The Death and Life of Great British Cities *

Stephan Heblich, Dávid Krisztián Nagy, Alex Trew, Yanos Zylberberg

Discussion Paper 25/797

15 July 2025

School of Economics

University of Bristol
Priory Road Complex
Bristol
BS8 1TU
United Kingdom

University of BRISTOL
School of Economics

The Death and Life of Great British Cities*

Stephan Heblich Dávid Krisztián Nagy Alex Trew
Yanos Zylberberg

July 15, 2025

Abstract

Does industrial concentration shape the life and death of cities? We identify settlements from historical maps of England and Wales (1790–1820), isolate exogenous variation in their late 19th-century size and industrial concentration, and estimate the causal impact of size and concentration on later dynamics. Industrial concentration has a negative effect on long-run productivity—independent of industry trends and consistent with cross-industry Jacobs externalities. A spatial model quantifies the role of fundamentals, industry trends, and Jacobs externalities in shaping industry-city dynamics and isolates a new, dynamic trade-off in the design of place-based policies

Keywords: Jacobs externalities; cities over time; quantitative economic geography.
JEL codes: F63; N93; O14; R13.

*Heblich: University of Toronto, CESifo, IfW Kiel, NBER; stephan.heblich@utoronto.ca; Nagy: CREi, Universitat Pompeu Fabra, BSE, CEPR; dnagy@crei.cat; Trew: University of Glasgow, CESifo, IZA; alex.trew@glasgow.ac.uk; Zylberberg: University of Bristol, CEPR; yanos.zylberberg@bristol.ac.uk. We thank Gabriel Ahlfeldt, Pierre-Philippe Combes, Kerem Cosar, Don Davis, Jonathan Dingel, Gilles Duranton, Fabian Eckert, Pablo Fajgelbaum, James Feigenbaum, Ed Glaeser, Laurent Gobillon, Gordon Hanson, Miklos Koren, Hans Koster, Oleg Itskhoki, Guy Michaels, Eduardo Morales, Henry Overman, Javier Quintana, Steve Redding, Jean-Marc Robin, Michele Rosenberg, Esteban Rossi-Hansberg, Olmo Silva, Robert Staiger, Daniel Sturm, Lin Tian, Elisabet Viladecans-Marsal, Pablo Warnes, David Weinstein, as well as participants in seminars and conferences, e.g., at HU Berlin, Birmingham, Bocconi, Bologna, Bristol, BU, CERGE-EI, Chicago Fed, CRED, CREi, CURE, Dortmund, the EEA meetings, FREIT, Georgetown, HEC, Helsinki, INSEAD, LSE, McGill, NBER SI, Oregon, Oslo, Penn State, Philadelphia Fed, Princeton, PSE, Reading, Richmond Fed, RWI, Sciences-Po, SED, Sheffield, SMU, Toronto, Tufts, UCL, UCLA, UEA, UPF, USC, Virginia, Washington, Warwick, the World Bank, and York for useful comments. We would also like to thank Clement Gorin for invaluable advice and the Cambridge Group for the History of Population and Social Structure, the British Library, Alexis Litvine and Gethin Rees for their help with data. Heblich and Zylberberg acknowledge support from the ANR/ESRC/SSHRC, through the ORA grant ES/V013602/1 (MAPHIS: Mapping History).

Many of the cities and regions that drove the industrial transformation of the nineteenth century have since declined in the twentieth century. The formerly thriving towns of Lancashire, the rust belt cities in the northeastern United States, or the Ruhr valley in Germany grew rapidly and employed generations of workers but have struggled to find continued economic success in the longer run. Both the rise of those cities and their reversal of fortune may be tied to the particular dynamics of their industries. One explanation is *external* to the cities: macroeconomic factors of demand and technological change drive the dynamics of industries within a country. The dynamics of cities are then simply tied to the aggregate dynamics of industries in which they have a comparative advantage. A second force relates to factors that are *internal* to the city: a city's industrial portfolio affects its long-run development, even when accounting for external industry trends. For instance, Glaeser et al. (1992) discusses the role of long-run, between-industry externalities in driving city growth—as motivated by the seminal work of Jacobs (1961, 1969) on *The Death and Life of Great American Cities*.

This paper focuses on internal drivers within cities and studies whether industrial structure shapes their life and death. We rely on unique data which characterizes the evolving spatial distribution of employment and industries in England and Wales over the course of two centuries. More specifically, (i) we identify settlements in the early nineteenth century from historical maps, isolate clusters of built-up areas, and delineate the footprint of potential cities; (ii) we follow the composition of these settlements during the period of rapid urban and industrial expansion in the nineteenth century using the first near-exhaustive picture of industries across Britain (a quasi-census based on baptism records, around 1817) and a micro-census when the structure of cities stabilizes (1881); and (iii) we measure their long-run economic performance using an array of high-quality contemporary data, including firm-level production information. We find a strong, negative effect of industrial concentration in 1881 on later city productivity, irrespective of secular industrial trends. We use this empirical evidence to motivate and estimate a quantitative spatial model of industries and cities, where heterogeneous cities produce and trade goods. Dynamics in the model arise from both city-specific externalities (à la Jacobs) and aggregate, exogenous industry trends. We find that the interaction of these dynamic forces with the early specialization of British cities explains a large share of current spatial inequalities. The present-day North/South productivity gap would be 40% lower in the absence of Jacobs externalities. The industrial concentration of British cities that arose during the nineteenth century plays a powerful role in long-term regional disparities because of these long-run externalities.

Britain provides the ideal setting to study this question. Through the course of the nineteenth century, significant macroeconomic changes included the adoption of labor-

saving technologies, the shift to large-scale, steam-powered production, and the rapid rise in domestic and international trade. These factors transformed the scale of cities and industries which in turn underpinned a sustained increase in economic growth. Alongside this growth, the (mostly) small settlements of the early nineteenth century gave way to a dense network of larger cities. By the end of the nineteenth century, the cities of England and Wales had emerged into markedly different economic structures: some cities had developed a diverse industrial base, while others had specialized in a narrow set of industries; some cities had grown larger, while others had not.

To make progress, we need: the early (pre-expansion) settlements and their economic composition; a way to understand what drives the heterogeneity in the manner of the settlement growth and industrial concentration up to the end of the nineteenth century; and modern data on productivity of those places up to the present day. The first empirical challenge is to locate and delineate potential cities of the early nineteenth century. We collect unique historical maps from around 1790–1820, develop a machine-learning algorithm to identify buildings, and use a delineation procedure to select urban settlements as sufficiently large clusters of built-up areas (De Bellefon et al., 2021). The output is a new dataset of early nineteenth century settlements and their boundaries that each had the potential to develop into cities by the late nineteenth century.[1]

Having established the early nineteenth century settlements, a second challenge is to isolate exogenous variation in the nature of their rise during the nineteenth century. Doing so will permit us to identify the causal impact of late nineteenth century industrial diversity and size on the long-run dynamics of cities. To predict industrial concentration, we rely on a measure of initial location advantages—the 2-digit industrial composition of early settlements in 1817—and create a "shift-share industrial Herfindahl index" which combines the initial city-sectoral shares with the aggregate growth of each sector. To predict late-nineteenth century city size, we borrow insights from the historical literature and conjecture that the fragmentation of land ownership in the immediate fringe of cities affected the pace at which they could grow in response to industrialization. The underlying argument is at the heart of the land assembly problem (see, e.g., Eckart, 1985; Strange, 1995): a higher degree of land ownership fragmentation in the city's fringe makes negotiations to acquire additional land for urban use more costly. We develop an algorithm to predict natural fault lines—multidimensional breaks in elevation, ruggedness, persistent soil attributes and water bodies—between potential agricultural land parcels, and

[1] A related challenge is to define the geography of cities over time and nest economic activity within these geographies. In our baseline strategy, we expand city boundaries uniformly (across cities, and across directions for each city) to match the average city growth over the nineteenth century, and we nest economic characteristics from the end of the nineteenth century onward at this level. An alternative procedure would consist in using the actual, "endogenous" city boundaries, e.g., as measured by historical maps around 1890–1900 and recent built-up data—such a procedure does not affect our empirical findings.

we compute the distribution of "natural farms" over England and Wales.[2] We use the density of such "farms" around each of the early settlements as an exogenous proxy for land ownership fragmentation (controlling for the *average* ruggedness and soil quality around city boundaries, see Saiz, 2010; Harari, 2020).[3]

We find that cities which specialize in a smaller number of industries at the end of the nineteenth century experience a decline in the long run. An additional 0.10 in the 2-digit industrial Herfindahl index in 1881—about 60% of a standard deviation—causes: an additional 3.7 percentage points in unskilled employment; 27% lower returns to labor; and a 24% lower firm productivity around 2010-2020, conditional on 4-digit industry fixed effects, the local share of turnover in the firm industry, and firm-specific factor use. By contrast, we find that a city's size in 1881 is less economically significant in the long-run—consistent with Gibrat's law. These results are obtained in a baseline specification which already controls for: aggregate industry dynamics; the composition of industrial employment around 1820; topography and the geography of development in these early times; transportation infrastructure, and access to resources and markets; agricultural productivity and exposure to the Grain Invasion (Coeurdacier et al., 2022; Heblich et al., 2024); the marginal land supply from the end of the nineteenth century onward; and initial built-up density at the city fringe. We subject these results to an array of robustness checks that include: adding controls for historical agricultural production, land values, and parliamentary enclosures (Heldring et al., 2022); controlling for the (endogenous) employment shares in agriculture, services, and manufacturing in 1881 to show that our findings are not driven by structural transformation; excluding 2-digit industries from our industrial concentration indices to show that our results are not driven by any sub-sectors in particular; providing alternative shift-share designs (e.g., leaving out the surrounding of the city to derive the "shifts"); showing that Jacobs externalities are local and operate within a 20 kilometer buffer; and assessing the extent of treatment

[2]We validate the predictive power of this measure of land fragmentation by comparing it with the actual concentration of ownership from micro-census records where land acreage is reported by landowners.

[3]Identification requires the two instruments to be relevant (H1) and exogenous (H2). (H1) Relevance justifies our effort to find independent sources of variation with sufficient explanatory power for our two endogenous variables: industrial concentration; and city size. Indeed, we show that geography is mostly predictive of city size, while industrial concentration is mostly explained by the nature of initial cottage industries. Our first stage reflects this observation: local topography constrains city growth without any effect on industrial diversity; and our predicted Herfindahl index mostly predicts industrial concentration. (H2) To support the exogeneity assumption, our empirical analysis controls for a large set of variables capturing: first-nature geography (e.g., initial area, elevation, slope, bulk density, agricultural productivity, soil characteristics); second-nature geography (e.g., latitude, market access, travel time to main cities, travel time to resources, waterways, transportation infrastructure); baseline city composition (deciles of employment, industrial concentration, agricultural employment shares in 1817); "fringe" variables (e.g., built-up density and predicted land ownership fragmentation within the contours of the settlement around 1800 and within a large 10 kilometer buffer); an industry-based (linear) shift-share or industry fixed-effects to clean for the later, aggregate evolution of industries; and fixed effects at the level of 39 counties.

heterogeneity—the interacted effect of industrial concentration and city size remains limited, and cities with higher share of services are *slightly* less affected by industrial concentration. The negative effect of industrial concentration is robust and very stable.

