

The Death and Life of Great British Cities*

Stephan Heblisch Dávid Krisztián Nagy Alex Trew
Yanos Zylberberg

July 5, 2021

Abstract

This paper studies how industries shape the life and death of cities. Using data on the population and industrial composition of English and Welsh cities over the course of two centuries, along with a novel measure of exogenous land fragmentation within a narrow ring around city borders, we show that: (i) cities with a more elastic land supply at the onset of the nineteenth century experience a swift structural transformation during which they specialize in a few key industries; (ii) the boom during the nineteenth century is, however, followed by a long-run bust. We develop a dynamic spatial model of cities and their industries over time to understand the forces which govern the life and death of cities. The model helps untangle the role of aggregate industry trends from city-specific externalities à la Jacobs in explaining long-run dynamics.

Keywords: structural transformation; specialization; cities over time; quantitative economic geography; land fragmentation.

JEL codes: F63, N93, O14, R13

*Heblisch: University of Toronto, CESifo, IfW Kiel, IZA, SERC; stephan.heblisch@utoronto.ca; Nagy: CREI, Universitat Pompeu Fabra and Barcelona GSE; dnagy@crei.cat; Trew: University of Glasgow; alex.trew@glasgow.ac.uk; Zylberberg: University of Bristol, CESifo, Alan Turing Institute; yanos.zylberberg@bristol.ac.uk. We thank participants in seminars, e.g., at CREI, for very useful comments and suggestions. We acknowledge support of the ANR/ESRC/SSHRC, through the ORA grant ES/V013602/1 (MAPHIS: Mapping History).

1 Introduction

Many of the great cities that drove the industrial transformation of the nineteenth century have declined in relative, and sometimes absolute, terms in the twentieth. The formerly thriving mining towns of south Wales, the rust belt cities in the north-eastern United States, the Ruhr valley in Germany – these locations once employed generations of workers but now struggle to find renewed economic success. The initial, often century-long, relative success in those cities, and subsequent slow reversals in fortune, can be tied to the dynamics of the industries that first rose to support them and then fell into under-utilization and depression. The life and death of these great cities points to the long-run costs of specialization—in industry and in services—as the process of creative destruction and technological change plays out at the level of entire cities.

This paper studies how the life and death of cities relates to the dynamic evolution of industries. To do so, we rely on a unique data-set that allows us to follow the population and industrial composition of cities in England and Wales over the course of the last two centuries. This context is ideal for our analysis. The economic geography of England and Wales was transformed in a process that was initiated by the invention of new technologies in key sectors in the 18th century and then accelerated in the 19th century after the conclusion of wars against France, as we describe below. Rapid growth in international trade and the resulting shift to large-scale, steam-powered industrial production transformed both the scale and location of industries that underpinned the sustained increase in aggregate growth. Similar dynamics subsequently played out in other parts of the world, and go on to this day in modern economies, both industry- and service-based.

We predict city growth *during industrialization* with a novel measure of exogenous land fragmentation around historical city borders. To this end, we identify towns in 1817—before trade openness accelerated after the end of the wars with France—as the set of potential future metropolises, inferring initial boundaries from historical maps. At this time, many workers in industry were located in small-scale, often rural, workshops, producing for home or local consumption. The trade shock induced a demand for large-scale production in export-oriented industries, which shifted production to cities.

To predict the potential for city growth, we borrow insights from the historical literature and conjecture that the concentration of land ownership in the immediate hinterlands of cities affected the pace at which they could respond to an increased demand for land. The underlying argument is at the heart of the land assembly

problem (see, e.g., [Eckart, 1985](#); [Strange, 1995](#)): a higher degree of fragmentation, or a lower concentration, makes negotiations to acquire additional land for urban use more costly.

To isolate exogenous variation in land concentration within a narrow ring around 1817 city borders, we exploit plausibly exogenous terrain characteristics including elevation, ruggedness, (time-persistent) soil attributes and water bodies. [We develop an algorithm to predict the natural delineation of fields from multidimensional breaks in these natural characteristics](#).¹ One may think of this algorithm as identifying natural fault lines between potential agricultural land parcels. Importantly, while the local characteristics of city hinterlands at the beginning of the nineteenth century did influence urban sprawl and the pace of structural transformation in the following decades, they did not (directly) affect the later evolution of cities during the twentieth century, all of them having grown out of this narrow ring and thus being subject to another topography at their now expanded borders.

Our reduced-form specifications exploit this exogenous variation in the ease or difficulty of cities to industrialize and grow in the mid-19th century. We present novel stylized facts about the short- and long-run relationships between industry structure and city prosperity. First, we show that cities with a better initial ability to grow underwent more rapid structural transformation. Along with this transformation, they specialized in the high-growth, export-oriented industries that emerged during the time. Second, the initial boom was followed by an eventual bust in the twentieth century that reversed many of the benefits from the boom. This average picture masks some differences across cities, mostly driven by the concentration of their sectoral advantages: cities with large comparative advantages in a few industries experienced both a larger boom, and a larger bust.

To interpret these empirical findings, we develop a multi-sector dynamic spatial model and study quantitatively the mechanisms behind cities' boom and bust. The model features a finite set of cities that trade the products of a finite set of industries. At any point in time, cities can differ in their sectoral productivities, their land supply elasticity, and their trade costs with other cities. Population is fully mobile across cities. City population and industrial structure can influence future productivity in a flexible way. In particular, the model allows for both dynamic Marshall–Romer and Jacobs externalities, which have been discussed as important determinants of city growth and industrial structure in the literature (see [Carlino](#)

¹We validate the predictive power of this measure of land fragmentation by comparing it with the actual observed concentration of ownership from land tax registers, Tithe records and later micro-census records where land acreage is reported by landowners.

and Kerr, 2015, for a review).² We provide a condition for the existence and uniqueness of equilibrium and offer an algorithm to quickly solve the model by extending the theoretical results of Allen et al. (2020) to a multi-sector, multi-location setup.

The model offers the following rationale behind the boom and subsequent bust observed in the data. In the short run, a decrease in trade costs leads to more industrial specialization. This increased specialization is particularly pronounced in the most centrally-located cities, which experience the largest gains from trade. As a result, population also reallocates towards these cities, rationalizing the short-run boom and the corresponding increase in specialization found in the data. In the long run, a **bust can follow the boom** for two possible reasons. First, industries that were successful earlier may turn into a period of decline over their life-cycle, e.g., because of an aggregate structural change of the economy (Ngai and Pissarides, 2007) or because of international competition and exposure to trade (Pierce and Schott, 2016). Second, dynamic Jacobs externalities, generating larger productivity gains in historically more diverse cities, might direct population and economic activity away from cities that were highly specialized in the past. In the last part of the paper, we illustrate the model’s ability to disentangle these two mechanisms by simulating the model on a simple (linear) geography.

We contribute to several strands of existing research. First, we connect to a quantitative literature studying the dynamic evolution of the spatial distribution of economic activity. The closest contributions in this literature are Allen and Donaldson (2020), Berkes et al. (2021), Eckert and Peters (2018) and Fajgelbaum and Redding (2021).³ Our main contribution to this literature lies in proposing a tractable multi-sector dynamic model with various dimensions of heterogeneity (e.g., trade costs, sectoral productivities and land supply elasticities) that we can theoretically characterize. Specifically, we provide a sufficient condition for the existence and uniqueness of the model’s equilibrium and offer an algorithm to solve the equilibrium in an extension to the recent results of Allen et al. (2020).

Second, we contribute to the literature on the drivers of the first industrial revolution (for a survey, see Clark, 2014). Stokey (2001) showed the quantitative importance of trade for structural transformation at a macroeconomic level, while Allen (2009) has argued that the mechanism at work is via the impact of trade on relative prices. We suggest another channel, in particular that trade induced the growth of

²Marshall–Romer externalities operate within industry, implying that cities specialized in a narrow set of industries are the ones primarily benefiting from these externalities. By contrast, Jacobs externalities operate across industries and hence favor cities with a diverse industrial base.

³For a more comprehensive survey of the quantitative spatial literature, we refer the interested reader to Redding and Rossi-Hansberg (2017) and Nagy (2021) who focuses on the use of quantitative spatial models to address historical questions.

manufacturing in cities, accelerating the transition to large-scale, export-orientated growth in the nineteenth century. The transition to large-scale production was underpinned by the human capital in ‘mechanics and tinkerers’ (Mokyr, 2009; Kelly et al., 2020; Hanlon, 2021) that took hold in cities. Lastly, the focus on the first structural transformation *and* the second transition away from industry in the twentieth century also links to the macroeconomic literature on growth and structural transformation summarized in Herrendorf et al. (2014).

Third, we relate to the literature on industries as drivers of urban evolutions (Duranton, 2007) and specifically contributions that discuss the negative effects of specialization (Glaeser et al., 1992; Duranton and Puga, 2001; Faggio et al., 2017; Hebllich et al., 2019) on city development. One feature of our model, responsible for the negative effect of specialization, is the crucial role played by between-industry spillovers à la Jacobs (1969) as drivers of development in the long run (Carlino and Kerr, 2015). The closest paper to ours in this literature is Henderson et al. (1995), which discusses the life-cycle of industries and cities. In a more specific application, Glaeser (2005) studies Boston’s development over nearly four centuries and points to the important role of human capital in reinventing Boston after periods of crisis and decline. Alternative mechanisms behind the rise and fall of cities could be distortion in the acquisition of human capital (see, e.g., Franck and Galor, 2017) or the life cycle of industries (Henderson et al., 1995). The key advantage of our empirical strategy is that it allows us to quantify the long-run causal impact of specialization on city outcomes, while it will account for each of these alternative explanations.

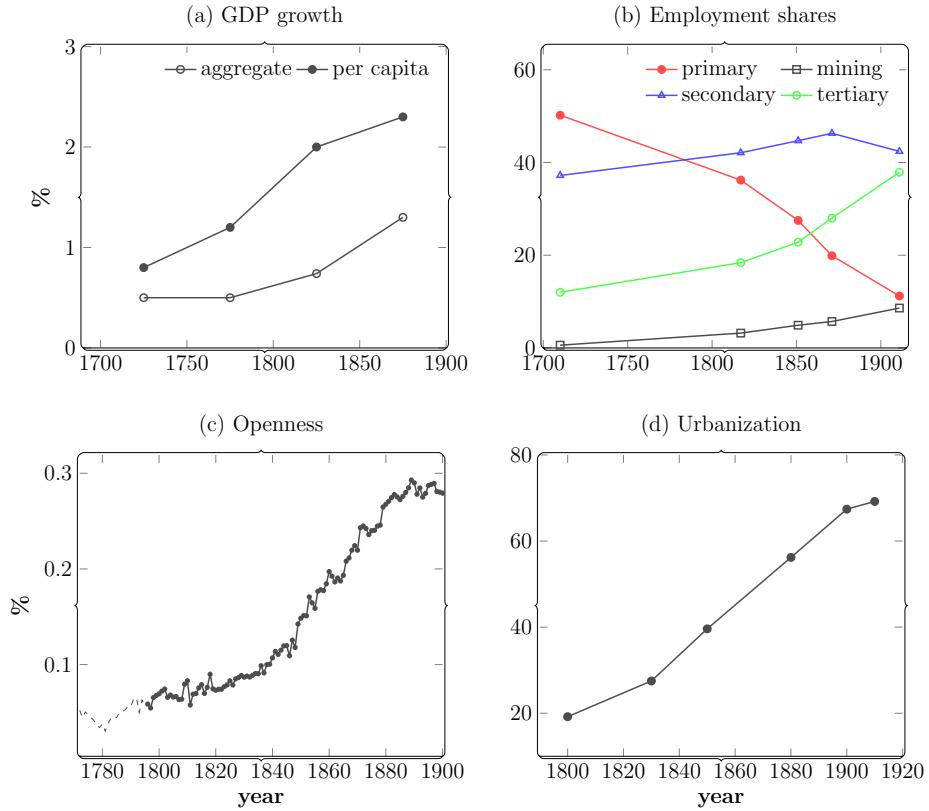
The remainder of the paper is organized as follows. Section 2 presents some historical context. Section 3 describes our data sources, data construction, and empirical strategy. Section 4 establishes a few key stylized facts, which motivate the structure of the model (Section 5). Finally, Section 6 takes the model to the data to disentangle the mechanisms behind the effect of industry specialization on the “birth” and “death” of cities and Section 7 briefly concludes.

