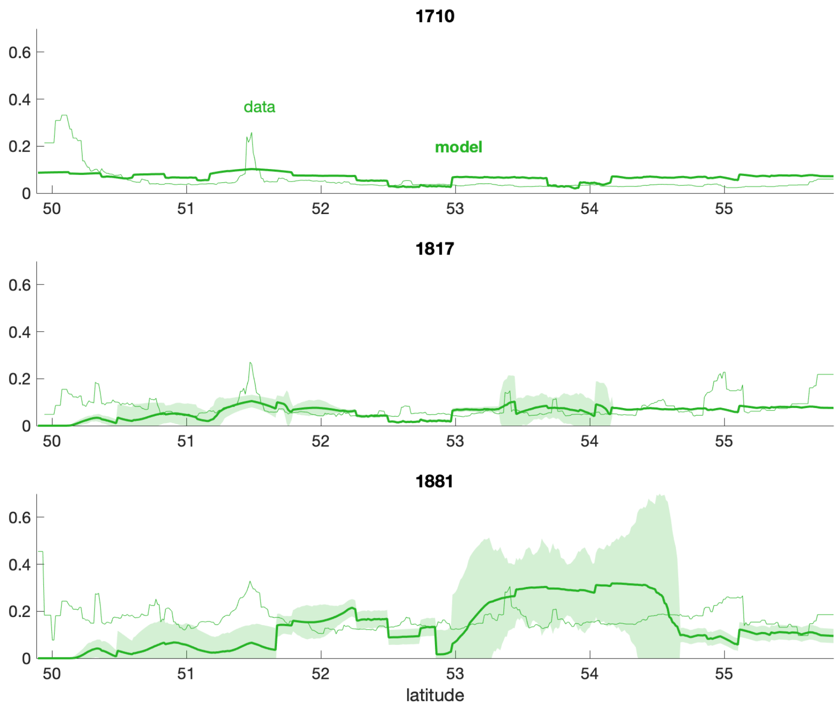


Figure 28: Endogenous Infrastructure: Tertiary Shares
(Average of 20 simulations ± 2s.d.)



B A two-region model

The model in the main text incorporates a continuum of locations which can then be directly connected to (a one-dimensional transformation of) the real world. However, it is reasonable to explore whether such an approach yields gains in terms of the ability to capture the dynamics of the industrial revolution over a simpler model with fewer points in space. If a simple two-region (e.g., north-south) model can replicate patterns of trade, aggregate labour shares, the location of innovation, the impact of transport costs and the implications for policy, then the continuum version is not a parsimonious one.

We present here a simplified version of the model in the main text wherein there are two regions, $r = \{S, N\}$. We think of these two regions as collapsing the southern and northern halves of the continuum of locations, such that the distance between region S and region N is $\frac{1}{2}$. In order to present some initial observations, we give results from a first period of the model calibrated to data for 1710. We also model transport costs as exogenous. We thus drop the time subscript and ignore transport labor.

Preferences Consumers face the same optimization problem as detailed in Section 3.1. Preferences are non-homothetic, agents consume where they work and labour is mobile, meaning that, in equilibrium, utility is equal across regions.

Production Land supply at each region is normalized to one. In the model with a continuum of land, we restrict each location to only one sector, but since there are many such locations a sector may on aggregate increase or decrease its overall land use. There can be three versions of a two-region model. We could assume that each region can be occupied by only one sector. With only two regions, such a restriction would result in complete and persistent specialization of each region which is not something we observe in reality. Alternatively, if multiple sectors can occupy the same region they may do so either with fixed land supply per sector per region or with competition for land between sectors within a region. As such, we permit multiple sectors to operate in each single region, sharing the land. In particular,

$$A(r) = Z_A(r)L_A(r)^{\alpha}G_A(r)^{(1-\alpha)} \tag{24}$$

$$M(r) = Z_M(r)L_M(r)^{\mu}G_M(r)^{(1-\mu)}, \tag{25}$$

where now $G_i(r)$ is a sector i firm's choice of land in region ℓ at time t. In the version of the model with fixed land supply, we set $G_i(r) = \frac{1}{2}$ for each i and all r.

We assume there is no diffusion of technology between regions.[23] Firm profits are thus,

$$p_i(r)Z_i\hat{L}_i(r)^{\imath}G_i(r)^{(1-\imath)} - w(r)\hat{L}_i(r) - R_i(r)\hat{G}_i(r), \tag{26}$$

where $\hat{G}_i(r)$ is optimal choice of land.

Trade We set κ to be exogenous and, since there are only two regions, we do not use a region index. We ignore labor employed in transport.

Equilibrium As previously, landowners maximise rental income. As such, all land is used so in equilibrium $G_A(r) + G_M(r) = 1$ for all r and, since different firms can now share a region, this implies $R_M(r) = R_A(r) = R(r)$. The first order conditions, along with these equilibrium conditions, yield regional land prices. Equilibrium in land then requires $\sum_{r\in\{S,N\}} R(r) = \bar{R}$, where $\bar{R}$ is the aggregate land rent.

Equilibrium in goods trade implies that the excess supply by a good in one region is, net of transport costs, equal to excess demand of that good in the other region. In particular,

$$\begin{aligned} x_i(S) - c_i(S)[L_M(S) + L_A(S)] - \kappa\,|x_i(S) - c_i(S)\,[L_M(S) + L_A(S)]| \\ +x_i(N) - c_i(N)[L_M() + L_A(N) = 0. \end{aligned} \tag{27}$$

Finally, equilibrium in the labour market requires that $\sum_\ell \sum_i L_i(r) = \bar{L}$.