To interpret these empirical findings, we develop a multi-sector dynamic spatial model and study the mechanisms behind the evolution of industries across cities and over time. 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 amenities, their land supply elasticities, and their trade costs with other cities. While the model can theoretically accommodate for a flexible structure in how employment and industrial structure influence productivity, e.g., agglomeration externalities, Marshall-Arrow-Romer externalities, and long-run Jacobs externalities (see Carlino and Kerr, 2015, for a review about important determinants of city growth), our model quantification will only be able to credibly parametrize agglomeration externalities and estimate long-run Jacobs externalities while conditioning on contemporaneous Marshall-Arrow-Romer externalities.[4] First, the model rationalizes the distribution of economic activity at the end of the nineteenth century: trade costs are low, leading to a specialization of cities in their comparative advantage sectors; and the availability of land disciplines the degree to which population reallocates towards these cities. Second, (a) aggregate industry trends and (b) long-run externalities imply that this late nineteenth century distribution of employment across industries and cities leads to markedly different city dynamics into the twentieth century. (a) Some industries that were successful enter a period of decline over their life-cycle, e.g., because of structural change (Ngai and Pissarides, 2007) or international competition (Pierce and Schott, 2016), and cities with a location advantage in originally flourishing but later floundering industries simply suffer by dint of their reliance in that sector. (b) Long-run Jacobs externalities generate productivity gains in historically more diverse cities. This force directs economic activity away from cities that were specialized in the past, irrespective of the aggregate trends in each particular sector. Both forces (a) and (b) may rationalize the decline of a place, but they have very different implications and it is crucial to separately identify their contributions.

In the final part of the paper, we estimate our quantitative model using several key inputs: city-industry employment data; trade costs and industry-specific factor shares in production; elasticities estimated by previous literature, including elasticities of substitution within and across industries, static agglomeration externalities, and migration elasticities; city-specific land supply elasticities (Drayton et al., 2024); and our empirical, exogenous variation to identify long-run Jacobs externalities. We use the estimated

[4]Jacobs externalities operate across industries and benefit all industries within cities with a diverse industrial base. By contrast, Marshall-Arrow-Romer externalities operate within industry.

model to perform two counterfactual experiments that quantify the roles of location fundamentals, industry life-cycles, and local externalities in shaping urban dynamics. First, we assess the contribution of long-run Jacobs externalities to current regional inequality in England and Wales: Eliminating these externalities would reduce the North-South productivity gap by nearly half—from about 35% to 20%. This spatial inequality reflects the interaction of persistent Jacobs externalities and the historically high industrial concentration of the nineteenth century, driven by trade openness and transportation improvements. Second, we evaluate stylized place-based industrial policies that alter the degree of industrial diversity—representing specialization-inducing or diversification-promoting local interventions—and quantify their normative implications.[5] This exercise makes salient the existence of a dynamic trade-off faced by local policymakers: while specialization yields short-term productivity gains by reinforcing comparative advantages, it may harm long-run outcomes by reducing industrial diversity. In a range of reasonable discount factors, both diversification and specialization policies may appeal to policymakers, but the latter are preferable to impatient ones. In practice, many place-based policies promote specialization—reflecting the preferences of policymakers who effectively assign lower weight to future generations.

The main contributions of our study are to estimate Jacobs externalities, to isolate their role in driving regional inequalities, and to discuss a novel trade-off in allocating economic activity across space—a dynamic trade-off based on industrial composition (in stark contrast with the static trade-off between agglomeration benefits and crowding costs discussed, among others, in Henderson, 1974; Ahlfeldt et al., 2015; Brinkman, 2016). These findings have direct implications for the design of place-based policies, which must weigh not only static efficiency considerations but also long-run consequences tied to local industrial structure. In their study of the Tennessee Valley Authority, Kline and Moretti (2014a) show that place-based interventions may be justified on efficiency grounds if they correct meaningful local market failures; otherwise, they risk being distortionary, with local gains offset by national losses. We complement this view by introducing an intertemporal dimension and show that more patient planners would favor diversification policies, while less patient ones would prefer specialization.[6]

[5] Our theory abstracts from formally modeling innovation or other possible interactions between firms of similar or different industries. For this reason, we cannot properly evaluate existing place-based industrial policies such as the "Smart Specialisation Strategy" implemented by the European Union or the innovation clusters (or hubs) in the United States.

[6] Our focus on the long-run impact of industrial structure entails a few limitations. First, our theoretical framework abstracts from detailed firm-level interactions and thus does not capture the full range of contemporary agglomeration effects. This reflects a deliberate emphasis on long-term structural trade-offs, rather than within-period market dynamics. Second, our empirical analysis identifies a reduced-form, causal relationship between industrial concentration and long-run productivity. While multiple channels may drive this relationship, their separate identification is left for future work.

We relate to several strands of existing research. First, our study adds to the literature discussing industries as drivers of urban growth (Duranton, 2007; Hanlon and Miscio, 2017), and specifically to research on the negative effects of industrial concentration (Glaeser et al., 1992; Duranton and Puga, 2001; Lin, 2011; Faggio et al., 2017; Heblich et al., 2019). The closest paper to ours is Henderson et al. (1995), which discusses the interconnected life-cycle of industries and cities. One feature of our model, responsible for the negative effect of industrial concentration, is the role played by externalities à la Jacobs (1969) as drivers of knowledge development in the long run. These dynamics bear directly on the normative evaluation of place-based policies, where the spatial allocation of activity is shaped by local spillovers and long-run externalities (e.g., Fajgelbaum and Gaubert, 2025; Kline and Moretti, 2014b; Gaubert et al., 2025). In a related application, Glaeser (2005) studies Boston over nearly four centuries and points to the role of human capital in reinventing the city after periods of crisis and decline; distortion to the acquisition of human capital is also a key mechanism behind the rise and fall of cities in Franck and Galor (2021). Relative to these earlier contributions, an advantage of our empirical design is to causally identify the long-run causal impact of industrial concentration, while also accounting for the life cycle of nationwide industries themselves.

Second, our work is inspired by theories of agglomeration economies—recently reformulated as sharing, matching, and learning effects, operating through production linkages, labor markets or knowledge creation (Duranton and Puga, 2004)—but mostly relates to their empirics (Combes and Gobillon, 2015). Our empirical approach focusing on Jacobs externalities borrows from previous literature in its measurement of industrial concentration, but markedly differs in its time horizon: we relate contemporary productivity to industrial structure at the end of the nineteenth century, thus capturing the long-run effects of localized knowledge spillovers (Carlino and Kerr, 2015).

Third, we connect to a quantitative literature studying the dynamic evolution of economic activity across space. The closest contributions in this literature are Allen and Donaldson (2020), Berkes et al. (2021), Caliendo et al. (2019), Nagy (2023), Eckert and Peters (2023) and Fajgelbaum and Redding (2022).[7] Our main contribution to this literature lies in proposing a multi-sector dynamic model with various dimensions of heterogeneity that can be taken to the data in a computationally tractable way.

Finally, our period of interest includes a key transformation of Britain—the industrial revolution—accompanied by urbanization, trade, and structural transformation. Trade has been discussed as a key mechanism behind the industrial revolution in a number of contributions, notably Stokey (2001) or Allen (2009); it accelerated the transition to large-

[7] We refer the interested reader to Redding and Rossi-Hansberg (2017) for a comprehensive survey of the quantitative spatial literature and to Nagy (2022) for their use in addressing historical questions.

scale, export-oriented (urban) growth, and induced the invention of the labor-saving technologies that underpinned Britain's transformation. In our model, external demand reallocates economic activity and industries across cities (thus relating to Heblich et al., 2024, about the Grain Invasion and the resulting heterogeneous structural transformation across space). Lastly, a macroeconomic literature, summarized in Herrendorf et al. (2014), studies growth and structural transformation. Our quantitative spatial model abstracts away from the mechanisms discussed in theories of structural change and considers relative prices across industries as given (in contrast with Eckert and Peters, 2023).

The remainder of the paper is organized as follows. Section 1 presents some historical context. Section 2 describes our data. Section 3 establishes empirical facts which motivate the structure of the model, described in Section 4. Finally, Section 5 takes the model to the data and discusses mechanisms behind the long-run dynamics of cities.

1 Context

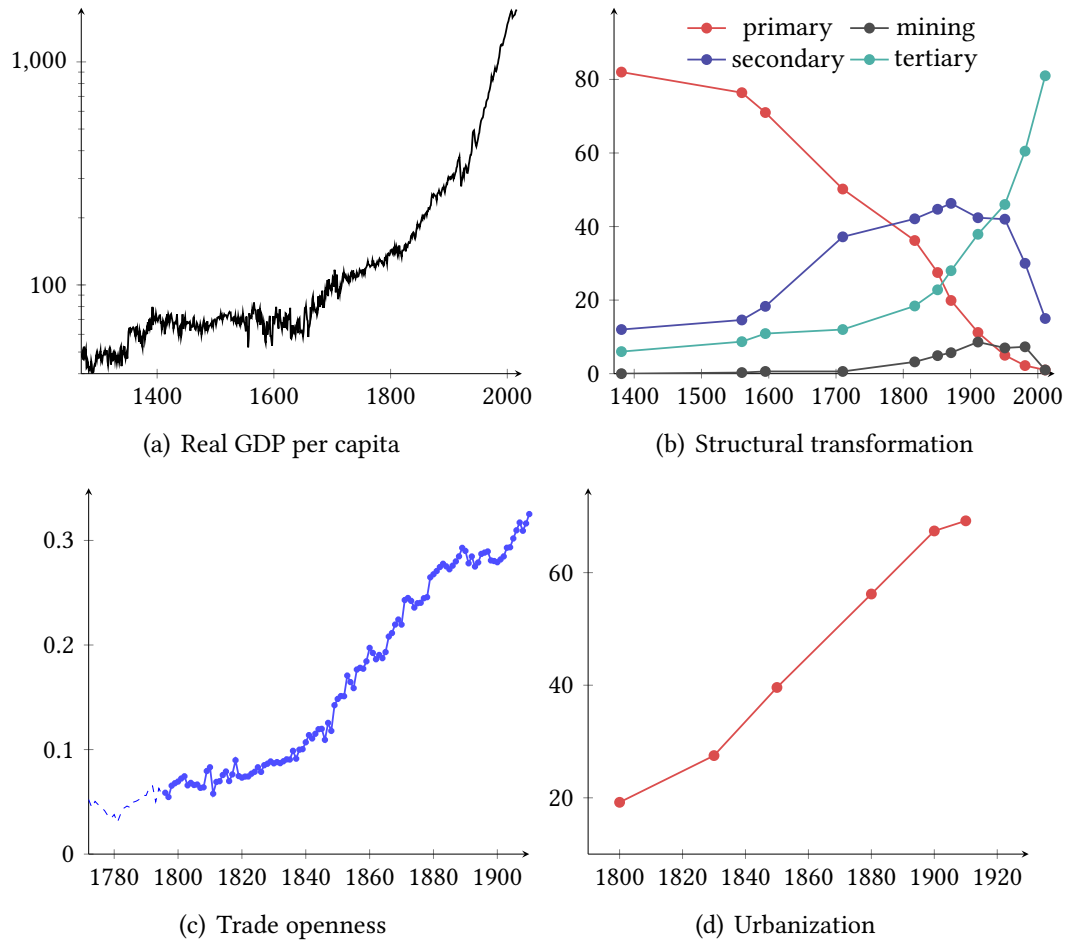
This section provides a brief overview of the context in which cities of our sample grow over the course of the nineteenth century: the industrial revolution in Britain, its early cottage industries, and the rise of urban agglomerations from early settlements.