2 The industrial revolution in England

This section provides a brief overview of Britain’s historical development over the past three centuries, focusing on the factors that led to the sustained increase in economic growth in the short- and also the long-run.

The industrial revolution can be broadly characterized by four stylized facts: (a) the emergence of new technologies in key sectors leading to sustained increases in growth rates of per capita incomes; (b) a growing share of employment in non-

Figure 1: The industrial revolution, 1700–1900



Notes: (a) Growth figures are from [Broadberry et al. \(2015\)](#). (b) The employment shares are classified according to the PST system described in [Wrigley \(2010\)](#). We report the available data for male adults in England and Wales ([Shaw-Taylor and Wrigley, 2014](#)). (c) Openness is defined as the sum of imports and exports as a share of GDP, using [Hills et al. \(2010\)](#) and [Broadberry et al. \(2015\)](#). (d) Urbanization is defined as the share of total population in cities over 5,000, from [Bairoch and Goertz \(1986\)](#).

agriculture; (c) the growth in domestic and international trade; and, (d) an increasing share of the population living in cities. How each of these fit together, and which are the causal elements in the context of the industrial revolution that first emerged in England, is not completely settled (see, for example, the survey in [Clark, 2014](#)).

Figure 1 depicts trends of these four dimensions over the period 1700–1900. While many of the key industrial technologies emerged in the mid-eighteenth century, growth in per capita output accelerated only in the early nineteenth century. A further puzzle is presented by the share of employment in the secondary sector, which is high as early as 1710 and grows only marginally until the mid-nineteenth century. Most striking are the dramatic changes in openness, which accelerates after 1820, and urbanization, which grows by nearly fifty percentage points over the century.

Two highly influential hypotheses on the causes of the industrial revolution are [Mokyr \(2009\)](#) and [Allen \(2009\)](#). For Mokyr, the industrial revolution was driven

by the emergence of “attitudes and aptitudes” (Mokyr, 2021)—a respect for entrepreneurs and inventors, and the growth of useful human capital—that begat the enrichment of society. Kelly et al. (2020) further makes the case that the distribution of useful human capital—mechanical workers—across counties in England was key. Hanlon (2021) corroborates this prospective with a specific focus on the professionalization of invention through the emergence of engineers. Allen, in contrast, emphasizes the demand for new technologies—high wages and low energy costs induced the capital-biased (and labor-saving) technical change that drove the growth in the export-oriented industries (see also Allen, 2021). Those high wages arose as a result of globalization and the increasing external demand for manufactured goods in which England had a comparative advantage.

We present in this paper an alternative channel by which trade induced the industrial revolution in England: the growing external demand for manufactured output caused a shift in modes of production away from partly rural, low-scale, domestic-oriented and water-powered production to urban, specialized, export-oriented and large-scale factories in which steam power dominated. This trade shock induced specialization that benefited the growth of the economy in the nineteenth century but sowed the seeds of decline in the twentieth.

Shaw-Taylor and Wrigley (2014) suggested that the 19th century shift of manufacturing production into towns outside of London was key to the success in the 19th century. As Crafts (1989) made clear, the majority of industrial employment in the early nineteenth century industry was small-scale production for local markets. Growth in export-oriented industry was key to the eventual rise in the average standard of living. Stokey (2001) builds and calibrates a macroeconomic model of Britain, incorporating the potential role of trade, energy cost and technical change at 1780 and 1850. Stokey finds that trade explains all the decline in agricultural production, over a quarter of the increase in manufacturing, and half of the increase in real wages. Harley and Crafts (2000) and Clark et al. (2014) conduct static computable general equilibrium (CGE) exercises. Both find trade to be an important part of the industrial transformation in the nineteenth century; Clark et al. finds that the welfare contribution of trade by 1850 approaches 30% of aggregate welfare.

Harley and Crafts (2000), Stokey (2001), Clark (2014) are all aggregate accounts of the role of trade, while Harley and Crafts and Clark also omit dynamic effects. Each of these exercises also ends in the mid-nineteenth century. Our argument goes further to incorporate *disaggregated and dynamic* consequences that emerged, for good and ill, over the subsequent two centuries. The interaction between trade,

urbanization and the change in the structure of production is key.⁴

The shift of industrial production from a rural to an urban setting has been studied extensively, most notably with the hypothesis that rural ‘protoindustrialization’ constituted an important stage in the progress to more general industrialization (Mendels, 1972). That such protoindustrialization caused the industrial revolution has not sustained later analysis (Ogilvie, 2008), but the existence of extensive proto-industries suggests a way to explain the high share of employment in the secondary sector seen in Figure 1 (panel b) in spite of low urbanization levels. In between the low-scale artisanal home-production of finished manufactures and the large-scale, city-centered factory production that typifies late-industrialization are the extensive and well-organized rural industries, some of which exported beyond the locality (Hudson, 2004; Goose, 2014). Early factory production was also frequently rural, relying on water power and with rural workers housed by entrepreneurs around production facilities (Trinder, 2000). While steam engines began to proliferate in the eighteenth century (Nuvolari et al., 2011), the transition to steam engines as the predominant motive power, and the “triumph of the factory system” that went along with it, was not complete until the mid-nineteenth century (Musson, 1976).

Factory production using new technologies concentrated in growing cities (see, e.g., Trew, 2014) and met the accelerating external demand after the 1820s. The interaction between the growing demand for scale that resulted from trade, and the spatial limits on industrial expansion was key (as Hennessy, 2006, p.103, suggests),

When size brought economies of scale with the growth of world markets
... a steelworks high up in a south Wales valley or a shipyard crammed
into the narrow banks of the lower Tyne experienced increasing disad-
vantage. Take the great plant at Dowlais ... Tucked between the hills
above Merthyr Tydfil almost into the uplands of the Brecon Beacons, its
site was hopeless...

Such constraints could be geographical or could be the remnants of historical property rights over land that persisted (Denman, 1958; Hoskins, 1988; Neeson, 1996; Hudson, 2004). In discussing the constraints on attaining scale in the smelting of iron in South Wales, Trinder (2000) notes that “[p]atterns of housing were dispersed, following patterns set by pre-existing fields and property boundaries rather than those of order and convenience” (p.820).

⁴Voigtländer and Voth (2013) considers a model in which urbanization and manufacturing demand result from non-homothetic preferences. In the presence of a non-monotonic relationship between income and death rates, the demographic shock of the Black Death causes the economy to shift to a high income (urbanized and industrialized) steady state which prefigured the industrial revolution in Europe.

3 Data

This section describes our data sources, the data construction, most notably, the creation of a predicted measure of land fragmentation from topography and soil characteristics, and provides descriptive statistics for the measure of land fragmentation.

3.1 Data sources

Census of England and Wales. The main data source is the Census of England and Wales, which provides a unique characterization of population and industrial composition at the level of about 11,500 parishes over the course of two centuries (1801–1911, 1971–2011). The census provides population counts from 1801 onward, but a precise decomposition of the labor force across occupations only after 1851 (when the micro-census records become available). We thus rely on a quasi-census based on (adult male) baptism records collected between c.1817 (referred to below as c.1817) in order to retrieve consistent industrial composition at the parish level before the time of trade-induced industrialization ([Shaw-Taylor and Wrigley, 2014](#)). One issue with census data is that the lowest administrative units—the parishes—are regularly redefined, merged or split over the course of the nineteenth century. We thus apply an “envelope” algorithm which considers the union of the different parishes covering the same points over time (see Appendix B.1). For instance, if a parish is split into two parishes in 1891, we would group the two sub-parishes from 1891 onward such as to create a consistent, unique parish from 1801 to the current day. This grouping is less relevant at the city level, as none of these re-compositions of lowest administrative units significantly affect the allocation of administrative units across cities.

Geography. To characterize the immediate neighborhood of cities and the (local, temporary) constraints to land supply, we gather high-quality raster maps at a disaggregated level: elevation (OpenLandMap, 30m resolution); soil organic carbon content (OpenLandMap, 250m resolution); soil bulk density (Soil bulk density, 250m resolution); a detailed soil classification (National Soil Resources Institute); and a dataset of all rivers and smaller streams in England and Wales.

Transportation and land ownership. We complement the previous data on population, occupation and geography with the transportation infrastructure (roads, navigable waterways, train lines and train stations), as provided by the Cambridge

Group for the History of Population and Social Structure. This dynamic characterization of transportation allows us to measure access to resources through the transportation network and trading costs across different cities (see Appendix B.1).

A crucial component of the empirical analysis consists in the construction of an exogenous measure of land fragmentation based on topography and soil characteristics (i.e., natural breaks between possible agricultural land parcels). One channel through which exogenous land fragmentation might put a strain on city growth is that it might contribute to fragmented land ownership and thus make the land harder to assemble. To validate this channel, we collect actual measures of land ownership concentration at the beginning of the nineteenth century (from land tax registers) and in 1851, 1861 and 1881 from micro-census records where land acreage is reported by landowners. Inferring land ownership concentration from micro-census records requires a systematic text analysis as the information has not been coded by the I-CeM project.

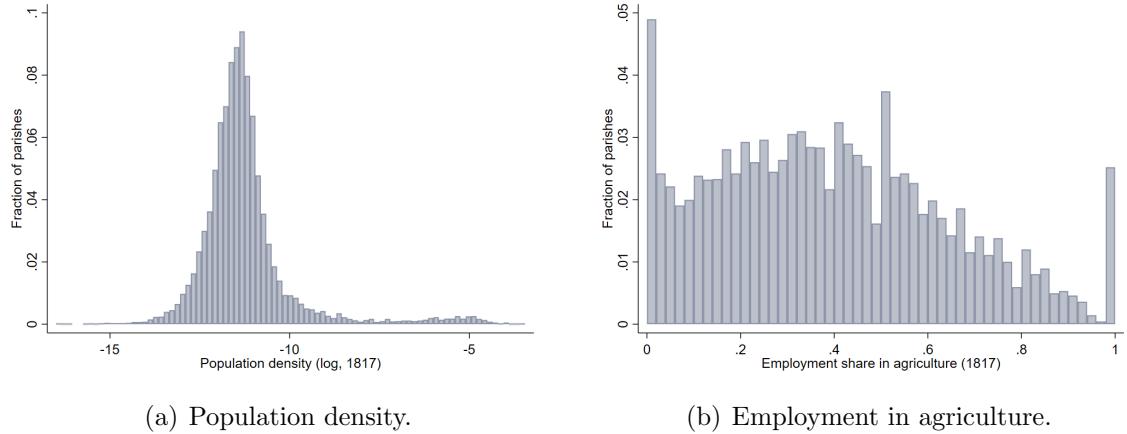
3.2 Data construction

This section describes how we construct a predictor ζ_c for industrialization in city c , which is exogenous to later city dynamics. The idea is to measure land fragmentation at the boundaries of city c , right before the time of rapid, trade-induced industrialization.