Calibration Initially retained baseline parameters are as in Table 1. Again, we can use the data to invert the model to obtain average initial production technologies, as in the continuum model. As before, we use occupational data that is normalized by the average width of each region (since the north is thinner than the south) We treat continuum-model locations $\ell = [1, 250]$ as region $r = S$, and locations $\ell = [251, 500]$ as region $r = N$. The sum of employment in each sector over those locations thus determines initial productivity in each region. We also use the same initial guesses for $\bar{u}$, p_M and $\bar{R}$.

Simulation A first observation can be based on comparative advantage. When we calibrate using data aggregated to two regions, there is only a slight comparative advantage to the south in manufacturing; $\frac{Z_M(S)}{Z_M(N)} = 1.22$ and $\frac{Z_A(S)}{Z_A(N)} = 1.20$. While the south initially has a large, spatially-concentrated secondary sector in London, its agriculture is, as described in Allen (2004), also more productive than the north. In a simulation of the model with the transport cost used in the continuum model, there are no net gains from trading between the regions.

[23]The absence of diffusion approximates the diffusion parameter chosen in the continuum model. With $\delta = 15$, only 0.06% of an innovation realised at the midpoint of the south arrives at the midpoint of the north ($e^{-\delta * \frac{1}{2}}$).

Second, we can compare simulation outcomes when we consider a transport cost which does yield trade in equilibrium. As a baseline, when transport costs are zero the excess supply of manufacturing in the south is only 2% of southern manufacturing output. When we set transport costs to be $\varkappa = 0.1$ (one third of that in the original parameterization), we obtain the results in Table 4. Regardless of whether land is fixed or chosen optimally, the initial aggregate labour share in agriculture is lower than that in both the data and the continuum model. This is a consequence of neglecting the spatial concentration of highly-productive manufacturing. With productivity averaged over large regions, the land required to satisfy demand for manufactured goods crowds-out that available for agriculture and, by consequence, labour demanded by agriculture is lower.

Table 4: Initial labor shares (exc. tertiary)

	Data	Continuum	2-region (lower κ) (endog. G)	2-region (lower κ) (fixed G)
$L_A(1710)/\bar{L}$	51.6%	50.6%	43.9%	44.9%
$L_M(1710)/\bar{L}$	48.4%	49.4%	56.1%	55.1%

Third, the existence of productive secondary locations that are also spatially concentrated matters for the location of an industrial takeoff. The location of innovative manufacturing is a product of, among other factors, the relative local productivity in manufacturing and the relative local cost of energy. A productivity disadvantage in some locations of the north relative to the south may, as we see in reality, not limit innovation if energy costs are also sufficiently low there, relative to the south. In the calibration of the continuum model, the peak southern manufacturing location is 2.39 times as productive as the peak northern manufacturing location. Despite that productivity difference, innovation begins in the north since energy costs six times more in the south. Based on a two-region calibration, southern manufacturing is, as noted, initially 1.22 times as productive as northern manufacturing. This suggests that the energy price differences in a two-region model could be far lower and still yield a northern industrial revolution – that is, the importance of very cheap energy in the north is only clear in the continuum model. Moreover, the existence of a few northern locations that are relatively more productive than northern locations on average is crucial. As we saw in the counterfactuals in Section 4.4, other locations in the north are not sufficiently productive, relative to the south, to take advantage of lower energy costs.

To summarize, a two-region model does not have trade in equilibrium with the original transport cost, misses aggregate labour shares by a wider margin and cannot explain a role for energy prices. Moreover, these comparisons with the continuum model are against two-region representations of the data. A two-region model cannot, for example, capture a different impact on growth of a spatially-disaggregated infrastructure policy. We can conclude that there is, in this instance, merit in treating space as a continuum of locations.

C Computation of Nash equilibria in Section 5

Section 5 presents a version of the model in which county-level investment in infrastructure depends on the investment decisions of other counties. At the beginning of each period, the model runs through the following algorithm[24] to compute the infrastructure equilibrium:

1. Where N is the number of counties, set P to be a $2^{N-1} \times N$ array containing an empty column 1 and all binary permutations for $\mathcal{I}(i)$ in columns $2 : N$.
2. Calculate optimal strategy for County 1:
   (a) For each permutation in P, calculate $T(\ell, t)$ and $\bar{\kappa}(\ell, t)$ for ℓ in counties $2 : N$.
   (b) Run full model both with County 1 investing and not investing for each permutation in P.
   (c) Calculate gain to County 1 of investment under each permutation in P.
   (d) Add County 1 investment decisions to the array of permutations P.
3. Calculate optimal strategy for County 2:
   (a) P contains County 1 decisions given permutations over $2 : N$, including the optimal response to the County 2 investment decision.
   (b) Run full model both with County 2 investing and not investing for each permutation in P.
   (c) Calculate gain to County 2 of investment under each permutation in P.
   (d) For each permutation, check that the County 2 decision is consistent with the County 1 decision.[25] Add consistent County 2 investment decisions to the array of permutations P and collapse P to unique remaining permutations.
4. Repeat Step 3 for counties $3 : N$, checking at each part (d) that the County decision is consistent with strategies in previous counties.
5. After calculating optimal strategy for County N, consider the investment permutations left in P,
   (a) If P is empty, set $\mathcal{I}(i) = 0$ for all i and pass to the full model.
   (b) If P contains more than one permutation, select that row in P which maximises $\int_\ell R(\ell, t)\mathrm{d}t$ and pass to full model.
   (c) If P contains one permutation, pass to the full model.

[24] Full MATLAB code is available online or direct from the author.

[25] Suppose that $N = 3$ and that at step 2 County 1 invests only if both County 2 and County 3 invest. At step 3, it may be profitable for County 2 to invest in all permutations of County 1 and 3. However, the investment by County 2 when County 3 invests and County 1 does not is not consistent with County 1's decision to invest.