The industrial revolution in Britain 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 income per capita; (b) a declining share of employment in 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 what are the factors that caused Britain to industrialize first, is not completely settled (see, e.g., the survey in Clark, 2014).[8]

Figure 1 depicts trends of these four dimensions. While many key industrial technologies emerge in the mid-eighteenth century, growth in output per capita accelerates particularly in the early nineteenth century (panel a), at a time when overall population

[8]Two highly influential hypotheses on the causes of the industrial revolution are formulated in Mokyr (2009) and Allen (2009). In Mokyr (2009), the industrial revolution is driven by the emergence of "attitudes" (a respect for entrepreneurs and inventors) and "aptitudes" (the growth of useful human capital, see Mokyr, 2021). Kelly et al. (2023) and Hanlon (2022) corroborate this view with a focus on mechanical workers and the professionalization of invention through the emergence of engineers. By contrast, in Allen (2009), (openness to) trade changes demand for the manufactured output, causing a shift in modes of production away from rural, low-scale, domestic-oriented and water-powered production to urban, specialized, export-oriented and large-scale factories in which steam power dominates (a change discussed in Crafts, 1989). More specifically, high wages due to external demand drive capital-biased, labor-saving technical change which fosters the growth in export-oriented industries (Allen, 2021). Stokey (2001) finds that trade explains all the decline in agricultural production, and a significant share of the increase in manufacturing and real wages (a finding qualitatively supported by Harley and Crafts, 2000; Clark et al., 2014).

Figure 1. The industrial revolution in Britain.



(a) Real GDP per capita

(b) Structural transformation

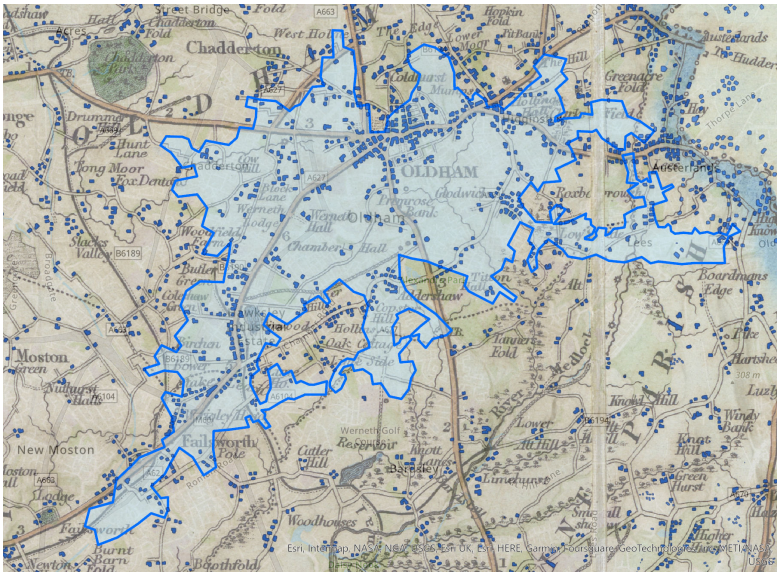
(c) Trade openness

(d) Urbanization

Notes: Panel (a) provides data on real GDP per capita from Broadberry et al. (2015) and national accounts, collated in Thomas and Dimsdale (2018). In panel (b), employment shares are classified according to the PST system—primary, secondary, tertiary—described in Wrigley (2010), which are respectively (and very broadly) agriculture, manufacturing/construction and services; we also separate out mining. We report the available data for male adults in England and Wales (Shaw-Taylor and Wrigley, 2014). In panel (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). In panel (d), urbanization is defined as the share of total population in cities over 5,000 inhabitants (Bairoch and Goertz, 1986).

was also growing rapidly. The share of employment in the secondary sector reaches a peak around 1871, having only marginally grown since 1710; and structural transformation then materializes in a swift rise in tertiary employment (panel b). Most striking and less commonly known are the dramatic changes in trade openness, which accelerates after 1820 (panel c), and urbanization, which grows by nearly fifty percentage points between 1820–1900 (panel d). During the nineteenth century, Britain grows from being a relatively rural economy to urbanization levels that are higher than its European comparators. We shed light on the patterns of such urbanization in the next section.

Figure 2. The urbanization of Oldham.

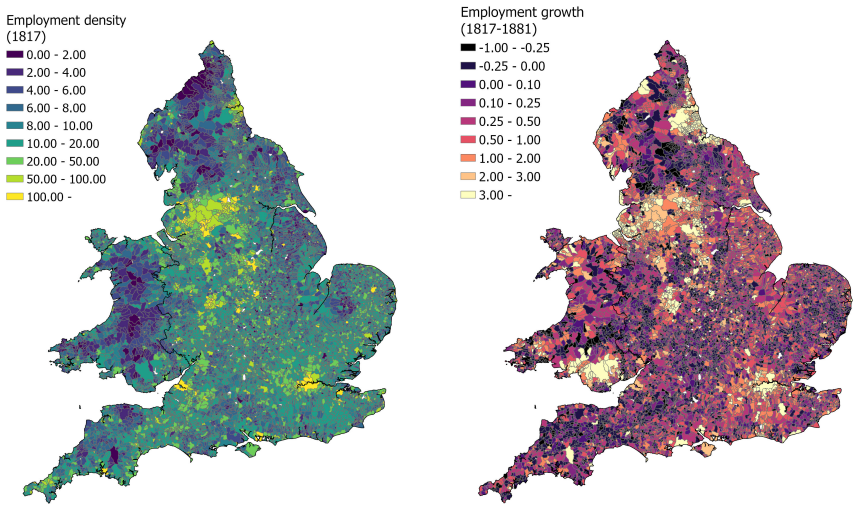


Notes: The underlying map is a county map of Lancashire around 1820 where built-up is indicated by darker blue rectangles (see Section 2 for a description of the built-up extraction process). The lighter blue area shows the urban boundaries of Oldham around 1880–1900, as defined by areas of contiguous built-up.

Proto-industrialization, the urbanization of Britain, and the rise of cities The shift of industrial production from a rural to an urban setting, and the concomitant growth in agglomerations, has been studied extensively. In the early eighteenth century, the low-scale artisanal home-production of finished manufactures gave way to well-organized rural industries, some of which exported beyond the locality (Hudson, 2004; Goose, 2014). This phenomenon is sometimes labeled "proto-industrialization" (Mendels, 1972; Ogilvie, 2008). The existence of extensive proto-industries rationalizes high employment shares in the secondary sector between 1700–1820 in spite of low urbanization levels (Figure 1). Factory production is, at that time, frequently rural, relying on water power and with rural workers housed by entrepreneurs around production facilities (Trinder, 2000). While steam engines begin to proliferate in the eighteenth century (Nuvolari et al., 2011), the transition to steam engines as the predominant motive power, and the associated large-scale, city-centered factory production, is not complete until the mid-nineteenth century (Musson, 1976). One illustration of the swift concentration of manufacturing production in urban centers is Oldham, in the North-East of Manchester. At the end of the eighteenth century, Oldham does not exist yet as a town: there are numerous hamlets where cottage industry takes place, including one called Oldham. By the end of the nineteenth century, Oldham is a factory town of about 140,000 inhabitants which produces a significant share of the whole country's textile output (almost a quarter of cotton production in 1911). Figure 2 provides the geographical illustration of such

rapid urbanization: the underlying map is from around 1820, with the 1820 built-up areas (dark blue polygons) overlaid with the urban boundary of Oldham around 1880–1900.

Figure 3. The geography of employment and growth in Britain.



(a) Employment density (1817) (b) Employment growth (1817–1881)

Notes: These two maps represent employment density in 1817 (panel a, as computed from a quasi-census of baptism records conducted between 1813–1820, and in number of recorded workers per square kilometers) and employment growth between 1817–1881 (panel b, where population in 1881 is calculated using micro-census records). The geographic unit is a parish (see Section 2).

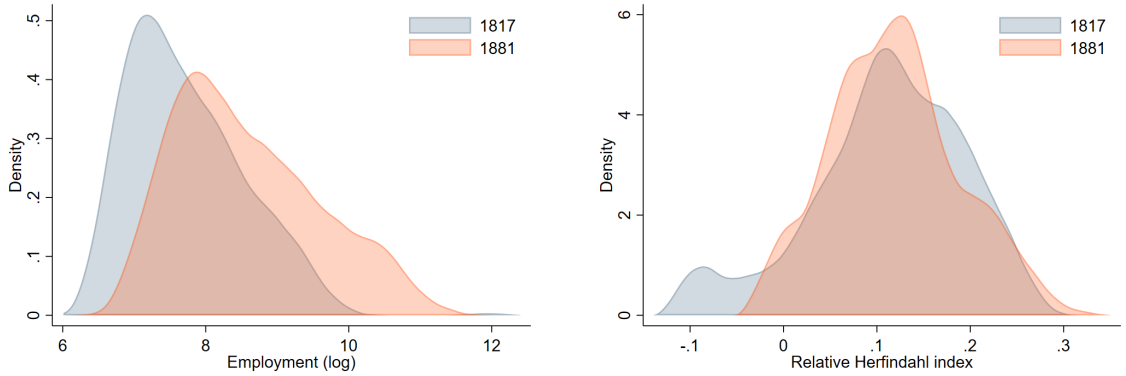
The example of Oldham is typical of a phenomenon whereby smaller settlements consolidate into towns, and small towns grow into cities with high concentration of manufactures. Stoke-on-Trent is another example: early potteries became boroughs, boroughs grew into towns, which would be amalgamated into Stoke-on-Trent—a leading producer of ceramics. This pattern is visible within each county of England and Wales. There is, however, a more aggregate geography of employment and growth that is worth discussing. As shown in panel (a) of Figure 3, employment is highly concentrated in a few regions in the early nineteenth century: Lancashire, North Yorkshire, the West Midlands, Northumberland, London and a few isolated cities (e.g., Bristol). The geography of employment growth is slightly more balanced across space (panel b).[9]

Settlements of the early nineteenth century give way to a dense network of cities by the end of the century. We shed some light on the extent and nature of this growth in Figure 4. In panel (a), we show the distribution of employment across cities in 1817 and in 1881, and we see that there is a massive population increase: cities grow by a

[9]While Lancashire, North Yorkshire, the West Midlands, Northumberland and London grow into large, densely populated regions, the South of Wales also grows significantly (due to a location advantage and proximity to coal). In other regions, growth is more unevenly distributed with a few cities concentrating most of the (new) workforce.

factor of 4 on average (while overall population increases by about 80% during the same period). In panel (b), we show the distribution of industrial concentration, as captured by a 2-digit industry-based Herfindahl index of employment. Cities are mechanically more specialized than the overall economy, but there exists large variation in the degree to which they specialize. We will see later that this variation is tied to the portfolio of location advantages that they hold across the different industries. In particular, trade induced a specialization in industries where cities hold a comparative advantage.

Figure 4. Specialization and growth in cities.



(a) Employment

(b) Herfindahl index

Notes: These two figures represent the distribution of employment (panel a) and industrial concentration (panel b) across 435 potential cities of England and Wales (see Section 2 for a definition of these potential cities and their boundaries). The distributions are shown in 1817 (blue) and in 1881 (orange). Specialization is captured by a Herfindahl index calculated as the difference between the city-specific Herfindahl index and a national one, $\tilde{h}_c = \sum_j s_{jc}^2 - \sum_j s_j^2$, where s_{ic} is the employment share in city c and industry i and s_i is the nationwide employment share in industry i. Industries are captured at the 2-digit level, based on the PST system—primary, secondary, tertiary—described in Wrigley (2010).

Land supply and the local geography of urban sprawl So far, we have described the determinants of urbanization and growth without discussing possible geographic constraints. During the nineteenth century, industrialization leads to rising demand for space within, then around, cities.[10] While factory production using new technologies concentrates in growing cities (see, e.g., Trew, 2014), some locations face difficulties in meeting the accelerating external demand after the 1820s, because of land constraints. Such constraints could be geographic but they could also be the remnants of historical property rights that present a land assembly-type problem (Denman, 1958; Hoskins, 1988; Eckart, 1985; Strange, 1995; Neeson, 1996; Hudson, 2004; Heldring et al., 2022). As Trinder (2000) notes, the "pattern of urban industrial growth in Britain before 1840 was

[10] "Land inside the older towns was acquiring a scarcity value [...] Open spaces inside the older towns vanished rapidly..." (Hoskins, 1988, p.185); "Manufactures ran up their mills, factories and works on the edge of existing towns[...]" (Hoskins, 1988, p.183).

untidy" (p.827). Where it was possible, towns were laid out in a grid pattern and experienced rapid growth. Where it was not, "[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). We rationalize this observation next and explore how land ownership patterns at the city's fringe shape the potential for urban expansion.