City boundaries at the onset of rapid industrialization. We implement two procedures to identify city boundaries at the onset of rapid industrialization. First, we collect early maps (Ordnance Survey maps in the South of England and in Wales, other sources in the North of England), and detect settlements and building imprints. We then aggregate these imprints to create a collection of tightly related buildings: a town or city. Second, we use our baptism records nested at the parish level in c.1817 to classify parishes as urban or rural and aggregate nearby, urban parishes into a town or city. We describe the latter procedure below.

The urban/rural classification procedure relies on a set of variables including (log) population density, 1-digit sectors and 1-digit occupations at the parish level, as measured around 1817. We consider a simple machine-learning, unsupervised algorithm, a k-means clustering using the Euclidean distance, in order to isolate two groups of parishes. As shown in Figure 2, there is a large disparity in population density and in agricultural employment across parishes. However, and fortunately, population density exhibits two modes in the data, a low-density mode and a high-density mode. This feature ensures that the clustering procedure is mostly driven by

Figure 2: Distribution of population density and agricultural employment share across parishes in c.1817.



Notes: Panels A and B respectively display the distributions of population density and agricultural employment share across parishes in c.1817.

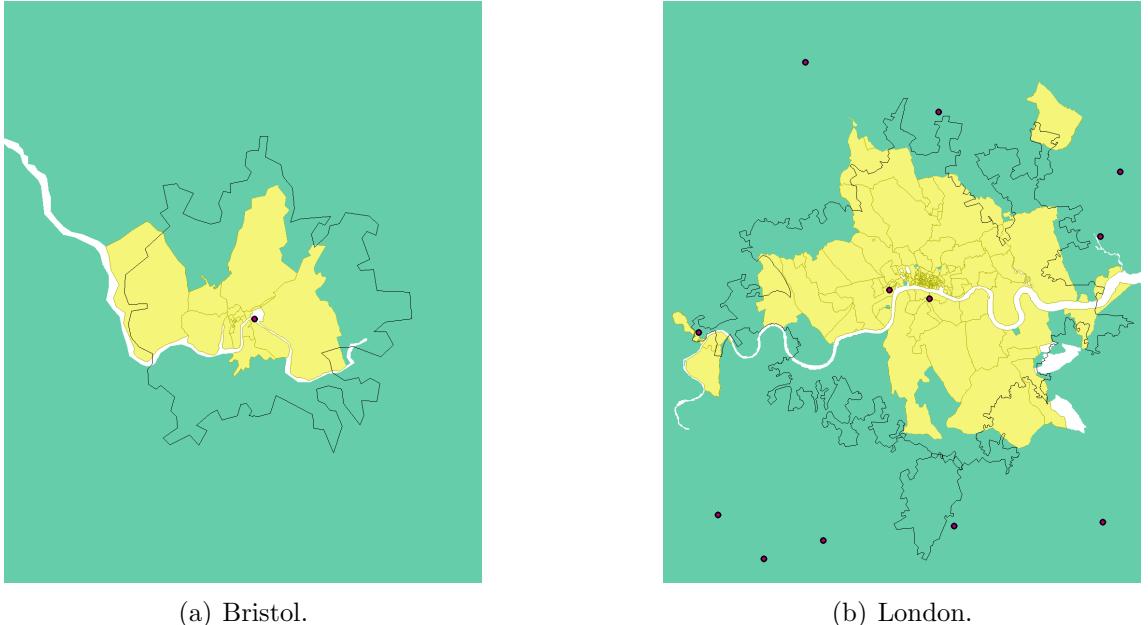
differences in population density, and that it identifies consistently proper towns from rural manufacturing settlements. Note, however, that mining settlements would be identified as urban (they are densely populated), even though they never led to a proper urbanization process as they did in some parts of the United States.

We then perform simple geographic operations to merge nearby, adjacent parishes into polygons which could be interpreted as towns or cities, even though they may include rural parts of urban parishes and a few urban areas within very rural parishes may be omitted. We show the outcome of the procedure in Figure 3 with a comparison between the clustering procedure in c.1817 versus city boundaries as extracted from maps in 1880–1890 (Bristol, left panel; London, right panel). One can see that both cities expand to some degree, but this expansion is not homogeneous across all directions. We discuss next how we predict the extent of such expansion with land ownership concentration.

Land fragmentation around city boundaries. Land ownership concentration across different parts of England and Wales was not only instrumental to the development of agriculture, as illustrated by the effect of enclosures (Neeson, 1996), it was also crucial in disciplining urban sprawl during the era of rapid industrialization.

Indeed, when land markets are not perfectly competitive and land parcels (and their rights of use) cannot be split arbitrarily, developing land at the fringe of cities may be a challenge. For instance, a textile mill requires a large parcel of flat land to construct a factory, but also possible access to water sources. When a suitable loca-

Figure 3: Urban parishes around Bristol and London: delineation in c.1817 versus city borders in 1880–1890.

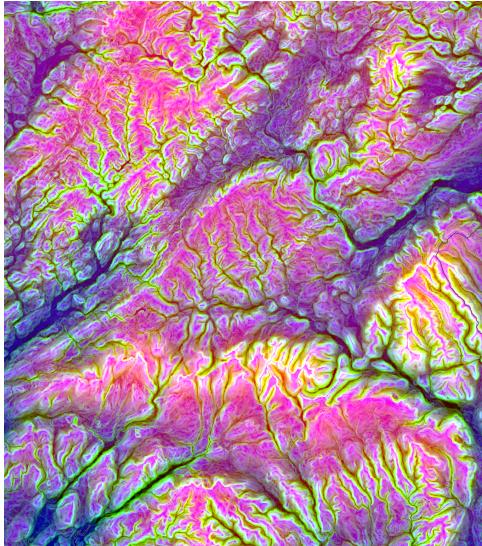


Notes: The clustering procedure classifies about 520 parishes as urban around 200 urban centers (dark lines: city boundaries as extracted from maps, 1880–1890; yellow: parishes classified by the clustering procedure as urban in c.1817).

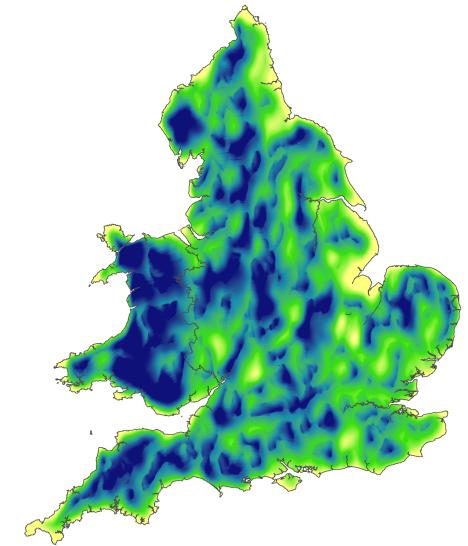
tion spans multiple land parcels, the possible buyer needs to engage in a multilateral bargaining in which the value of the marginal parcel increases as the buyer acquires rights to use for other parcels. The number of different parties then matters. This issue is a “standard” hold-up problem, which has been labelled as the land assembly problem in this specific context (see, e.g., [Eckart, 1985](#); [Strange, 1995](#)). The consequences for urban sprawl are straightforward: high land concentration at the fringe of the city makes negotiations to develop the land for urban use costly. As a result, cities with concentrated ownership in their immediate hinterlands have a more elastic land supply in the short run and may better respond to sudden bursts in land demand. The rapid industrialization induced by a surge in external trade in the mid-19th century, as evidenced in Section 2, is one such shock.

A major issue with land ownership concentration in the immediate hinterlands of cities is that it may reflect past and future city dynamics. The relationship between land ownership concentration and urban sprawl may thus be “contaminated” by omitted variation—the relative productivity of land between urban and rural use inducing different structure of ownership—but also reverse causality—land owners less willing to own large plots around cities most likely to expand. For these reasons,

Figure 4: A multi-band raster covering England and Wales and a low-resolution map of predicted fragmentation.



(a) Topography (South Wales).



(b) Fragmentation (England and Wales).

Notes: The left panel displays the three topographic bands of the multi-band (elevation, slope, Multi-Scale Topographic Position Index) raster covering England and Wales in the RGB spectrum. Each band is standardized between 0 and 1. The right panel displays a low-resolution map of predicted fragmentation across England and Wales.

we would like to consider a measure of land ownership concentration with the following characteristics: (i) the measure, as evaluated within a neighborhood of city boundaries at the beginning of the nineteenth century, should predict urban sprawl and the pace of industrialization in the following decades; (ii) the measure should not (directly) affect the later evolution of cities during the twentieth century, when conditioned on the right control variables (e.g., the elasticity of land supply in later periods, once cities have expanded beyond the narrow ring and are thus subject to another topography at their borders).

We construct a plausibly exogenous measure of land fragmentation by exploiting fine-grained terrain characteristics including elevation, ruggedness, time-invariant soil attributes and water bodies. We leave the details of the procedure to Appendix B.2 and only summarize its main steps below. First, we combine these different dimensions of soil characteristics into a multi-band raster covering England and Wales at a resolution of 30m (see the left panel of Figure 4).

Second, we use an image segmentation algorithm which groups pixels by their proximity in actual space (i.e., along the physical distance) and in the space as defined by the bands of the raster. These algorithms are typically unsupervised and used to create super-pixels in images and isolate homogeneous (and contiguous)

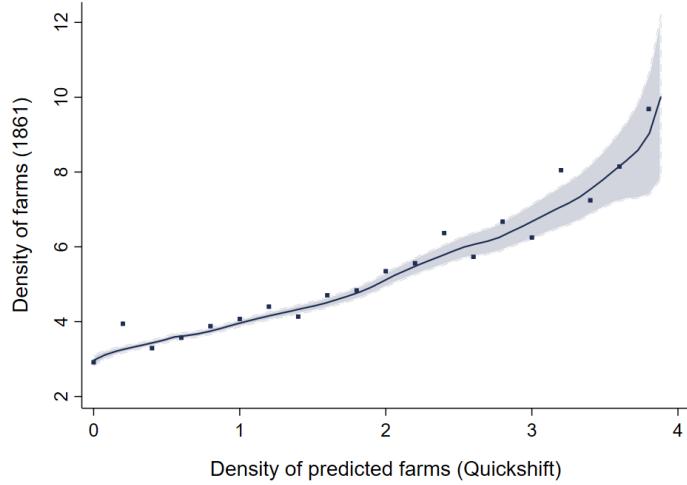
color zones. In such cases, the multi-band raster is constituted of three bands (R: red, G: green, B: blue). In our application, the raster may contain many more bands, but the principle is the same: the algorithm maximizes a weighted sum of target distances within constituted superpixels, with a weight allocated to physical distance relative to the “color distance.” A color superpixel is, in our application, a patch of land with homogeneous topography and soil characteristics: a typical agricultural parcel, e.g., as delineated by enclosures in some parts of England and Wales. Within the class of local mode-seeking algorithms able to perform this classification, we opt for Quickshift, which is disciplined by three parameters: a scale parameter (the size of a typical farm), a maximum physical distance which disciplines the extent to which the algorithm looks for neighbors, and a relative weight between distance in the multi-bands-space and physical distance. The latter typically disciplines how compact the predicted farms will be. We display an example of super-pixelization in Appendix B.2.

Third, we isolate the propensity for the city to grow at different times over the course of two centuries by drawing buffers of different widths (e.g., 1, 2, 3, 5, 10 kms) around city boundaries at the onset of the rapid industrialization. One can think about the narrow rings as predicting the propensity for cities to industrialize at the onset of the nineteenth century and the wider rings as controlling for later land supply elasticities. The quantitative model developed in Section 5 will allow for cities to face varying land supply elasticities over time, in part to capture the previous intuition.

An exogenous measure of land fragmentation. The predicted measure of land fragmentation should be correlated with the actual concentration of land ownership. We validate the predicted measure of land fragmentation by comparing it with actual farm density as collected from micro-census records in 1861 across all parishes of England and Wales (see Figure 5).