2 Data

This paper combines data on the evolving economic geography of England and Wales over 200 years. In this section, we present our data sources and describe the identification of early urban settlements.

2.1 Data sources

Census of England and Wales The main data source for population and employment is the Census of England and Wales, which provides a highly detailed 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) Anglican baptism records collected between 1813 and 1820 (referred to below as 1817) in order to retrieve consistent 2-digit industrial composition at the parish level before the time of rapid industrialization (Shaw-Taylor and Wrigley, 2014). One issue with census data is that the smallest 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 transitive closure of the different parishes covering the same points over time (see Appendix A.1).

Firm data We access, through a secure server, high-quality contemporary data, most notably firm-level production information: the Annual Business Survey (ABS), which reports turnover, purchases, employment costs, capital expenditure, stocks, and, value added for a representative selection of firms (about 2.5% of all firms); the Business Register and Employment Survey (BRES), which reports detailed employment for a representative selection of firms (about 4% of all firms); and, the Business Structure Database (BSD), which reports aggregate employment and turnover for all businesses. A challenge is that these firms are nested within geographies (NUTS-3 or local authority) and industries (4-digit SIC) which differ from the nineteenth century data. A corollary is that we will lose some geographic granularity in analyzing firm productivity and that we will not

be able to map industrial codes in earlier censuses with later industry classifications—which does not prevent us from controlling for 4-digit industry-fixed effects.

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

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 fragmentation from micro-census records (in 1851 and 1861) where land acreage is reported by landowners. Note that inferring land ownership fragmentation from micro-census records requires textual analysis as the information has not been coded by the I-CeM project.[11]

We complement the previous data on population, occupation and geography with the transportation infrastructure (main roads, navigable waterways, train lines and train stations), as provided by the Cambridge Group for the History of Population and Social Structure (Shaw-Taylor et al., 2024). This dynamic characterization of transportation allows us to measure access to resources through the transportation network and trading costs across different cities over time (see Appendix A.1).

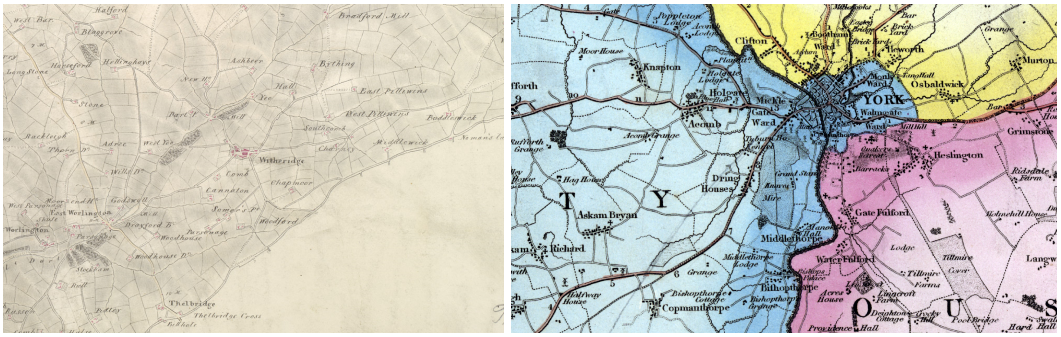
Urban settlements at the onset of industrialization In the early nineteenth century, the population of Britain mostly lives in rural settlements or small towns. That starting point allows for a significant reshaping of the urban network, with the rise of new cities sprawling around existing hamlets. This also represents a challenge: how can we identify potential cities and their boundaries?

Our approach exploits early county maps and Ordnance Survey drawings, mostly produced between 1790–1830, that cover the whole territory of England and Wales. We digitize, geo-reference and process these maps through computer vision.[12] We provide

[11] The occupational category in micro-censuses provided by I-CeM is extracted and inferred from a free text entry detailing occupation; one example for agricultural occupations would be "Farmer, 5 acres, 4 men, 2 boys". We use such free text entry to systematically collect information about acres and employment.

[12] We rely on about 350 Ordnance Survey drawings (covering Southern England and Wales) and 10 larger county maps (covering Chester, Cumberland, Derbyshire, Durham, Lancashire, Lincolnshire Northumberland, Nottinghamshire, Westmoreland, and Yorkshire). These collections markedly differ in

Figure 5. A rich set of Ordnance Survey drawings and county maps (1790–1830).



(a) South Molton (Devon) (b) York (Yorkshire)

Note: Panel (a) displays the surroundings of South Molton in Devon, a typical example of Ordnance Survey drawings (2 inches to the mile, 1804). Panel (b) displays the surroundings of York in Yorkshire, a typical example of county maps (4/5-inches to the mile, 1828). While built-up tends to be indicated with red in Ordnance Survey drawings, it is typically represented as dark rectangles and gray blocks in county maps. See Appendix A.1 for a description of these maps.

illustrations in Figure 5 centered around South Molton (Devon, Ordnance Survey drawing) and York (Yorkshire, county map). Identifying objects of interest on these maps is challenging (see, e.g., Combes et al., 2022, for a discussion of map digitization through machine learning). In what follows, we briefly describe our approach to identifying built-up areas, and we leave details to Appendix A.1.

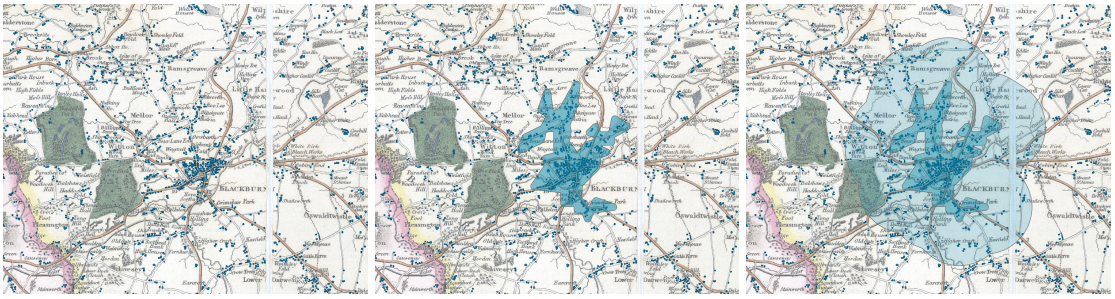
To identify built-up areas, we use a U-Net—a neural network commonly used to detect objects of interest within images. We train the model by manually labeling built-up areas across the different map tiles and drawings, and validate it through the use of a validation sample. Due to the very large sample of labeled buildings and their heterogeneity, overfitting is a second-order concern; we however apply various transformations (i.e., flipping, zooms, rotations) to each training batch, which allows us to impose invariance properties to these transformations. Our procedure leads to a good classification of built-up areas (our main object of interest), with a precision of about 0.96.

At that time, however, isolated built-up areas and small settlements are a frequent occurrence. The settlements that would later consolidate into towns and grow into cities are large, dense or spread across contiguous hamlets. To identify future cities, we develop a procedure to select urban settlements as sufficiently large clusters of built-up areas.

Figure 6 illustrates our approach with the city of Blackburn. We first detect built-up areas on the raw map with the previous algorithm (panel a). We then follow the procedure developed in De Bellefon et al. (2021) to identify a nucleus of high density and contiguous areas of excess building density. We illustrate the outcome of this procedure

their display of built-up areas; they also use different symbols and colors within collections.

Figure 6. Clustering procedure.



(a) Detecting built-up areas (b) Clusters and city boundaries (c) Predicted sprawl

Notes: The underlying map is the county map of Lancashire drawn by G. Hennet. The blue rectangles are built-up areas detected by our algorithm; the blue area in panel (b) is the outcome of our clustering algorithm adapting De Bellefon et al. (2021). In panel (c), the lighter blue area is predicted urban sprawl based on average urban sprawl across all towns. While these areas can overlap, at least in principle, we record no such overlaps for our baseline sample of 435 cities.

in panel (b) of Figure 6.[13] We then construct predicted boundaries for these cities at the end of the nineteenth century by assuming that towns and cities all grow at the same proportional rate across the country and do so homogeneously in any direction (panel c). In practice, towns and cities all expand to some degree, but this expansion is not homogeneous across all directions and not homogenous across cities. We discuss in Section 2.2 how we predict the extent of such expansion, and therefore cities' late-nineteenth century population, with land ownership fragmentation.

2.2 Predicting population and industrial concentration

Late-nineteenth century population and industrial concentration are primarily explained by geography, and the pre-expansion distribution of location advantages.

The heterogeneous rise of Great British cities Our procedure to detect urban settlements around 1790–1820 identifies more than 800 clusters and their initial boundaries. Throughout the paper, we consider a sample of settlements excluding those with an agricultural share of employment in the top quartile (above 60%) and a population below 3,000 inhabitants in 1817.[14] We are left with 435 potential cities, 7 of those being isolated

[13] An alternative procedure is described in Arribas-Bel et al. (2021).

[14] We provide robustness checks showing that this selection does not affect our headline estimates for the long-run impact of industrial concentration. Note that boundaries of urban settlements change during the period, and they do so endogenously. To strike a balance between comparing employment numbers within similar areas across time and dealing with endogenous urban sprawl, we associate employment within each 2-digit industry in 1817 and in 1851–1911 (and in 1971–2020) to a city c by intersecting the predicted boundaries around 1880–1900 with parish boundaries, allocating employment using the area share of the intersection. Predicted boundaries are calculated by assuming a homogenous, proportional growth of all urban areas in every direction between 1817–1881. In summary, we construct all our em-

in their ceremonial county. By the end of the nineteenth century in Britain, as urbanization slows, some of these cities are specialized in a few industries while others are more diverse. This process is disciplined by geography, trade and the distribution of location advantages. We now provide a few descriptive statistics illustrating the heterogeneous rise of Great British cities.[15]

Table 1. The role of initial industries and geography.

Adjusted R-squared	Herfindahl index (1881)	Employment (1881)
Initial industrial mix (1817)	0.782	0.453
+ Geography	0.838	0.724
Observations	435	435

Notes: The Herfindahl index (1881) is constructed from industrial employment shares (s_{ic}^{81} for city c and industry i) in 1881 as $\sum_j \left(s_{jc}^{81}\right)^2$. Employment (1881) is the (log) employment growth between 1817 and 1881, $\ln\left(L_c^{81}\right) - \ln\left(L_c^{17}\right)$. *Initial industrial mix (1817)* includes: dummies corresponding to deciles of employment in 1817, deciles of industrial concentration in 1817, and deciles of agricultural employment share in 1817; (log) population density in 1821; and the industry-based shift-shares, g_c and χ_c, later described in this Section 2.2. *Geography* includes: (log) area of the initial city outline; minimum, maximum, and average elevation; slope, bulk density, latitude; travel time to the closest city, to the closest market town, to each major port (London, Liverpool, Plymouth), and to coal; density of roads (1830), train lines (1851), and waterways; share of arable agriculture; suitability to grow wheat, oat, grass, and rye; share of heavy soil (NSRI); built-up density (within the contours of the settlement around 1800 and within 1, 2, 3, 5, 10 kms); and the predicted land ownership fragmentation within the contours of the settlement around 1800 and within a 10-kilometer buffer. Note that the Adjusted R-squared with only geographic variables are respectively 0.700 (Herfindahl index, 1881) and 0.654 (Employment, 1881).