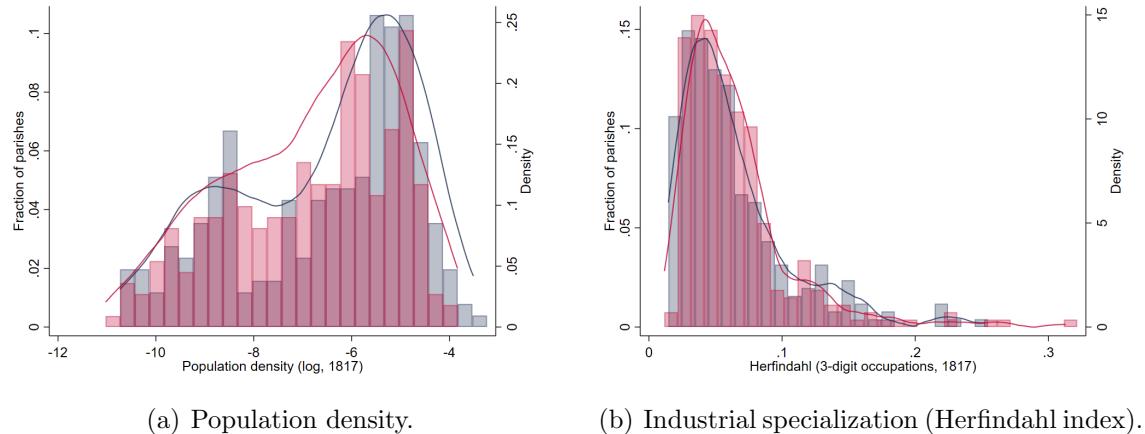
While this exercise shows that our measure of land fragmentation does predict land ownership fragmentation, it does not rule out that the measure is related to city dynamics before industrialization. Thus, we further validate the land fragmentation measure by comparing the internal structure of cities with different degrees of land fragmentation in their “external crusts,” as calculated at the onset of the nineteenth century. Figure 6 provides the equivalent of a balance test comparing the distribution of population density and industrial specialization (calculated as a Herfindahl index across 3-digit occupations) for cities with above- and below-median predicted land fragmentation in their immediate hinterlands. We do not find very

Figure 5: Validation of the predicted measure of land fragmentation.



Notes: This Figure displays the measure of predicted fragmentation versus actual farm density as collected from micro-census records in 1861 across all parishes of England and Wales. We create 20 bins of density and the dots represent the average actual farm density within each bin. The lines are locally weighted regressions on all observations. Note that the conditional correlation between the two measures is 0.25 (once conditioned on the separate topographic and soil characteristics).

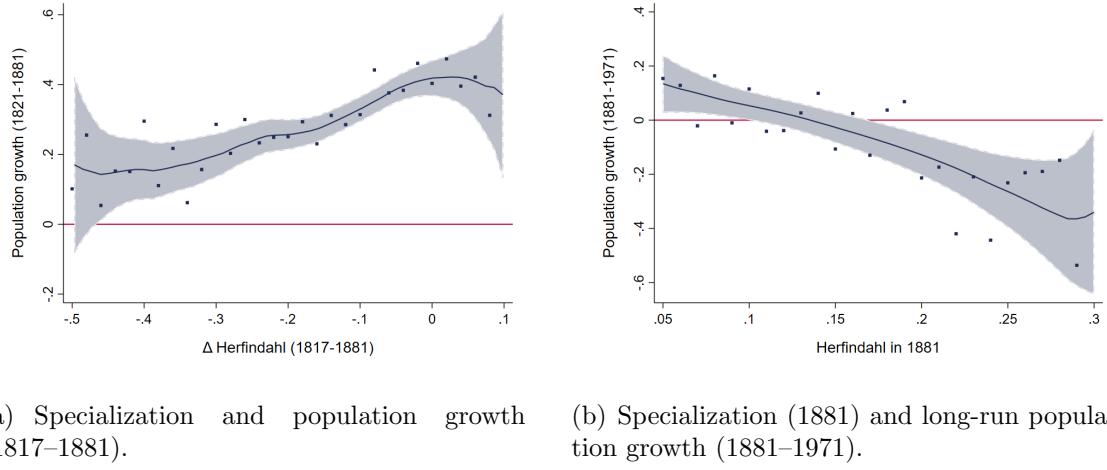
Figure 6: A balance test for cities with different predicted land fragmentation in their immediate hinterlands.



Notes: This Figure displays the distribution of population density (left panel) and industrial specialization (calculated as a Herfindahl index across 3-digit occupations, right panel) for cities with above- (red) and below-median (blue) predicted land fragmentation in their immediate hinterlands.

marked differences.

Figure 7: Specialization and urbanization during the nineteenth and twentieth centuries.



(a) Specialization and population growth (1817–1881).
(b) Specialization (1881) and long-run population growth (1881–1971).

Notes: Panel a shows the joint dynamics of specialization and population growth over the course of the nineteenth century (1817–1881) across parishes of England and Wales. Panel b shows how specialization in 1881 correlates with longer-run population growth between 1881–1971. Herfindahl indices are based on 3-digit occupations in 1817 and in 1881. The lines are locally weighted regressions on all observations.

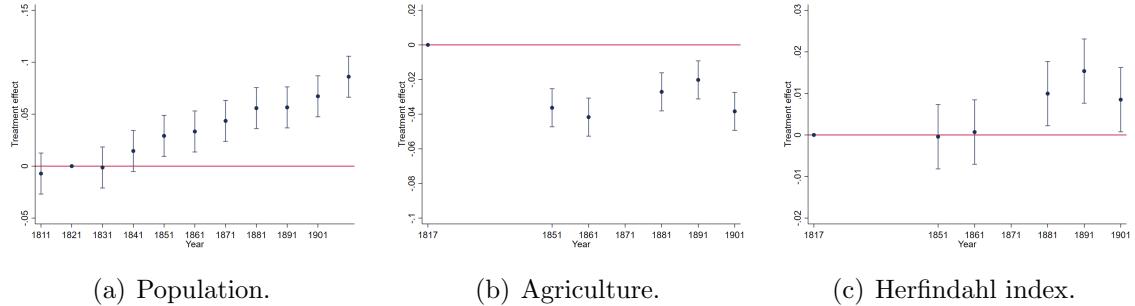
4 Stylized facts

In this section, we build upon the previous data sources and provide the following stylized facts: 1. Growing cities become more specialized in the short run, but they perform poorly in the long run; 2. Low land fragmentation around initial city boundaries predicts population growth, structural transformation and specialization in the short run; 3. Low land fragmentation around initial city boundaries predicts poor city performance in the long run.

Population growth and specialization. Figure 7 shows the joint dynamics of specialization and population growth over the course of the nineteenth and twentieth centuries (1817–1881, 1881–1971) across parishes of England and Wales. Growth appears to entail an increase in industrial specialization between 1817 and 1881 (panel a).

This is confirmed by a causal analysis using predicted land fragmentation as an exogenous measure for the capacity of cities to grow during the course of the nineteenth century. In this exercise (shown in Figure 8), we regress an outcome y_{ct} for city c at time t on city-FEs, time-FEs, a dummy equal to 1 for below-median land fragmentation (the “treatment”) interacted with time-FEs and the following controls interacted with time-FEs: travel time to coal/London/nearest city, yield,

Figure 8: Land fragmentation, specialization and urbanization during the nineteenth century.



Notes: This Figure displays the time-varying coefficients in front of the treatment at different periods in time; 1817 is the baseline year in panel b and c, 1821 is the baseline year in panel a. Panel a uses (log) population as the dependent variable; panel b uses the employment share in agriculture as the dependent variable; panel c uses Herfindahl indices based on 3-digit occupations as the dependent variable. The regression includes as explaining variables: city-FEs, time-FEs, a dummy equal to 1 for below-median land fragmentation (the “treatment”) interacted with time-FEs and a set of controls interacted with time-FEs (travel time to coal/London/nearest city, yield, elevation/slope).

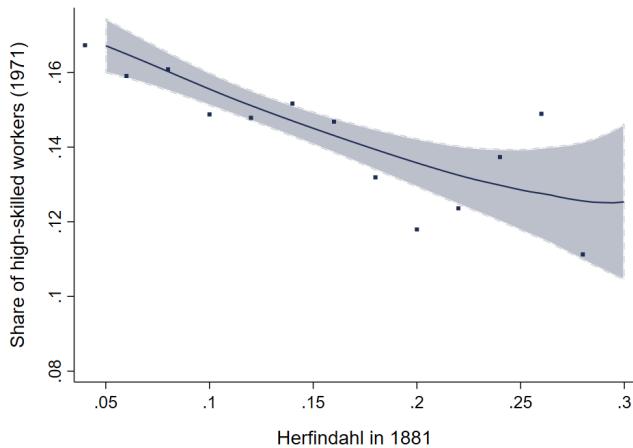
elevation/slope. Figure 8 displays the time-varying coefficients of the treatment at different periods in time; 1817 is the baseline year in panel b and c, 1821 is the baseline year in panel a.

We find that low land fragmentation, inducing lower cost of urban sprawl, predicts population growth (panel a of Figure 8), structural transformation (panel b) and specialization (panel c) between 1817/21–1881. More specifically, we use (log) population as the dependent variable in panel a, the employment share in agriculture in panel b, and Herfindahl indices based on 3-digit occupations in panel c. Cities with below-median predicted land fragmentation see their population increase gradually, and become about 10% larger by the end of the nineteenth century. This urbanization process induces a lower employment share in agriculture in their immediate hinterlands and a more specialized workforce from 1881 onward. This specialization appears to be detrimental in the longer run, as we see next.

Specialization and longer-run dynamics. The previous specialization pattern, as induced by structural transformation and urbanization, has long-run effects on city dynamics. Indeed, panel b of Figure 7 shows how specialization in 1881 correlates with longer-run population growth between 1881–1971. Less specialized cities in 1881, with a Herfindahl index around 0.05, are about 40% larger in 1971 than in 1881, compared to more specialized cities in 1881, with a Herfindahl index around 0.20.

Given the previous evidence, this pattern may indicate a mere reversion to the mean. It does not. Indeed, we show in Figure 9 that specialized cities in 1881

Figure 9: Specialization and longer-run dynamics.



Notes: This Figure displays the relationship between the Herfindahl index in 1881 (3-digit occupations) and the share of high-skilled workers in 1971 (professionals, managers etc.), but the findings are robust to adding clerks and/or skilled manual workers. The lines are locally weighted regressions on all observations.

experience a bust during the course of the twentieth century, which leaves them less attractive, even in absolute terms, than other cities—and even before the swift decrease in manufacturing employment accompanying Thatcher’s reforms. More specifically, the population of high-skilled workers in 1971 for less specialized cities in 1881, with a Herfindahl index around 0.05, is 25% larger than for more specialized cities in 1881, with a Herfindahl index around 0.20. We find that they are also less populated and remain very specialized in unreported tests.

These stylized facts provide evidence about the joint dynamics of urbanization and specialization in cities, showing that the fate of cities is tightly related to that of their industries. The next section provides a more structured approach, by developing a quantitative model of cities and their industries over time.

5 A multi-sector dynamic spatial model

To provide a framework in which we can rationalize the previous stylized facts and study the connections between a city’s short- and long-run specialization and development, we develop a multi-sector, dynamic model of cities. To make the model suitable for quantitative analysis, we theoretically characterize equilibrium existence and uniqueness and offer an algorithm to solve the model efficiently, building on the results of [Allen et al. \(2020\)](#). Finally, we simulate the model on a stylized geography to illustrate how it can replicate the empirical facts documented in the previous

section.