To quantify the respective role of initial industrial mix and geography in predicting urban dynamics across urban settlements, we conduct a simple variance decomposition exercise explaining industrial concentration in 1881—constructed from employment shares in 1881, s_{ic}^{81}, in city c and 2-digit industry i, as $\sum_j \left(s_{jc}^{81}\right)^2$—and employment growth between 1817 and 1881. We first control for initial industries; we then add a large set of geographic indicators (topography, market access, transportation, agricultural potential, built-up density at the city fringe). Table 1 shows that initial industries are important in explaining the heterogeneous rise of Great British cities. However, they are much more important in explaining industrial concentration than growth: the set of "initial industries" variables explains 78% of the variation in industrial concentration across cities in 1881. By contrast, the same set of variables only explains 45% of the variation in growth across cities. The crucial factor to explain urban growth throughout the nineteenth century is geography: adding geographic factors to the previous regressions raises the ad-

ployment/population variables within predicted expanded cities as of the end of the nineteenth century.

[15] In Appendix A.3, we illustrate the role of initial industries in shaping the industrial structure of cities, and we discuss two empirical regularities of urban development: the Gibrat's law, and the Zipf law.

justed R-squared from 0.45 to 0.72. There are two types of geographic factors which predict urban development: (i) market access and connectivity; and (ii) topography and factors which influence land supply at the fringe of existing urban settlements, e.g., local agricultural conditions.[16]

In summary, the nineteenth century sees the rise of different cities. Our previous evidence hints at pre-expansion comparative advantages as a strong predictor for industrial concentration while urban growth is partly driven by geography.

An exogenous predictor for industrial concentration To predict industrial concentration in city c, $\sum_j \left(s_{jc}^{81}\right)^2$, we combine city-specific employment shares across 2-digit sectors $i \in \{1, \dots, I\}$ in 1817, s_{ic}^{17}, with aggregate sectoral employment growth between 1817–1881 across 2-digit sectors i, $g_i > 0$, into a predicted Herfindahl index of industrial concentration:

$$\chi_c = \frac{\sum_j \left(\sigma_{jc}^{81}\right)^2}{\left(\sum_j \sigma_{jc}^{81}\right)^2}$$

where $\sigma_{ic}^{81} = s_{ic}^{17} \cdot g_i$. This prediction, relying on the initial nature of cottage industries and the differential expansion of 2-digit sectors, constitutes the first building block of our empirical approach. We construct a more standard, linear shift-share predictor, $g_c = \sum_j \sigma_{jc}^{81} = \sum_j s_{jc}^{17} g_j$, based on aggregate shifts and the initial composition of employment within the predicted urban area, which will be used as a *control* in our empirical specification. Our exogenous prediction for city size will indeed rely on local geography.

Following the recent advances in identification and inference in shift-share designs, we also consider a "purer" shift-share instrument for industrial concentration. First, we consider $\chi_c = \sum_j \left(s_{jc}^{17} g_i\right)^2 / \sum_j \left(s_{jc}^{17}\right)^2$, thus verifying the property that city-specific shares $\left\{ \left(s_{ic}^{17}\right)^2 / \sum_j \left(s_{jc}^{17}\right)^2 \right\}_i$ sum up to 1 (Adão et al., 2019; Borusyak et al., 2022). Identification then requires the shifts—functions of 2-digit sectoral growth at the national level—to be sufficiently numerous and quasi-random. Second, we also consider a leave-out approach, excluding a buffer of 10 kilometers around the city, to construct shifts that are only related to aggregate demand, rather than local supply factors.

An exogenous predictor for city size based on land fragmentation We now describe how we construct a predictor ζ_c for the population of city c that is exogenous to later city dynamics. The idea is to identify land fragmentation as induced by local gradients in soil conditions, around the historical boundaries of an urban settlement (before the time of rapid industrialization).

[16] The role of agricultural hinterlands in fueling urban growth has been discussed in Matsuyama (1992), Coeurdacier et al. (2022), Heblich et al. (2024) and Nagy (2023). Our instrument will rely on another theoretical mechanism, the land assembly problem, and we will control for local agricultural productivity.

Land ownership fragmentation across different parts of England and Wales was not only instrumental to the development of agriculture, as illustrated by the effect of enclosures (Heldring et al., 2022), but it was also crucial in disciplining city growth 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 location spans multiple land parcels, the possible buyer needs to engage in multilateral bargaining in which the value of the marginal parcel increases as the buyer acquires the rights to use other parcels. The number of different parties then matters. This issue is a "standard" hold-up problem, which has been labeled as the land assembly problem in this specific context (see, e.g., Eckart, 1985; Strange, 1995). The consequences for city growth are straightforward: high land fragmentation at the fringe of the city makes negotiations to develop the land for urban use costly. As a result, cities with fragmented ownership in their immediate fringe have a less elastic land supply in the short run and may be worse at responding to sudden bursts in land demand. The rapid industrialization induced by technological progress and subsequent surge in external trade in the mid-nineteenth century, as evidenced in Section 1, is one such shock.

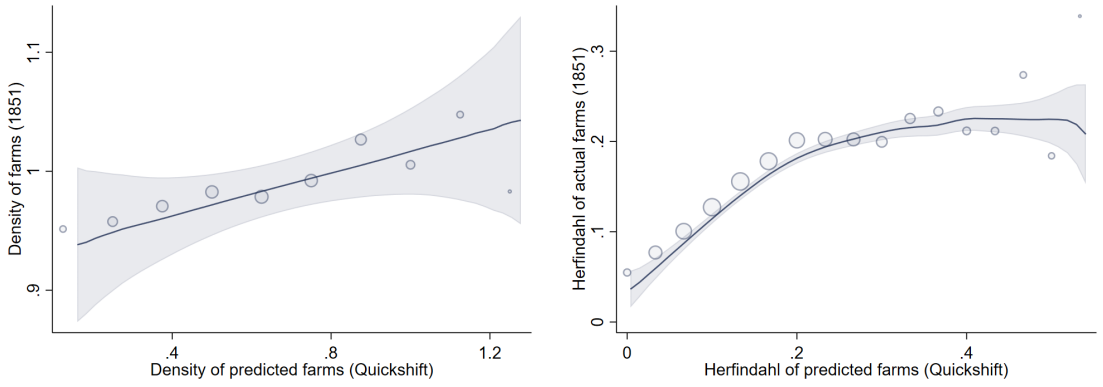
A major issue with land ownership fragmentation in the immediate fringe of cities is that it may be influenced by future city dynamics. The relationship between land ownership fragmentation and city growth may be "contaminated" by omitted variation—the relative productivity of land between urban and rural use inducing different ownership structure—but also reverse causality—land owners being more or less willing to own large plots around cities most likely to expand.[17] For these reasons, we would like to consider a measure of fragmentation with the following characteristics: (i) the measure, as evaluated within a neighborhood of city boundaries at the beginning of the nineteenth century, should predict city growth in the subsequent 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 and are 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 A.2 and only summarize its main steps below. First, we combine these different dimensions of characteristics into a multi-band raster covering England and Wales at a resolution of

[17] A recent contribution discusses the role of enclosure acts on crop yields and land inequality (Heldring et al., 2022). Our source of variation to explain land fragmentation will be, in essence, orthogonal to the one used in their analysis. However, we do control for the specific role of enclosures in a robustness check.

30m (elevation, slope, bulk, soil classes, zones delineated by waterways). Second, we use an unsupervised segmentation algorithm which isolates homogeneous (and contiguous) "color zones" or superpixels. In the present application, the raster contains more than three bands, but the principle is the same as color segmentation: the algorithm maximizes a weighted sum of target distances within constituted superpixels, with a weight allocated to physical distance relative to the "value distance." A superpixel with similar values is, here, 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. Third, we compute the density of predicted farms (or superpixels) in the fringe of cities by drawing buffers of fixed (e.g., 1, 2, 3, 5, 10 kms) and relative widths (e.g., proportional to the initial area and calibrated on the average city growth over the period, used for our baseline measure ζ_c) around city boundaries at the onset of rapid industrialization. One can think of the narrow, relative rings as predicting the propensity for cities to grow over the nineteenth century and of the wider rings as controlling for later land supply elasticities. The quantitative model developed in Section 4 will allow for cities to face varying land supply elasticities over time, in part to capture the previous intuition.

Figure 7. Validation of the predicted measure of land fragmentation.



(a) Density (b) Land concentration

Notes: The left panel displays the measure of predicted fragmentation computed at the parish level, ζ_p, versus actual farm density as collected from micro-census records in 1851 across 11,500 parishes of England and Wales. We create about 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). The right panel repeats the same exercise with the predicted and actual Herfindahl concentrations of farm ownership based on reported acres across parishes of England and Wales.

The predicted measure of land fragmentation, ζ_c, should be correlated with the actual fragmentation of land ownership. We validate the predicted measure of land fragmentation by comparing it with actual farm density and farm concentration as collected from

micro-census records in 1851 across all parishes of England and Wales (see Figure 7.[18]

3 Empirical facts

In this section, we derive motivating facts about the long-run dynamics of cities.

3.1 Empirical strategy

A general formulation for Jacobs externalities A traditional explanation for the interconnected dynamics of industries and cities is based on the nature of comparative advantages. Consider an industry i in city c at time t, and let $\mathcal{T}_{ict}$ denote its total factor productivity. Ignoring local externalities and assuming a constant location advantage $\mathcal{A}_{ic}$, (revenue-based) total factor productivity can be written as,

$$\mathcal{T}_{ict} = \mathcal{A}_{ic} \cdot \mathcal{P}_{it},$$

where $\mathcal{P}_{it}$ captures macroeconomic factors of demand (and technological change) affecting the dynamics of industries within the country. Under this formulation, the dynamics of cities, as disciplined by their portfolio of productivity across industries $\left\{\mathcal{A}_{jc}\right\}_{j\in I}$, reflects the dynamics of industries in which they have a comparative advantage. We will modify the previous framework by introducing static and *long-run* externalities,

$$\mathcal{T}_{ict} = \mathcal{A}_{ic} \cdot \mathcal{P}_{it} \cdot f_i\left(\left\{L_{jc,t}\right\}_{j\in I}\right) \cdot g_i\left(\left\{L_{jc,t-1}\right\}_{j\in I}\right),$$

where the function f_i capture static spillovers and g_i map the local industrial structure in $t-1$ to the distribution of productivity in period t. In practice, the data will not allow us to keep such level of generality, and we will consider externalities of the form,

$$f\left(\left\{L_{jc1}\right\}_{j\in I}\right) = L_{ic1}^{\mu}, \quad g\left(\left\{L_{jc0}\right\}_{j\in I}\right) = L_{c0}^{\rho} \cdot \left[\sum_{j\in I}\left(\frac{L_{jc0}}{L_{c0}}\right)^2\right]^{-\iota},$$

where: period 0 will be the year 1881 and period 1 will be contemporaneous outcomes; L_{ic1}^{μ} will capture static Marshall-Arrow-Romer externalities operating within industries; and L_{c0}^{ρ} will be long-run agglomeration spillovers while the term between brackets will measure the impact of long-run Jacobs externalities.

[18] This exercise shows that our measure of land fragmentation does predict land ownership fragmentation; however, it does not rule out that the measure is related to city dynamics before industrialization. We further validate the land fragmentation measure by comparing urban settlements with different degrees of land fragmentation in their "external crusts," as calculated at the onset of the nineteenth century (1817), in Appendix A.3. We do not find evidence that settlements with different degrees of land fragmentation were different in their population and industrial concentration before the era of rapid industrialization.

Identification The estimation of the previous relationship presents challenges, e.g., because of omitted variation correlated with long-run productivity and the historical portfolio of industries. Our main empirical strategy will be to consider the following empirical (log) equivalent,

$$\ln(\mathcal{T}_{ic1}) = \mu l_{ic1} + \rho l_c - \iota h_c + \eta_i + \gamma \mathbf{X}_c + \varepsilon_{ic1}, \tag{1}$$

where: measures of population (l_c) and industrial concentration (h_c) in period 0 are instrumented by our geological predictor of city size (ζ_c) and historical determinants of industries before the swift changes in aggregate demand (χ_c); $\eta_i = \ln(\mathcal{P}_{i1})$ captures industry trends; and our approach to properly isolating the fundamentals of city c, $\ln \mathcal{A}_{ic}$, will consist in conditioning the specification on a rich set of controls capturing first-nature geography, second geography, and the baseline industrial composition in 1817.

Identification of the long-run parameters (ρ, ι) relies on the assumptions that (ζ_c, χ_c) predict population and industrial concentration, and that they do not have a direct effect on later productivity. The latter requires conditioning specification (1) on: granular controls for initial industries in 1817 to ensure that the variation induced by the predictor for industrial concentration (χ_c) comes from the shifts, i.e., the differential dynamics of sectors during the time of rapid industrialization and market expansion; and land supply factors relevant to cities after 1881.

In practice, we will estimate Equation (1) in its "most structural form" in the quantitative exercise of Section 5, and consider different variations of the previous equation in this section. One reason is that a few illustrative outcomes will be at the city level, thus changing our approach to controlling for industry-specific factors.

Measurement and sensitivity to alternative choices The previous estimation of externalities relies on a few choices, e.g., related to the spatial extent of such spillovers, to treatment heterogeneity and to the (log) linear nature of the estimated relationship (Combes and Gobillon, 2015).[19] In the baseline specification(s), we will assume a (log) linear relationship, that the spillovers are homogenous, and that they exclusively operate *within* cities (motivated by the localized nature of Jacobs externalities, see, e.g., Carlino and Kerr, 2015)—assumptions that we will relax in a sensitivity analysis.

The literature also discusses measurement issues: (i) how to construct an industrial concentration or diversity index; and (ii) which outcomes to consider. Our baseline mea-

[19]The empirics of agglomeration externalities usually distinguishes static and dynamic externalities. Our approach differs from that of the literature in that we consider *long-run* externalities, thus relating industrial structure at the end of the nineteenth century to contemporary outcomes—only relying on earlier measures for identification purposes.

sure is a Herfindahl industrial concentration index constructed from the shares of local employment within industries; we will consider alternative measures and associated instruments in robustness checks, e.g., measures of local distance to the national industrial portfolio. Regarding the latter issue, our approach will rely on (i) city-level proxies for economic development including wages, (ii) an empirical measure of total factor productivity at the firm level, and (iii) a model-inferred measure of total factor productivity at the city/industry level (relying on the structure of the model and the distribution of employment and wages across cities and industries).

3.2 Motivating facts

The (exogenous) rise of different cities To establish a causal link between population, industrial concentration and the subsequent dynamics of cities, we need to isolate exogenous variation among the numerous factors shaping their transformation during the nineteenth century. Our empirical strategy relies on a specification with two instruments: land ownership fragmentation at the fringe of urban settlements (ζ_c), intuitively predicting (log) *population* when the structure of cities stabilizes; and our "shift-share" Herfindahl index (χ_c), predicting *industrial concentration*.

We report the main estimates of the first stage in Table 2. First, predicted industrial concentration does induce more industrial concentration in 1881. The pass-through between the predicted Herfindahl index (χ_c) and the actual one (h_c) is around 0.51, i.e., a 0.10 higher predicted Herfindahl index translates into a 0.051 higher Herfindahl index in 1881. Second, we find that predicted land fragmentation impacts the capacity of cities to grow during the era of industrialization, and hence their population in the late nineteenth century. We find that one standard deviation in land fragmentation (about 0.5) decreases employment by about 6% (0.5×0.12). Importantly, the instrument for urban growth does not predict industrial concentration in any significant way, conditional on our set of controls: we need some (conditional) orthogonality between the two sources of variation, given that our objective is to untangle as much as possible an effect passing through industrial concentration from an effect of city size. In that spirit, we will report the conditional F-statistics following Sanderson and Windmeijer (2016), and conditioning on the other endogenous variable(s).

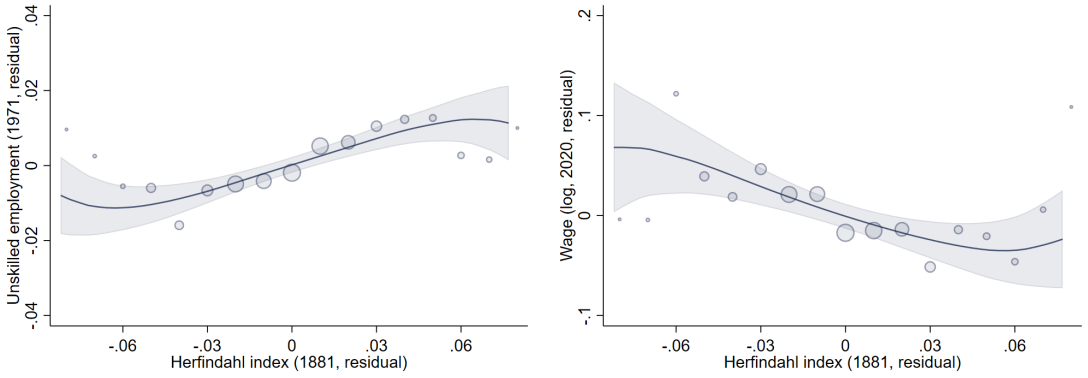
Industrial concentration causes long-run decline The concentration of industries appears to be detrimental in the longer run. Figure 8 shows how industrial concentration in 1881 correlates with a measure of urban deprivation in the longer run, i.e., the share of low-skilled workers in 1971 (derived following Heblich et al., 2021, see panel a), before the swift decline in manufacturing employment. Panel (b) shows how industrial

Table 2. Predicting late-nineteenth century industrial concentration and population.

	Herfindahl index (1881) (1)	Employment (1881) (2)
Predicted Herfindahl index (χ_c)	0.512	-1.910
	(0.059)	(0.587)
	[0.390]	[-0.221]
Land fragmentation (ζ_c)	-0.004	-0.122
	(0.004)	(0.035)
	[-0.019]	[-0.090]
Observations	428	428

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). Standardized effects are reported in square brackets. The dependent variables are the Herfindahl index of industrial concentration in 1881 (h_c, column 1) and (log) employment in 1881 (l_c, column 2). The set of baseline controls include: dummies corresponding to deciles of employment in 1817, deciles of industrial concentration in 1817, and deciles of agricultural employment share in 1817; (log) population density in 1821; the industry-based shift-share, g_c, described in Section 2.2; (log) area of the initial city outline; minimum, maximum and average elevation; slope, bulk density, and latitude; travel time to the closest city, to the closest market town, to each major port (London, Liverpool, Plymouth), to coal; density of roads (1830), train lines (1851), and waterways; share of arable agriculture (Tithe survey); suitability to grow wheat, oat, grass, and rye; share of heavy soil (NSRI); built-up density (within the contours of the settlement around 1800 and within 1, 2, 3, 5, 10 kms); and the predicted land ownership fragmentation within the contours of the settlement around 1800 and within a 10-kilometer buffer. The specification also adds fixed-effects at the level of 39 ceremonial counties and a shift-share control based on employment shares in 1881 and aggregate employment growth between 1881–1971. The two instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2 and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c.

Figure 8. Industrial concentration appears to be detrimental in the longer run.



(a) Unskilled employment (1971) (b) Wage (2020)

Notes: Panel (a) displays the relationship between industrial concentration in 1881 and the share of unskilled workers in 1971 (following the definition used in Heblich et al., 2021) across our 435 cities; both measures are cleaned for all controls used in Table 2. The dots represent the average residualized share of unskilled workers in 1971 within each of 20 bins grouping residualized Herfindahl index across cities, and the line is a locally weighted regression on all observations. Panel (b) displays the relationship between residualized industrial concentration in 1881 and residualized (log) wage in 2020.

concentration in 1881 correlates with a more recent measure of wage—labor income estimates for small areas based on the Family Resources Survey in 2020. The estimated gradients imply that less specialized cities in 1881, with a Herfindahl index around 0.05, would have about 2 percentage points fewer unskilled workers in 1971 and 5% higher wages, compared to specialized cities with an index of 0.10.

Table 3. The long-run effect of industrial concentration and city population.

	Unskilled employment (1971) (1)	Wage (2020) (2)
Herfindahl index (1881, h_c)	0.369	-2.656
	(0.148)	(0.888)
Employment (1881, l_c)	0.038	-0.307
	(0.030)	(0.177)
Observations	428	428
F-stat (h_c)	15.30	15.30
F-stat (l_c)	14.94	14.94

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). In column (1), the dependent variable is the share of unskilled workers in 1971 (following the definition used in Heblich et al., 2021); in column (2), the dependent variable is a measure of (log) wages in 2020. The main endogenous variables are the Herfindahl index of industrial concentration in 1881 (h_c) and (log) employment in 1881 (l_c). The set of baseline controls is the same as in Table 2. Note that we control for a shift-share control based on employment shares in 1881 and aggregate employment growth between 1881–1971, to clean for sector-specific trends. The two instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2 and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

We now use the exogenous variation in economic structure at the end of the nineteenth century to provide causal evidence on how patterns of urban development influence long-run city dynamics. We first shed light on the long-run dynamics of cities by regressing contemporary outcomes in city c, y_c^1, on urban development in 1881, $\mathbf{x}_c^0 = (h_c, l_c)$,

$$y_c^1 = \alpha + \beta \mathbf{x}_c^0 + \gamma \mathbf{X}_c + \varepsilon_c,$$

while instrumenting $\mathbf{x}_c^0$ with our exogenous predictors $\mathbf{p}_c = (\chi_c, \zeta_c)$ and conditioning on a large set of covariates $\mathbf{X}_c$. This specification differs from specification (1) and its later structural estimation in two main dimensions. First, and for the sake of interpretation, we consider a linear specification in industrial concentration, with $h_c = \sum_{j \in I} \left(\frac{L_{jc0}}{L_{c0}}\right)^2$. Second, the equation is specified at the city level, implying that we cannot properly control for static Marshall-Arrow-Romer externalities and for industry fixed effects—a caveat that we will address in firm- or industry-level specifications. The effect of industrial concentration on long-run city performance might indeed reflect two distinct

forces: nationwide industry decline might hurt cities specialized in the declining industries; and industrial concentration might have a direct effect on future city productivity, à la Jacobs (1969). We aim at cleaning our estimates from the former effect by controlling for aggregate industry trends from 1881 to the recent period in a shift-share design.

Table 3 shows that industrial concentration is detrimental in the long run: cities with a 0.10 higher Herfindahl index in 1881 have a 0.037 higher share of unskilled workers in 1971—about 80% of a standard deviation (column 1), and wages are 26.5% lower in 2020 (column 2). City size has some predictive power, but its effect is one order of magnitude smaller and not statistically significant: a 20% larger city in 1881 has a 0.008 higher share of unskilled workers in 1971 and 6% lower wages in 2020.

Table 4. The long-run effect of industrial concentration and city population—firm-level data.