5.1 Setup

The model involves a finite number of cities $c = 1, \dots, C$ and industries $i = 1, \dots, I$. Time is discrete and is indexed by $t = 0, 1, \dots$. Within each industry, every city produces its own variety that consumers view as different from the varieties produced in other cities. There is an exogenous number \bar{L} of workers in the economy. Each worker lives for one period, maximizing her utility from the consumption of varieties. Each worker is endowed with one unit of labor. The worker decides which industry to work for and which city to live in. For simplicity, we abstract from moving costs between cities and industries.

In what follows, we describe the four main building blocks of the model: workers' preferences, the production technology, the equilibrium within a time period t , and the dynamic process that links subsequent periods to each other.

Workers' preferences. If a worker who lives at time t decides to work in industry i and to live in city c , she chooses her consumption levels to maximize her utility

$$U_{ict} = \max \left[\sum_{j=1}^I \left(\sum_{d=1}^C (q_{jdt})^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\sigma-1}{\epsilon-1}} \right]^{\frac{\sigma}{\sigma-1}}, \quad (1)$$

subject to the budget constraint

$$\sum_{j=1}^I \sum_{d=1}^C p_{jdt} q_{jdt} \leq w_{ict} + R_{ct}, \quad (2)$$

where q_{jdt} is the worker's consumption of the city- d variety in industry j , p_{jdt} is the price of this variety, w_{ict} is the wage that prevails in the city-industry, and R_{ct} is the worker's share of land rents that are redistributed to workers living in the city. When choosing her city, industry and consumption, the worker takes all prices and wages as given. We assume that varieties are substitutes, i.e., the within-industry elasticity of substitution ϵ is greater than one.

Technology. Varieties are produced by perfectly competitive firms. The representative firm producing the city- c variety in industry i at time t faces the production function

$$Y_{ict} = \tilde{\gamma} \tilde{T}_{ict} L_{ict}^\gamma H_{ict}^{1-\gamma}, \quad (3)$$

where \tilde{T}_{ict} is the TFP of industry i in city c at time t , L_{ict} is the number of workers hired by the firm, and H_{ict} is the amount of land used by the firm. $\tilde{\gamma} = \gamma^{-\gamma} (1 - \gamma)^{-(1-\gamma)}$ is a constant that simplifies the subsequent formulas.

Varieties can be traded across cities, but they are subject to iceberg trade costs. We denote the iceberg trade cost prevailing between cities c and d in industry i at time t by τ_{icdt} . Naturally, we assume that trade costs are always non-negative, which amounts to $\tau_{icdt} \geq 1$.

Land is supplied in each city according to the supply function

$$H_{ct} = r_{ct}^{\zeta_{ct}-1}, \quad (4)$$

such that r_{ct} is the land rent and $\zeta_{ct} - 1$ is the land supply elasticity. We let the exogenous parameter driving this elasticity, ζ_{ct} , vary both across cities and over time—this feature is essential to mirror the heterogeneity across cities uncovered in the empirical analysis. It is natural to assume that $\zeta_{ct} \geq 1$, i.e., the supply function is never downward-sloping. Land rents are fully redistributed to workers who live in the city.

Within-period equilibrium. Before we turn to presenting the dynamic evolution of sectoral TFP levels, we set up the equilibrium within a given time period t for *given* TFP levels in that period. Separating the within-period equilibrium from the overall dynamic equilibrium of the model will later prove useful for the theoretical characterization of equilibria.

In this within-period equilibrium, we impose that the labor market clears in each city; the land market clears in each city:

$$\sum_{i=1}^I \frac{1-\gamma}{\gamma} w_{ict} L_{ict} = r_{ct} H_{ct}, \quad (5)$$

markets clear for each variety:

$$(w_{ict} + R_{ct}) L_{ict} = \sum_{d=1}^C \left(\frac{P_{idt}}{P_{dt}} \right)^{1-\sigma} \left(\frac{p_{ict} \tau_{icdt}}{P_{idt}} \right)^{1-\epsilon} \sum_{j=1}^I (w_{jdt} + R_{dt}) L_{jdt}, \quad (6)$$

where P_{idt} is the CES price index of industry- i varieties in city d ,

$$P_{idt} = \left[\sum_{c=1}^C p_{ict}^{1-\epsilon} \tau_{icdt}^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}}, \quad (7)$$

and P_{dt} is the CES price index of all consumption goods in city d ,

$$P_{dt} = \left[\sum_{i=1}^I P_{idt}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (8)$$

and free labor mobility equalizes utility across industries and cities at any point in time:

$$\bar{U}_t = U_{ict}. \quad (9)$$

Dynamic evolution of TFP. We now present the assumptions on how sectoral TFP levels evolve over time. We allow the TFP of each industry to be influenced by agglomeration externalities. In particular, externalities in period t can take the form

$$\tilde{T}_{ict} = T_{ict} L_{ct}^\alpha f_i \left(L_{c,t-1}, \{L_{jc,t-1}\}_{j \in I} \right), \quad (10)$$

where T_{ict} is the exogenous fundamental productivity of industry i in city c at time t . That is, **agglomeration externalities** may not only depend on the current population of city c (as standard in the literature) as well as on its past population (as in [Allen and Donaldson, 2020](#)), but also on the city's sectoral composition in the past. This process, which links the productivity of city-industries to the spatial and sectoral distribution of employment in the previous period, is responsible for the dynamics of the model and, crucially, underlies the joint evolution of cities and industries.

Equation (10) is a flexible formulation of externalities that allows for dynamic within-industry (Marshall–Romer) and cross-industry (Jacobs) externalities. It is a generalization of the dynamic TFP process in the one-sector model of [Allen and Donaldson \(2020\)](#).

5.2 Equilibrium existence, uniqueness and solution algorithm

In this section, we extend the theoretical results of [Allen et al. \(2020\)](#), and we establish a sufficient condition under which the model's equilibrium exists, is unique, and can be solved by using a simple algorithm. While the condition itself is relatively involved, it is only a function of the model's structural parameters $\alpha, \gamma, \epsilon, \sigma$ and ζ_{ct} . Hence, for any combination of these parameters, one can check for equilibrium existence, uniqueness and tractability by checking whether the condition holds.

As the model's dynamics are uniquely characterized by Equation (10), the existence and uniqueness of the model's equilibrium naturally hinges on the existence and uniqueness of the within-period equilibrium for a given distribution of fundamentals, past populations $L_{c,t-1}$ and past sectoral employment levels $L_{jc,t-1}$. For a

given within-period equilibrium in $t - 1$, Equation (10) uniquely determines TFP (excluding static agglomeration externalities) in the subsequent period, implying that the entire dynamic equilibrium is unique. Therefore, we now proceed with characterizing the within-period equilibrium.

The within-period equilibrium conditions can be reduced to a system of $3IC$ equations (as shown in Appendix A.1),

$$\begin{aligned} x_{ict}^1 &= \sum_{j=1}^I \sum_{d=1}^C \left(x_{jdt}^2 \right)^{\frac{\alpha+\gamma}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \frac{\epsilon-1}{\sigma-1}} \left(x_{jdt}^3 \right)^{-\left[1 - \frac{1}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \right]} K_{icjdt}^1 \\ x_{ict}^2 &= \sum_{j=1}^I \sum_{d=1}^C \left(x_{jdt}^1 \right)^{\frac{\sigma-1}{\epsilon-1}} K_{icjdt}^2 \\ x_{ict}^3 &= \sum_{j=1}^I \sum_{d=1}^C \left(x_{jdt}^1 \right)^{\frac{\sigma-\epsilon}{\epsilon-1}} \left(x_{jdt}^2 \right)^{-\left[1 - \frac{\alpha+\gamma}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \frac{\epsilon-1}{\sigma-1} \right]} \left(x_{jdt}^3 \right)^{\frac{1}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)}} K_{icjdt}^3 \end{aligned} \quad (11)$$

where the $3IC$ unknowns x_{ict}^1 , x_{ict}^2 and x_{ict}^3 can be obtained from equilibrium prices, wages and population levels through the following change in variables:

$$\begin{aligned} x_{ict}^1 &= P_{ict}^{1-\epsilon} \\ x_{ict}^2 &= w_{ict}^{1-\sigma} \\ x_{ict}^3 &= w_{ict}^{1+(\gamma+\frac{1-\gamma}{\zeta_{ct}})(\epsilon-1)} L_{ct}^{1+(\frac{1-\gamma}{\zeta_{ct}}-\alpha)(\epsilon-1)} \end{aligned} \quad (12)$$

and K_{icjdt}^1 , K_{icjdt}^2 and K_{icjdt}^3 are the following functions of exogenous variables:

$$\begin{aligned} K_{icjdt}^1 &= \begin{cases} \left(\frac{1-\gamma}{\gamma} \right)^{-\frac{1-\gamma}{\zeta_{dt}}(\epsilon-1)} \hat{T}_{jdt}^{\epsilon-1} \tau_{jdct}^{1-\epsilon} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \\ K_{icjdt}^2 &= \begin{cases} (\gamma \bar{U}_t)^{1-\sigma} & \text{if } c = d \\ 0 & \text{otherwise} \end{cases} \\ K_{icjdt}^3 &= (\gamma \bar{U}_t)^{1-\sigma} \left(\frac{1-\gamma}{\gamma} \right)^{-\frac{1-\gamma}{\zeta_{dt}}(\epsilon-1)} \hat{T}_{jct}^{\epsilon-1} \tau_{jcdt}^{1-\epsilon} \end{aligned}$$

such that

$$\hat{T}_{ict} = T_{ict} f_i \left(L_{c,t-1}, \{L_{jc,t-1}\}_{j \in I} \right)$$

is the part of TFP that is exogenous in period t . In the following theorem, we theoretically characterize the existence and uniqueness of the solution to (11), and

hence of the model's equilibrium more generally.

Theorem 1. *The solution to (11), and hence the equilibrium of the model, exists and is unique under a condition that only depends on the values of structural parameters $\alpha, \gamma, \epsilon, \sigma$ and ζ_{ct} .*

Proof. See Appendix A.2. □

As the condition guaranteeing equilibrium existence and uniqueness is relatively involved, we only show it in Appendix A.2. The next theorem shows that, under the same condition as the one guaranteeing existence and uniqueness, a simple procedure can be used to solve the equilibrium on the computer.

Theorem 2. *Given Theorem 1, the equilibrium of the model can be solved by a simple algorithm that constitutes iterating on (11), then expressing equilibrium price indices, wages, population and sectoral employment levels by inverting (12).*

Proof. See Appendix A.2. □

Theorem 2 greatly improves the computational tractability of the model, despite the various dimensions of heterogeneity (in sectoral productivities, land supply elasticities and trade costs) that the model features. This is what allows us to use the model for the quantitative analysis of the next section.