Turnover (ABS)	(1)	(2)
Herfindahl index (1881, h_c)	-3.378	-2.376
	(1.716)	(1.176)
Employment (1881, l_c)	-0.311	-0.213
	(0.043)	(0.029)
Observations	2,192,591	2,192,591
Local industrial share	No	Yes

Notes: A unit of observation is a firm/year in the Annual Business Survey (ABS). Standard errors are reported between parentheses and are clustered at the NUTS-3 level. The dependent variable is (log) annual sales/turnover. The main endogenous variables are the Herfindahl index of industrial concentration in 1881 (h_c) and (log) employment in 1881 (l_c). The two instruments are the "shift-share" predictor of industrial concentration (χ_c) and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c. The baseline controls are the same as in Table 3. In columns (1) and (2), we control for year fixed effects, and (log) production factors (material, employment, capital). In column (2), we control for SIC-4 fixed effects and the local share of turnover/employment in the firm industry.

In Table 4, we exploit firm-level data between 2008–2018, nest all our geographic variables at the NUTS-3 level (local administrative units), and replicate our previous empirical strategy in a stacked regression setting with (log) sales as the dependent variable while controlling for year fixed effects, production factors, 4-digit industry fixed effects (η_i, see column 2), and the local share of turnover/employment in the firm industry (to account for Marshall-Arrow-Romer externalities, see column 2). The estimate of interest can thus be interpreted as an effect on "real" total factor productivity, orthogonal to the firm industry. We find that cities with a 0.10 higher Herfindahl index in 1881 are 24% less productive, when a 20% larger city in 1881 is 4% less productive.

Our empirical identification of long-run externalities relies on strong assumptions about their "extent" across space and across industries. First, in our baseline strategy,

we ignore the importance of linkage intensity across different sectors. For instance, a few sectors might be intrinsically much more connected to other sectors, which could alleviate the long-run impact of industrial concentration. Second, we limit the spatial extent of externalities to cities and ignore the industrial structure of nearby cities in the estimation of such externalities. The main reasons for doing so are the limits imposed by our causal identification strategy: causally estimating spillovers across industries or across cities is challenging. We however shed some light on these issues next.

Robustness checks Industrial concentration is a key driver of the long run underperformance of British cities. We now provide a series of robustness checks to support this empirical finding. This section summarizes a lengthier discussion in Appendix A.4.

First, we provide a series of robustness checks to support our identification hypothesis. Our identification narrative is that land fragmentation affects urban development through land supply at the fringe of early urban settlements; this narrative rules out the existence of competing mechanisms arising, e.g., from agricultural productivity (Coeurdacier et al., 2022), agricultural mechanization at the fringe of cities (Caprettini and Voth, 2020) or from the emergence of different informal institutions in places with different structure of land ownership (Heldring et al., 2022). To alleviate concerns about these competing channels, we control in our baseline specification for: the marginal land supply from the end of the nineteenth century onward; the share of arable in local production (which defines the exposure to the Grain Invasion; see Heblich et al., 2024); the potential yield of the most common crops; and access to resources (to reduce concerns that energy is the main factor explaining the rise and then fall of Great British cities).[20] In Appendix A.4, we further control for agricultural productivity, land market prices before 1817, and parliamentary enclosures.

Our argument is that the industrial concentration effect operates through intertemporal externalities à la Jacobs. In Appendix A.4, we show that our findings are orthogonal to the secular process of structural transformation and the reallocation of labor across wider sectors (e.g., between agriculture and manufacturing, or towards services), we test that the results are not driven by any of the nineteenth-century 2-digit industries, and we adopt a specification accounting for linkages across sectors (measured using occupational transitions for employed males in the micro-censuses of 1851 and 1861). We also provide support for our shift-share design by: considering an exogenous predictor for industrial concentration based on a "pure" shift-share formula; constructing an instrument by omitting the city and its surrounding from the derivation of the "shifts"; and

[20] In unreported checks, we further show that the vast majority of cities grow out of this narrow ring by the end of the nineteenth century, and that there are no differential zoning policies at the fringe of cities—depending on their initial land fragmentation (e.g., green belts, or social housing policies).

adopting other specifications considered in the literature (e.g., using the logarithms of the Herfindahl index of industrial concentration or its inverse as discussed in Combes and Gobillon, 2015, or controlling for the number of active industries).

Second, we quantify the spatial extent of Jacobs externalities in a specification with three endogenous variables—population, local industrial concentration and nearby industrial concentration—where nearby industrial concentration and its instrument incorporate nearby settlements within 10, 20, and 40 kilometers. We find support for our baseline strategy, which ignores the industrial structure of nearby cities in estimating dynamic externalities: spatial spillovers appear limited beyond 10 kilometers.

Third, the previous evidence ignores treatment heterogeneity. In Appendix A.4, we document: the interacted effect of industrial concentration and size; treatment heterogeneity in the incidence of services, manufacturing, transportation—cities that are highly service-intensive are slightly less prone to a specialization curse; and treatment heterogeneity in the "potential" of the historical industry portfolio (evaluated through the more recent dynamics of industries). Treatment heterogeneity is limited.

Fourth, we conduct a sensitivity analysis around the baseline specification(s): with a different specification for the land fragmentation instrument (with a fixed buffer, dropping bulk density from the algorithm delineating parcels, using SLIC rather than Quickshift to segment space into potential agricultural parcels); with different inference to account for spatial correlation; with different cut-offs to define urban settlements; and with outcomes computed at other dates.

These robustness checks leave our headline empirical finding—the detrimental long-run effect of industrial concentration—unchanged. The stylized facts presented in this section provide evidence about the joint dynamics of urbanization and industrial diversity in cities, showing that the fate of cities is tightly related to that of their industrial structure. The next section provides a more structural approach, by developing a quantitative model of cities and their industries over time which captures both the spatial linkages across cities and their industries in a given period, as well as possible intertemporal externalities.

4 A multi-sector dynamic spatial model

In this section, we develop a spatial, multi-sector, dynamic model of cities.

4.1 Setup

The model involves a finite number of cities, $c \in \{1, \dots, C\}$, and industries, $i \in \{1, \dots, I\}$. Time is discrete and is indexed by t. Within each industry, every city produces its own

variety that consumers view as different from varieties produced in other cities. There is an exogenous number of workers in the economy, $\bar{L}$. Each worker lives for one period, is endowed with one unit of labor, and maximizes her utility from the consumption of varieties.[21] The worker decides which industry to work for and which city to live in.

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.

Preferences If a worker m who lives at time t decides to work in industry i and to reside in city c, she chooses her consumption levels to maximize her utility,

$$U_{ict}^{m} = \max_{\{q_{jdt}^{m}\}_{j}} \left\{ a_{ct}^{m} \left[\sum_{j=1}^{I} \left(\sum_{d=1}^{C} \left(q_{jdt}^{m} \right)^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \right\}, \tag{2}$$

subject to the budget constraint,

$$\sum_{j=1}^{I} \sum_{d=1}^{C} p_{jdt} q_{jdt}^{m} \leq w_{ict} + R_{ct}, \tag{3}$$

where a_{ct}^{m} denotes the level of amenities enjoyed by the worker in city c, while the rest of the worker's utility is drawn from her consumption of varieties. More precisely, q_{jdt}^{m} denotes the worker's consumption of the city-d variety in industry j, p_{jdt} denotes 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. When choosing her city, industry and consumption, the worker takes amenities, prices and wages as given. We assume that varieties are substitutes, and they are more substitutable within than across industries, implying $1 < \sigma < \epsilon$.

Amenities are a combination of three factors,

$$a_{ct}^{m} = \bar{a}_{c} L_{ct}^{-\lambda} \varepsilon_{ct}^{m}, \tag{4}$$

where: $\bar{a}_{c}$ is the fundamental amenity level of city c, stemming from natural characteristics such as climate; $L_{ct}^{-\lambda}$ is a congestion externality that makes cities with a larger population L_{ct} less pleasant places to live; and ε_{ct}^{m} is an idiosyncratic taste shock for city c that is drawn from the following Fréchet distribution,

$$Pr\left[\varepsilon_{ct}^{m} \leq z\right] = e^{-z^{-1/\eta}}, \tag{5}$$

[21] In light of each period corresponding to a century in the quantification, we view this as a realistic assumption. We present a version of the model with infinitely lived, forward-looking workers in Appendix B.1.

and independently so across workers, cities and time periods. If η is high, then workers' utility is influenced to a large extent by their idiosyncratic tastes for cities; they are likely to settle in a city that they like, rather than in a city that offers them economic opportunities. The same would be true if workers faced high costs of moving across cities: η can also be interpreted as a parameter driving the severity of mobility frictions across cities.

Technology Varieties are produced by perfectly competitive firms. The representative firm produces the city-c variety in industry i at time t by combining workers (L_{ict}) and land (H_{ict}) as follows,

$$Y_{ict} = \Gamma \mathcal{T}_{ict} L_{ict}^{\gamma} H_{ict}^{1-\gamma}, \tag{6}$$

where $\mathcal{T}_{ict}$ is the Total Factor Productivity (TFP) of industry i in city c at time t. The constant $\Gamma = \gamma^{-\gamma}(1-\gamma)^{-(1-\gamma)}$ is introduced to simplify 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}. We assume that trade costs are non-negative, which amounts to $\tau_{icdt} \geq 1$.

Land is supplied in each city according to the supply function,

$$H_{ct} = r_{ct}^{\zeta_{ct}-1}, \tag{7}$$

where r_{ct} is the 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 in order 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 TFP, we set up the equilibrium within a given time period t for a *given* distribution of TFP in that period. In this within-period equilibrium, we impose that the labor market clears in each city, the land market clears in each city,

$$\sum_{j=1}^{I} \frac{1-\gamma}{\gamma} w_{jct} L_{jct} = r_{ct} H_{ct}, \tag{8}$$

markets clear for each variety,

$$(w_{ict} + R_{ct}) L_{ict} = \sum_{d=1}^{C} \left(\frac{P_{idt}}{P_{dt}}\right)^{1-\sigma} \left(\frac{p_{ict}\tau_{icdt}}{P_{idt}}\right)^{1-\epsilon} \sum_{j=1}^{I} \left(w_{jdt} + R_{dt}\right) L_{jdt}, \tag{9}$$

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

$$P_{idt} = \left[\sum_{c=1}^{C} p_{ict}^{1-\epsilon} \tau_{icdt}^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}}, \tag{10}$$

P_{dt} is the price index across all goods in city d,

$$P_{dt} = \left[\sum_{j=1}^{I} P_{jdt}^{1-\sigma} \right]^{\frac{1}{1-\sigma}}, \tag{11}$$

and each worker chooses the city and industry that offers them the highest utility.

Dynamic evolution of productivity We now describe how the distribution of productivity evolves over time. We allow the productivity of each industry to be influenced by agglomeration externalities of the form,

$$\mathcal{T}_{ict} = T_{ict} L_{ct}^{\alpha} g_i \left(L_{c,t-1}, \left\{ L_{jc,t-1} \right\}_{j \in I} \right), \tag{12}$$

where T_{ict} is the exogenous fundamental productivity of industry i in city c at time t. Agglomeration externalities depend on the current population of city c, as standard in the literature, past population (as in Allen and Donaldson, 2020), but also (past) sectoral composition. 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 (12) is a flexible formulation of externalities that generalizes the process described in the one-sector model of Allen and Donaldson (2020).