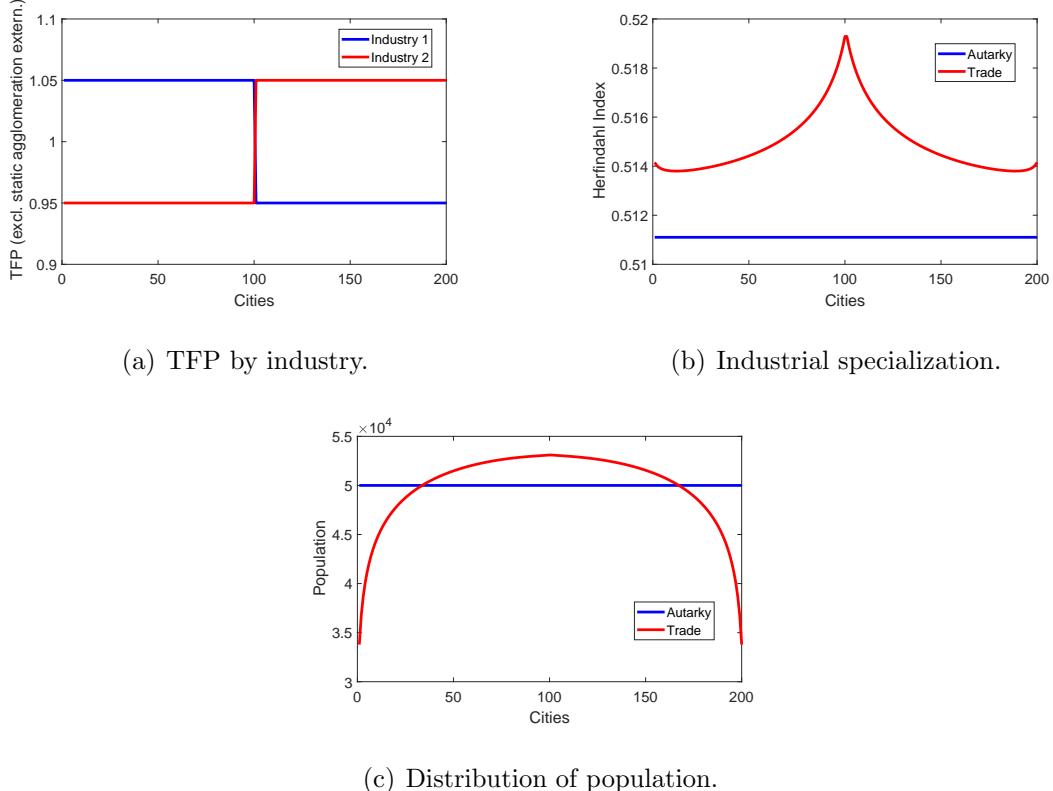
5.3 Illustration: a linear economy

In this section, we simulate the model on a simple geography to illustrate how it can rationalize the boom and bust of cities documented in Section 4. We focus on a country with 200 cities arranged on a line. There are two industries in the country. At the beginning of period 0, the 100 cities to the West of the line's midpoint have a high TFP of 1.05 in industry 1 but a low TFP of 0.95 in industry 2. The pattern is reversed in the 100 Eastern cities (top left panel of Figure 10).

Using the algorithm offered by Theorem 2, we simulate this stylized economy both under autarky and under trade. In either scenario, we simulate the economy for two subsequent time periods, period 0 and period 1. Under autarky, we assume that cities' trade costs with one another are infinitely high. Under trade, we assume that the cost of trading between cities c and d takes the form

$$\tau_{icdt} = (1 + dist_{cd})^\phi$$

Figure 10: Model simulation over a linear economy: period 0.



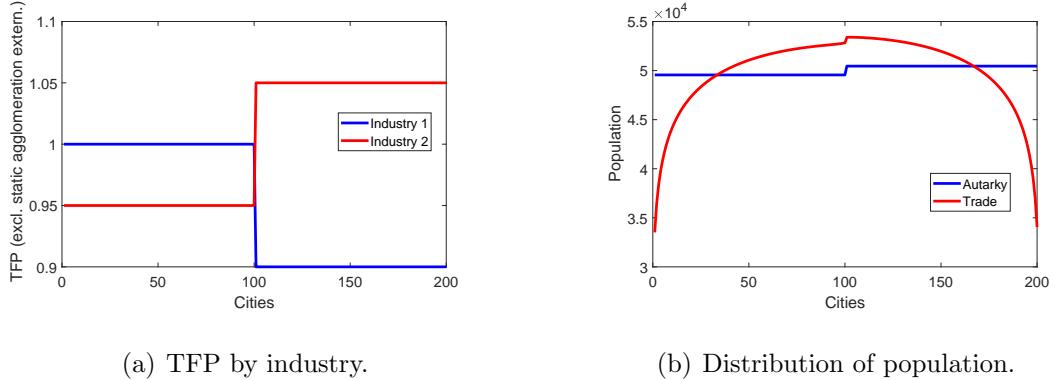
Notes: The values for the structural parameters are set as follows: $\alpha = 0.06$; $\gamma = 0.65$; $\epsilon = 5$; $\phi = 0.25$; $\sigma = 4$; $\bar{L} = 10,000,000$; $\zeta_{ct} = 1$.

in both industries and time periods, where $dist_{cd}$ denotes the Euclidean distance between cities c and d . That is, trade is cheaper between cities that are geographically close.

We set the values of the model's structural parameters to central values used in the literature ($\alpha = 0.06$, $\gamma = 0.65$, $\epsilon = 5$, $\phi = 0.25$, $\sigma = 4$). We set the country's total population to 10 million, which roughly equals the working population of England and Wales in the beginning of the 19th century. Finally, we set $\zeta_{ct} = 1$ for every city and time period for simplicity. This implies a land supply elasticity of zero (that is, fixed land supply) in each city.

Short run. We first look at the patterns of specialization and the distribution of population in period 0, i.e., in the short run. The top right panel of Figure 10 shows how specialized cities are under autarky (blue line) and under trade (red line), measured by their Herfindahl index across industries. Under autarky, all cities feature the same degree of specialization. This is not surprising as we made

Figure 11: Model simulation over a linear economy: period 1 under differential industry trends.



Notes: values of structural parameters set to: $\alpha = 0.06$; $\gamma = 0.65$; $\epsilon = 5$; $\phi = 0.25$; $\sigma = 4$; $\bar{L} = 10,000,000$; $\zeta_{ct} = 1$.

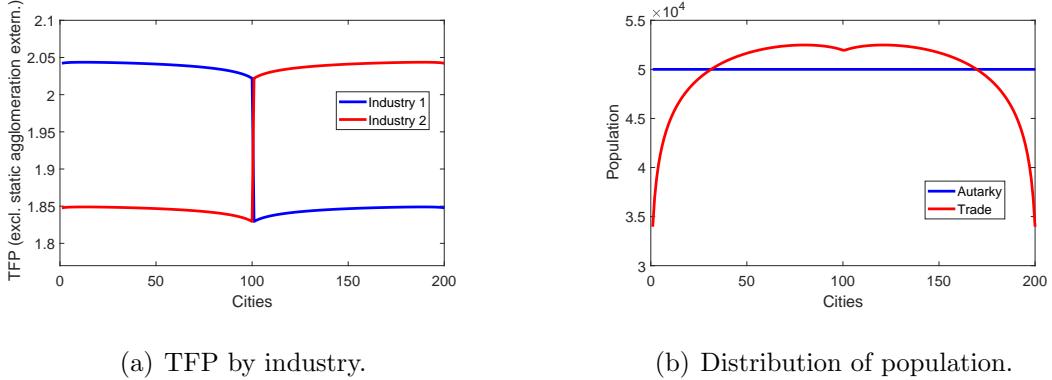
symmetric assumptions about their fundamentals; while Western and Eastern cities differ in what they are good at and therefore *what* they specialize in, they do not differ in the *degree* to which they specialize.

Under trade, cities that are near the center specialize more. Indeed, they are the ones with the best access to trade with other cities, and hence the largest room for specializing according to their comparative advantage. As the bottom panel of Figure 10 illustrates, population also reallocates towards these central cities as they benefit from trade through their increased specialization. This stands in stark contrast with the autarky scenario in which population is evenly distributed across cities. Thus, a short-run boom caused by trade favors cities in the center, which are able to gain from specialization and attract more people as a consequence.

Long run under differential industry trends. We study the distribution of economic activity in the long run (period 1) in two different cases. In the first case, we assume that industries differ in their productivity dynamics for reasons outside the model. More precisely, we assume that the TFP of industry 1 decreases by 0.05 uniformly across cities, while the TFP of industry 2 stays the same as in period 0 (left panel of Figure 11). Such differential industry trends may be due to nationwide productivity trends associated with structural transformation, as in Ngai and Pissarides (2007), or increased international competition, as in Pierce and Schott (2016). We abstract from other sources of TFP evolution, such as dynamic externalities, by setting $f_i(\cdot) = 1$ in Equation (10).

A comparison between the right panel of Figure 11 and the bottom panel of Figure 10 shows that population reallocates from the West towards the East as a

Figure 12: Model simulation over a linear economy: period 1 under dynamic Jacobs externalities.



Notes: values of structural parameters set to: $\alpha = 0.06$; $\gamma = 0.65$; $\epsilon = 5$; $\phi = 0.25$; $\sigma = 4$; $\bar{L} = 10,000,000$; $\zeta_{ct} = 1$.

result of differential industry trends. This is not surprising as Western cities were the ones specializing in industry 1, hence they are the ones that suffer from declining TFP in this sector. Therefore, this case of the model can rationalize the boom and the subsequent bust of center-West cities, which initially specialized in the industry that starts declining after period 0.

Long run under dynamic Jacobs externalities. In the second case in which we simulate the model in the long run, we abstract from differential industry trends but allow for dynamic agglomeration externalities. More precisely, we assume dynamic Jacobs externalities of the form

$$f_i(\cdot) = \left[\sum_j \left(\frac{L_{jc,t-1}}{L_{c,t-1}} \right)^2 \right]^{-1}$$

in Equation (10). This formulation implies that cities less specialized in period 0 (those with a lower Herfindahl index, $\sum_j \left(\frac{L_{jc,0}}{L_{c,0}} \right)^2$) see faster TFP growth by period 1. The resulting TFP distribution is presented in the left panel of Figure 12. Although all cities see an increase in TFP relative to period 0 (top left panel of Figure 10), central cities that were more specialized see a more modest increase than cities with a more diverse industrial base.

As the right panel of Figure 12 illustrates, this long-term disadvantage of initially more specialized cities has the ability to reverse the hump-shaped population distribution of period 0 (bottom panel of Figure 10). As a result, cities in the center see a period-1 bust after their period-0 boom.

How to tell apart differential industry trends and Jacobs externalities?

Illustrating the model for the simple geography of this section highlights that differential industry trends and dynamic Jacobs externalities both have the ability to rationalize cities' boom and bust. That said, they do have different implications. In particular, differential industry trends imply that cities that hosted declining sectors suffer in the long-run, no matter their degree of industrial specialization. Dynamic Jacobs externalities, by contrast, imply that specialized cities suffer in the long run, no matter their industries. This is the insight that allows us to disentangle these two mechanisms quantitatively in Section 6.

6 Quantitative analysis [TBC]

7 Concluding remarks

We have presented evidence that the extent of fragmentation of land at the fringes of cities in the early nineteenth century affected their growth rate and sectoral specialization over the course of the industrial revolution. The rise of international trade played an important role in accelerating the reshaping of the economic geography of England and Wales at this critical time. Armed with exogenous variation in the capacity of cities to grow, we further identify inter-temporal costs of the early success that arose over the twentieth century. Using our quantitative model, we suggest a mechanism: While many locations can indeed gain from being able to grow and specialize in response to structural transformation, this specialization comes at future costs as those locations fail to acquire the dynamic Jacobs externalities that sustain otherwise more diverse cities.

Future work will look to estimate the structural parameters of the model using the data. We will then be in a position to use the model to conduct counterfactual analysis to ask, for example, how much the growth in trade translated into faster urbanization and city growth. We can also consider whether real aggregate gains arose, or whether the trade-induced changes caused merely a reorganization of activity across space. In that vein, we can also ask about the role of policy in a manner that is relevant for modern economies. Even where attempts to replicate a ‘Silicon Valley’-type success story in different parts of the world by focusing on particular, growing sectors may be beneficial in the short-run, the long-run consequences of policy-driven specialization are clear from the costs borne by those cities in the formerly thriving hotspots of the industrial revolution that struggle to recover their previous advantages today.

References

- Allen, Robert C.**, *The British Industrial Revolution in Global Perspective*, Cambridge University Press, 2009.
- , “The Interplay among Wages, Technology, and Globalization: The Labor Market and Inequality, 1620-2020,” in “The Handbook of Historical Economics,” Elsevier, 2021, pp. 795–824.
- Allen, Treb and Dave Donaldson**, “Persistence and Path Dependence in the Spatial Economy,” Technical Report, National Bureau of Economic Research 2020.
- , **Costas Arkolakis, and Xiangliang Li**, “On the Equilibrium Properties of Network Models with Heterogeneous Agents,” Technical Report, National Bureau of Economic Research 2020.
- Bairoch, Paul and Gary Goertz**, “Factors of Urbanisation in the Nineteenth Century Developed Countries: A Descriptive and Econometric Analysis,” 1986, p. 21.
- Berkes, Enrico, Ruben Gaetani, and Martí Mestieri**, “Cities and Technological Waves,” Technical Report, mimeo 2021.
- Broadberry, S. N., B. M. S. Campbell, Alexander Klein, Mark Overton, and Bas van Leeuwen**, *British Economic Growth, 1270-1870*, New York: Cambridge University Press, 2015.
- Carlino, Gerald and William R Kerr**, “Agglomeration and innovation,” *Handbook of regional and urban economics*, 2015, 5, 349–404.
- Clark, Gregory**, “‘The Industrial Revolution’, in Aghion, P. and Durlauf, S.N. (eds.),” *Handbook of Economic Growth*, 2014. Elsevier.
- , **Kevin Hjortshøj O’Rourke, and Alan M. Taylor**, “The Growing Dependence of Britain on Trade during the Industrial Revolution,” *Scandinavian Economic History Review*, May 2014, 62 (2), 109–136.
- Crafts, Nicholas F. R.**, “British Industrialization in an International Context,” *Journal of Interdisciplinary History*, 1989, 19 (3), 415.
- Denman, Donald R.**, *Origins of Ownership: A Brief History of Land Ownership and Tenure in England from Earliest Times to the Modern Era*, George Allen & Unwin, 1958.
- Duranton, Gilles**, “Urban Evolutions: The Fast, the Slow, and the Still,” *American Economic Review*, March 2007, 97 (1), 197–221.
- and **Diego Puga**, “Nursery cities: Urban diversity, process innovation, and the life cycle of products,” *American Economic Review*, 2001, 91 (5), 1454–1477.

- Eckart, Wolfgang**, “On the land assembly problem,” *Journal of Urban Economics*, 1985, 18 (3), 364–378.
- Eckert, Fabian and Michael Peters**, “Spatial Structural Change,” Technical Report, mimeo 2018.
- Faggio, Giulia, Olmo Silva, and William C. Strange**, “Heterogeneous agglomeration,” *Review of Economics and Statistics*, 2017, 99 (1), 80–94.
- Fajgelbaum, Pablo and Stephen Redding**, “Trade, Structural Transformation and Development: Evidence from Argentina 1869-1914,” Technical Report w20217 2021.
- Franck, Raphaël and Oded Galor**, “Industrial Development and Long-Run Prosperity,” NBER Working Papers 23701, National Bureau of Economic Research, Inc 2017.
- Glaeser, Edward L.**, “Reinventing Boston: 1630–2003,” *Journal of Economic Geography*, 2005, 5, 119–153.
- Glaeser, Edward L, Hedi D Kallal, Jose A Scheinkman, and Andrei Shleifer**, “Growth in cities,” *Journal of Political Economy*, 1992, 100 (6), 1126–1152.
- Goose, Nigel**, “Regions, 1700–1870,” in Roderick Floud, Jane Humphries, and Paul Johnson, eds., *The Cambridge Economic History of Modern Britain*, second ed., Cambridge University Press, October 2014, pp. 149–177.
- Hanlon, W. Walker**, “The Rise of the Engineer: Inventing the Professional Inventor During the Industrial Revolution,” Technical Report, mimeo 2021.
- Harley, C. Knick and N. F. R. Crafts**, “Simulating the Two Views of the British Industrial Revolution,” *The Journal of Economic History*, 2000, 60 (3), 819–841.
- Heblich, Stephan, Marlon Seror, Hao Xu, and Yanos Zylberberg**, “Industrial clusters in the long run: evidence from Million-Rouble plants in China,” Technical Report, CESifo Working Paper No. 7682 2019.
- Henderson, Vernon, Ari Kuncoro, and Matt Turner**, “Industrial development in cities,” *Journal of political economy*, 1995, 103 (5), 1067–1090.
- Hennessy, Peter**, *Never Again Britain 1945-1951*, London [u.a.: Penguin Books, 2006.
- Herendorf, Berthold, Richard Rogerson, and Akos Valentinyi**, “Growth and Structural Transformation,” in Philippe Aghion and Steven Durlauf, eds., *Handbook of Economic Growth*, Vol. 2 of *Handbook of Economic Growth*, Elsevier, 2014, chapter 6, pp. 855–941.

Hills, Sally, Ryland Thomas, and Nicholas Dimsdale, “The UK Recession in Context — What Do Three Centuries of Data Tell Us?,” *Bank of England Quarterly Bulletin*, 2010, p. 15.

Hoskins, William G., *The Making of the English Landscape*, London: Hodder and Stoughton, 1988.

Hudson, Pat, “Industrial Organisation and Structure,” in Paul Johnson and Rodrick Floud, eds., *The Cambridge Economic History of Modern Britain: Volume 1: Industrialisation, 1700–1860*, Vol. 1, Cambridge: Cambridge University Press, 2004, pp. 28–56.

Jacobs, Jane, *The Economy of Cities*, New York: Vintage, 1969.

Kelly, Morgan, Joel Mokyr, and Cormac O’Grada, “The Mechanics of the Industrial Revolution,” June 2020.

Mendels, Franklin F., “Proto-Industrialization: The First Phase of the Industrialization Process,” *The Journal of Economic History*, March 1972, 32 (1), 241–261.

Mokyr, Joel, *The Enlightened Economy: An Economic History of Britain 1700–1850*, Yale University Press, 2009.

— , “Attitudes, Aptitudes, and the Roots of the Great Enrichment,” in “The Handbook of Historical Economics,” Elsevier, 2021, pp. 773–794.

Musson, Albert E., “Industrial Motive Power in the United Kingdom, 1800-70,” *The Economic History Review*, 1976, 29 (3), 415–439.

Nagy, Dávid, “Quantitative economic geography meets history: Questions, answers and challenges,” *Regional Science and Urban Economics*, 2021.

Neeson, Jeanette M, *Commoners: common right, enclosure and social change in England, 1700-1820*, Cambridge University Press, 1996.

Ngai, L. Rachel and Christopher A. Pissarides, “Structural Change in a Multisector Model of Growth,” *American Economic Review*, 2007, 97 (1), 429–443.

Nuvolari, Alessandro, Bart Verspagen, and Nick von Tunzelmann, “The Early Diffusion of the Steam Engine in Britain, 1700–1800: A Reappraisal,” *Clio-metrica*, October 2011, 5 (3), 291–321.

Ogilvie, Sheilagh, “Protoindustrialization,” in “The New Palgrave Dictionary of Economics,” London: Palgrave Macmillan UK, 2008, pp. 1–6.

Pierce, Justin R and Peter K Schott, “The surprisingly swift decline of US manufacturing employment,” *American Economic Review*, 2016, 106 (7), 1632–62.

Redding, Stephen J. and Esteban Rossi-Hansberg, “Quantitative Spatial Economics,” *Annual Review of Economics*, 2017, 9, 21–58.

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

Stokey, Nancy L., “A Quantitative Model of the British Industrial Revolution, 1780–1850,” *Carnegie-Rochester Conference Series on Public Policy*, December 2001, 55 (1), 55–109.

Strange, William C., “Information, holdouts, and land assembly,” *Journal of Urban Economics*, 1995, 38 (3), 317–332.

Trew, Alex, “Spatial Takeoff in the First Industrial Revolution,” *Review of Economic Dynamics*, 2014, 17, 707–25.

Trinder, Barrie, “Industrialising Towns 1700–1840,” in Peter Clark, ed., *The Cambridge Urban History of Britain: Volume 2: 1540–1840*, Vol. 2 of *The Cambridge Urban History of Britain*, Cambridge: Cambridge University Press, 2000, pp. 805–830.

Voigtländer, Nico and Hans-Joachim Voth, “The Three Horsemen of Riches: Plague, War, and Urbanization in Early Modern Europe,” *The Review of Economic Studies*, April 2013, 80 (2), 774–811.

Wrigley, E. Anthony, “The PST System of Classifying Occupations,” 2010. Paper 1 at www.hps.cam.ac.uk/research/projects/occupations/britain19c/papers/.

A Theory appendix

A.1 Derivation of Equation (11)

By free mobility across industries, nominal wages equalize across them in each city:

$$w_{ict} = w_{ct}$$

Plugging this result into Equation (5), we obtain total land rents in city c as

$$R_{ct}L_{ct} = r_{ct}H_{ct} = \frac{1-\gamma}{\gamma}w_{ct}L_{ct}$$

from which

$$w_{ct} + R_{ct} = \frac{1}{\gamma}w_{ct}. \quad (13)$$

Also, from Equation (4), we get

$$r_{ct} = \left(\frac{1-\gamma}{\gamma}\right)^{1/\zeta_{ct}} w_{ct}^{1/\zeta_{ct}} L_{ct}^{1/\zeta_{ct}}. \quad (14)$$

By perfect competition, the factory gate price of each variety c in industry i becomes equal to its marginal cost of production in equilibrium:

$$p_{ict} = T_{ict}^{-1} w_{ct}^\gamma r_{ct}^{1-\gamma} = \left(\frac{1-\gamma}{\gamma}\right)^{\frac{1-\gamma}{\zeta_{ct}}} T_{ict}^{-1} L_{ct}^{\frac{1-\gamma}{\zeta_{ct}}} w_{ct}^{\gamma + \frac{1-\gamma}{\zeta_{ct}}} \quad (15)$$

where we used Equation (14). As a result, we can write the price index of industry i , (7), as

$$P_{idt} = \left[\sum_{c=1}^C \left(\frac{1-\gamma}{\gamma} \right)^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} T_{ict}^{\epsilon-1} L_{ct}^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} w_{ct}^{-\left(\gamma + \frac{1-\gamma}{\zeta_{ct}}\right)(\epsilon-1)} \tau_{icdt}^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}} \quad (16)$$

and market clearing condition (6) as

$$w_{ct}L_{ict} = \left(\frac{1-\gamma}{\gamma} \right)^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} T_{ict}^{\epsilon-1} L_{ct}^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} w_{ct}^{-\left(\gamma + \frac{1-\gamma}{\zeta_{ct}}\right)(\epsilon-1)} \sum_{d=1}^C P_{dt}^{\sigma-1} P_{idt}^{\epsilon-\sigma} w_{dt} L_{dt} \tau_{icdt}^{1-\epsilon} \quad (17)$$

where we also used Equation (13).

By Equation (9), utility (real income) equalizes across cities in equilibrium, im-

plying

$$\bar{U}_t = \frac{w_{ct} + R_{ct}}{P_{ct}} = \frac{1}{\gamma} \frac{w_{ct}}{P_{ct}}$$

from which

$$P_{ct} = (\gamma \bar{U}_t)^{-1} w_{ct}. \quad (18)$$

Plugging this result into equations (8) and (17) and rearranging Equation (16), we obtain the system of equations

$$P_{ict}^{1-\epsilon} = \sum_{d=1}^C \left(\frac{1-\gamma}{\gamma} \right)^{-\frac{1-\gamma}{\zeta_{dt}}(\epsilon-1)} \hat{T}_{idt}^{\epsilon-1} L_{dt}^{-\left(\frac{1-\gamma}{\zeta_{dt}}-\alpha\right)(\epsilon-1)} w_{dt}^{-\left(\gamma+\frac{1-\gamma}{\zeta_{dt}}\right)(\epsilon-1)} \tau_{idt}^{1-\epsilon} \quad (19)$$

$$w_{ct}^{1-\sigma} = (\gamma \bar{U}_t)^{1-\sigma} \sum_{i=1}^I P_{ict}^{1-\sigma} \quad (20)$$

$$\left(\frac{1-\gamma}{\gamma} \right)^{\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} w_{ct}^{1+\left(\gamma+\frac{1-\gamma}{\zeta_{ct}}\right)(\epsilon-1)} L_{ct}^{1+\left(\frac{1-\gamma}{\zeta_{ct}}-\alpha\right)(\epsilon-1)} = (\gamma \bar{U}_t)^{1-\sigma} \sum_{i=1}^I \sum_{d=1}^C \hat{T}_{ict}^{\epsilon-1} P_{idt}^{\epsilon-\sigma} w_{dt}^{\sigma} L_{dt} \tau_{icdt}^{1-\epsilon} \quad (21)$$

where

$$\hat{T}_{ict} = T_{ict} f_i \left(L_{c,t-1}, \{L_{jc,t-1}\}_{j \in I} \right)$$

is the part of TFP that is exogenous in period t . Applying the change in variables in (12) and recalling $w_{ict} = w_{ct}$, Equation (11) immediately follows from (19), (20) and (21).

A.2 Proofs

Proof of Theorem 1. Treating an industry-location pair (i, c) as the equivalent of a location, one can see that (11) is (almost) a special case of the systems of equations studied in Allen et al. (2020). The only exception is that some of the K_{icjdt}^1 and K_{icjdt}^2 are zero, while Allen et al. (2020) assume that every term on the right-hand side of the system is strictly positive. However, it is straightforward to see that part (i) of Theorem 1 in Allen et al. (2020) applies even if some terms on the right-hand side of the system are zero, as long as the entire right-hand side is strictly positive, as in our case. Thus, as in part (i) of Theorem 1 in Allen et al. (2020), the solution to (11) exists and is unique if the largest eigenvalue of matrix \mathbf{A} is strictly less than one in absolute value, where \mathbf{A} is the 3-by-3 matrix such that

$$\left| \varepsilon_{icjdt}^{hh'} (x_{jdt}^1, x_{jdt}^2, x_{jdt}^3) \right| \leq (\mathbf{A})_{hh'}$$

and $\varepsilon_{icjdt}^{hh'}$ is the elasticity of the (j, d) term on the right-hand side of the h 'th equation in (11) with respect to $x_{jdt}^{h'}$. Looking at Equation (11), one can see that these elasticities are given by

$$\varepsilon_{icjdt}(x_{jdt}^1, x_{jdt}^2, x_{jdt}^3) = \begin{bmatrix} 0 & \frac{\alpha+\gamma}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \frac{\epsilon-1}{\sigma-1} & -\left[1 - \frac{1}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)}\right] \\ \frac{\sigma-1}{\epsilon-1} & 0 & 0 \\ \frac{\sigma-\epsilon}{\epsilon-1} & -\left[1 - \frac{\alpha+\gamma}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \frac{\epsilon-1}{\sigma-1}\right] & \frac{1}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \end{bmatrix}$$

and therefore matrix \mathbf{A} is given by

$$\mathbf{A} = \begin{bmatrix} 0 & \max_{dt} \left| \frac{\alpha+\gamma}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \frac{\epsilon-1}{\sigma-1} \right| & \max_{dt} \left| 1 - \frac{1}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \right| \\ \frac{\sigma-1}{\epsilon-1} & 0 & 0 \\ \left| \frac{\sigma-\epsilon}{\epsilon-1} \right| & \max_{dt} \left| 1 - \frac{\alpha+\gamma}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \frac{\epsilon-1}{\sigma-1} \right| & \max_{dt} \left| \frac{1}{1+(\frac{1-\gamma}{\zeta_{dt}}-\alpha)(\epsilon-1)} \right| \end{bmatrix}.$$

Thus, the solution to (11) exists and is unique if the largest eigenvalue of \mathbf{A} is strictly less than one in absolute value.

Once the solution to (11) exists and is unique, the change in variables in (12) can be uniquely inverted to express equilibrium price indices, wages and population levels as a function of this solution. Specifically,

$$P_{ict} = (x_{ict}^1)^{\frac{1}{1-\epsilon}}, \quad (22)$$

$$w_{ict} = w_{ct} = (x_{ict}^2)^{\frac{1}{1-\sigma}} \quad \text{for any } i \quad (23)$$

and

$$L_{ct} = \left(\frac{x_{ict}^3}{w_{ct}^{1+(\frac{1-\gamma}{\zeta_{ct}}-\alpha)(\epsilon-1)}} \right)^{\frac{1}{1+(\frac{1-\gamma}{\zeta_{ct}}-\alpha)(\epsilon-1)}} \quad \text{for any } i, \quad (24)$$

implying that the within-period equilibrium of the model exists and is unique. Finally, as Equation (10) characterizes the unique dynamic evolution of productivity as a function of within-period equilibria, the entire equilibrium of the model exists and is unique.

Proof of Theorem 2. Under the same sufficient condition as the one guaranteeing existence and uniqueness, part (i) of Theorem 1 in Allen et al. (2020) implies that the solution to (11) can be found by iterating on this system. That is, for any initial

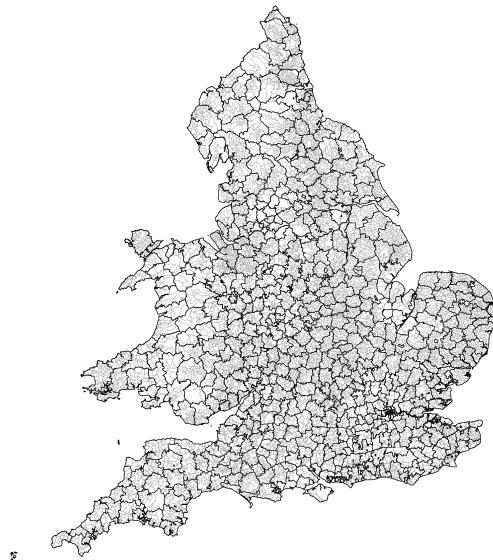
guess of $(x_{ict}^1, x_{ict}^2, x_{ict}^3)$, one can plug these values into the right-hand side of (11), update the guess by calculating the left-hand side, plug the new values into the right-hand side, and so on. Applying this process iteratively, the values of $(x_{ict}^1, x_{ict}^2, x_{ict}^3)$ are guaranteed to converge to the unique solution. Equilibrium prices, wages and population levels can then be obtained from Equations (22), (23) and (24). Finally, equilibrium sectoral employment levels can be obtained from Equation (17) as

$$L_{ict} = (\gamma \bar{U}_t)^{1-\sigma} \left(\frac{1-\gamma}{\gamma} \right)^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} \hat{T}_{ict}^{\epsilon-1} L_{ct}^{-\left(\frac{1-\gamma}{\zeta_{ct}}-\alpha\right)(\epsilon-1)} w_{ct}^{-[1+(\gamma+\frac{1-\gamma}{\zeta_{ct}})(\epsilon-1)]} \sum_{d=1}^C P_{idt}^{\epsilon-\sigma} w_{dt}^\sigma L_{dt} \tau_{icdt}^{1-\epsilon}. \quad (25)$$

B Data appendix

B.1 Data sources

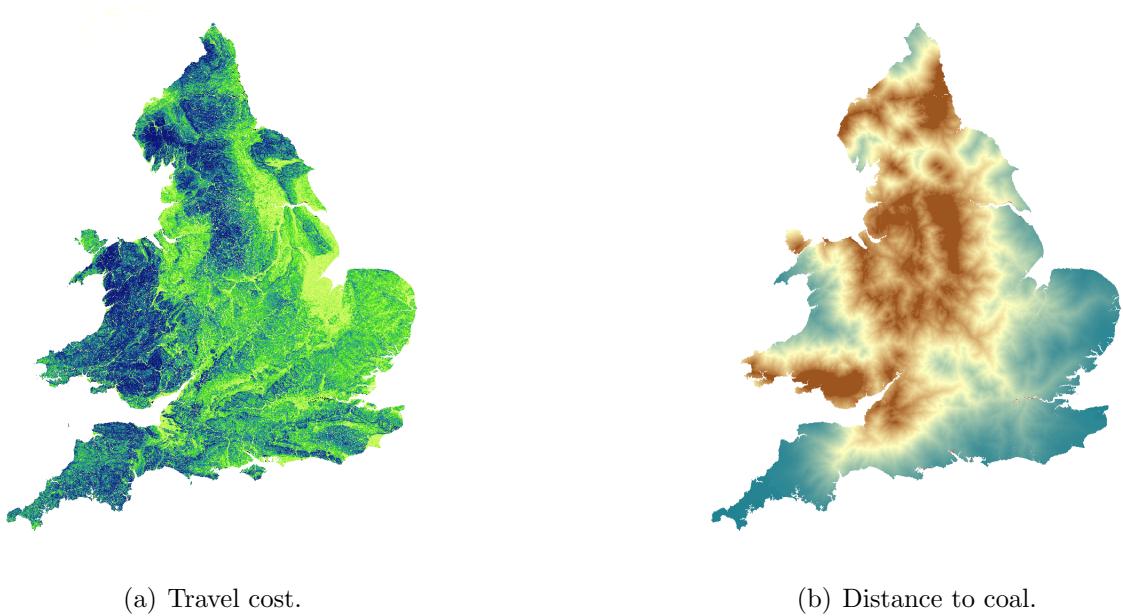
Figure B1: Consistent parishes across England and Wales.



Notes: This Figure displays the output of the transitive closure algorithm implemented by the Cambridge Group for History of Population and Social Structure. Consistent mappable units based on parishes are displayed in gray; registration districts are displayed with black borders.

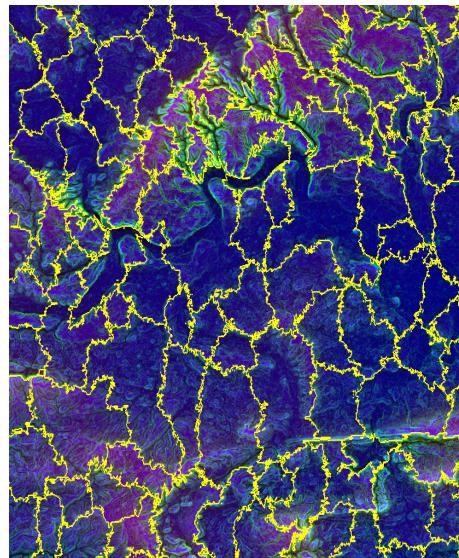
B.2 Data construction

Figure B2: Travel cost and distance to coal across England and Wales, as computed around 1817.



Notes: The left panel displays the raster of transport costs as calculated using the transport network at the beginning of the 19th century, and a penalization accounting for the local elevation gradient (yellow: low, green: medium, blue: high). The right panel displays the minimum travel time from the nearest coal field (red: low, blue/green: high).

Figure B3: Example of the Quickshift output.



Notes: This Figure shows the output of our image segmentation algorithm which groups pixels by their proximity in actual space (i.e., along the physical distance) and in the space as defined by topographic and soil characteristics. More specifically, the image segmentation algorithm is a “Quickshift” procedure disciplined by three parameters: a scale parameter (the size of a typical farm), a maximum physical distance which disciplines the extent to which the algorithm looks for neighbors, and a relative weight between distance in the multi-bands-space and physical distance.