4.2 Solving the model

In this section, we propose an algorithm to solve for the equilibrium of the model. This algorithm relies on reducing the within-period equilibrium conditions to a system of $3 \times I \times C$ equations (as shown in Appendix B.2),

$$\begin{aligned}
x_{ict}^1 &= \sum_{j=1}^{I} \sum_{d=1}^{C} \left(x_{jdt}^2 \right)^{\frac{\alpha+\gamma}{\kappa_{dt}} \frac{\epsilon-1}{\sigma-1}} \left(x_{jdt}^3 \right)^{-\left(1-\frac{1+\lambda+\eta}{\kappa_{dt}}\right)} K_{icjdt}^1 \\
x_{ict}^2 &= \sum_{j=1}^{I} \sum_{d=1}^{C} \left(x_{jdt}^1 \right)^{\frac{\sigma-1}{\epsilon-1}} K_{icjdt}^2 \\
x_{ict}^3 &= \sum_{j=1}^{I} \sum_{d=1}^{C} \left(x_{jdt}^1 \right)^{-\frac{\epsilon-\sigma}{\epsilon-1}} \left(x_{jdt}^2 \right)^{-\left(1-\frac{\alpha+\gamma}{\kappa_{dt}} \frac{\epsilon-1}{\sigma-1}\right)} \left(x_{jdt}^3 \right)^{\frac{1+\lambda+\eta}{\kappa_{dt}}} K_{icjdt}^3
\end{aligned} \tag{13}$$

where

$$\kappa_{dt} = 1 + \lambda + \eta + \left[\frac{1-\gamma}{\zeta_{dt}} - \alpha + \left(\gamma + \frac{1-\gamma}{\zeta_{dt}}\right)(\lambda+\eta)\right](\epsilon - 1) \tag{14}$$

is a combination of structural parameters, and the $3\times I\times C$ unknowns x^1_{ict}, x^2_{ict} and x^3_{ict} can be obtained from equilibrium prices, wages and population levels through the following change in variables:

$$\begin{aligned} x^1_{ict} &= P^{1-\epsilon}_{ict} \\ x^2_{ict} &= w^{1-\sigma}_{ict} L^{(\lambda+\eta)(\sigma-1)}_{ct} \\ x^3_{ict} &= w^{1+\left(\gamma+\frac{1-\gamma}{\zeta_{ct}}\right)(\epsilon-1)}_{ict} L^{1+\left(\frac{1-\gamma}{\zeta_{ct}}-\alpha\right)(\epsilon-1)}_{ct} \end{aligned} \tag{15}$$

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

$$K^1_{icjdt} = \begin{cases} \left(\frac{1-\gamma}{\gamma}\right)^{-\frac{1-\gamma}{\zeta_{dt}}(\epsilon-1)} \overline{T}^{\epsilon-1}_{jdt} \tau^{1-\epsilon}_{jdct} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$K^2_{icjdt} = \begin{cases} \left(\gamma \overline{U}_t\right)^{1-\sigma} \bar{a}^{\sigma-1}_c & \text{if } c = d \\ 0 & \text{otherwise} \end{cases}$$

$$K^3_{icjdt} = \left(\gamma \overline{U}_t\right)^{1-\sigma} \left(\frac{1-\gamma}{\gamma}\right)^{-\frac{1-\gamma}{\zeta_{dt}}(\epsilon-1)} \overline{T}^{\epsilon-1}_{jct} \bar{a}^{\sigma-1}_d \tau^{1-\epsilon}_{jcdt}$$

where,

$$\overline{T}_{ict} = T_{ict} g_i \left(L_{c,t-1}, \left\{L_{jc,t-1}\right\}_{j\in I}\right) \tag{16}$$

is the productivity component which is exogenous in period t.

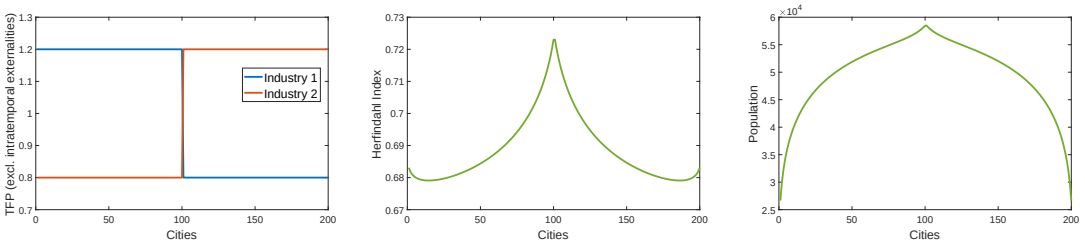
The solution algorithm consists of guessing an initial distribution of x^1_{ic0}, x^2_{ic0} and x^3_{ic0} in period 0. This is followed by inserting the guesses on the right-hand side of system (13), obtaining an updated guess of x^1_{ic0}, x^2_{ic0} and x^3_{ic0}, and iterating on system (13) until convergence. With the equilibrium values of x^1_{ic0}, x^2_{ic0} and x^3_{ic0} in hand, one can express period-0 price indices, wages, population and sectoral employment levels by inverting the system (15). Finally, one can obtain period-1 TFP from Equation (16) and repeat the procedure for subsequent periods, if any.

4.3 An illustration in a linear economy

In this section, we simulate the model on a simple geography to illustrate how it can rationalize: (i) the rise of (different) cities as influenced by trade and comparative advantages; and (ii) the long-run dynamics of cities, as documented in Section 3.

A linear economy We consider a country with 200 cities uniformly arranged on a line and two industries: in period 0, the 100 cities to the West of the country centroid have a TFP of 1.2 in industry 1, and a TFP of 0.8 in industry 2; the pattern is reversed for the 100 Eastern cities (panel a of Figure 9).

Figure 9. The linear economy under trade in period 0.



(a) Productivity (b) Industrial concentration (c) Employment

Notes: The values for the structural parameters are set as follows: $\alpha = 0.06$; $\gamma = 0.65$; $\epsilon = 5$; $\phi = 0.25$; $\sigma = 4$; $\bar{L} = 10,000,000$; $\zeta_{ct} = 2$. Panel (a) displays productivity as a function of longitude and across the two industries. Panels (b) and (c) display: a Herfindahl index of industrial concentration; and total employment, respectively.

Using the algorithm proposed in Section 4.2, 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 trade costs between cities c and d takes the following form,

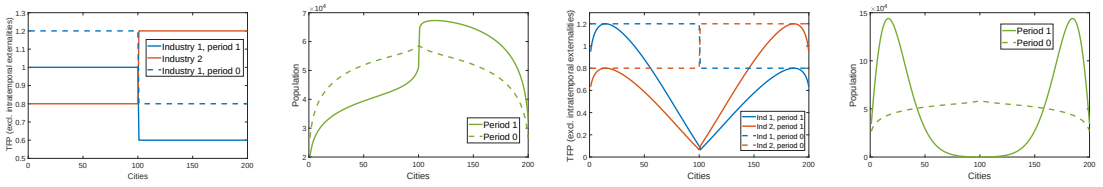
$$\tau_{icdt} = (1 + d_{cd})^{\phi},$$

in both industries and time periods, where d_{cd} denotes the Euclidean distance between cities c and d. We set the values of 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 total population to 10 million, which roughly equals the working population of England and Wales at the beginning of the 19th century. Finally, we set $\zeta_{ct} = 2$ for every city and time period for simplicity, i.e., a land supply elasticity of 1 in each city.

The rise of (different) cities We first look at the patterns of industrial concentration and the distribution of population in period 0. Panel (b) of Figure 9 shows how cities specialize under trade, as measured by their Herfindahl index across industries. Cities that are near the center—with the best access to trade with other cities and the largest room for specializing according to their comparative advantage—specialize more. Population also reallocates towards these central cities as they benefit from trade through their increased industrial concentration (panel c). This stands in stark contrast with the autarky

scenario in which industrial diversity and population would remain evenly distributed across cities, due to their symmetric fundamentals.

Figure 10. The linear economy in period 1 (A: differential industry trends, B: dynamic Jacobs externalities).



(a) Productivity (A) (b) Employment (A) (c) Productivity (B) (d) Employment (B)

Notes: In both scenarios, the values of structural parameters are set to: $\alpha = 0.06$; $\gamma = 0.65$; $\epsilon = 5$; $\phi = 0.25$; $\sigma = 4$; $\bar{L} = 10,000,000$; $\zeta_{ct} = 2$. Panel (a) and (c) display productivity as a function of longitude and across the two industries. Panels (b) and (d) display total employment as a function of longitude. The left panels are associated with scenario (A), featuring differential industry trends. The right panels are associated with scenario (B), featuring long-run Jacobs externalities. We provide complements to the analysis in Appendix B.4.

The long-run dynamics of cities We study the distribution of economic activity in the long run (period 1) under two scenarios. The first scenario assumes that industries differ in the evolution of aggregate productivity for exogenous reasons: the TFP of industry 1 uniformly decreases by 0.2 across cities, while the TFP of industry 2 remains the same as in period 0. 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 long-run externalities by setting $g_i(\cdot) = 1$ in Equation (12). The left panels of Figure 10 show that population reallocates from the affected Western cities towards the East as a result of differential industry trends. Central, formerly specialized cities in their respective regions are however larger than others.

The second scenario conversely abstracts from differential industry trends and allows for long-run agglomeration externalities of the form,

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

This formulation implies that less specialized cities, with a lower Herfindahl index, experience higher TFP in period 1 as shown in the bottom panels of Figure 10. In this example, Jacobs externalities lead to a stark reversal of fortune: central cities, which specialize heavily in period 0 and gain from such specialization, would lose in period 1 due to the dynamic consequences of their extreme industrial concentration.

How to tell apart differential industry trends from Jacobs externalities? Both of them

have the ability to rationalize the death of (some) specialized cities. However, differential industry trends imply that cities *with a comparative advantage in declining sectors* suffer in the long-run; by contrast, long-run Jacobs externalities induce an industrial concentration penalty, irrespective of the nature of comparative advantage.

5 Quantitative analysis

We now use the previous spatial model to quantify the role of local long-run externalities and their distributional consequences on the allocation of economic activity across space.

5.1 Parametrization, model inversion, and estimation

The quantification of our model requires: (i) a parametrization of key model elasticities; (ii) using its structure to recover unobserved city-specific fundamentals that rationalize the observed employment by city-industry and wages by city (a model inversion); and (iii) introducing the recovered productivities within a similar empirical framework as the one discussed in Section 3.1 to estimate Equation (12), i.e., the equation that drives the dynamic evolution of productivity.

Given its recursive structure, the model can be taken to the data in any period, and independently so. For reasons that will become clear later in this section, we primarily use contemporary data—covering the "long-run" or period 1—to estimate the model.

Parametrization and observable data We parametrize key elasticities of the model from the empirical literature—focusing primarily on recent times (i.e., period 1). We set the elasticity of substitution across sectors, σ, to be equal to 4 (Bernard et al., 2003) and an elasticity of substitution across varieties, ϵ, equal to 5 (Simonovska and Waugh, 2014). Our modeling of productivity features a static agglomeration externality, as usual in quantitative models fitting a distribution of workers across cities of very different sizes (e.g., discussed as scale economies in Henderson, 1974). We consider a static agglomeration externality, α, of 0.07 (following Graham and Gibbons, 2019, which reports the latest available agglomeration-productivity elasticity estimates for the United Kingdom). The dispersion parameter of idiosyncratic location tastes, η, and the congestion parameter, λ, are tied together in the equilibrium conditions (see, for instance, Equations 13 or 14); they both serve as mitigation forces to limit differences in city sizes. We thus set $\lambda + \eta$ to match an average migration elasticity of 0.5 (Head and Mayer, 2021). We use granular elasticities of housing supply at the Middle Layer Super Output Area (MSOA) level from Drayton et al. (2024) and map these estimates to the expanded urban area around our 435 cities in order to parametrize ζ_{c1} in each city c. These elasticities, obtained from a

similar procedure as in Baum-Snow and Han (2024), are quite low, around 1.2 on average, illustrating the (recent) institutional constraints to development in cities of England and Wales. We calibrate trade costs using data on the evolving transportation infrastructure within a procedure that is best described in Appendix B.5.

Finally, our quantitative exercise requires a labor share in production, data on employment per industry and per city (L_{ict}), and data on wages at the city level w_{ct}. We rely on Feinstein (1972) and set a labor share of production at 0.77, arguably in the high range of such estimates. As for employment by 2-digit industry (88 categories), we use the Business Register and Employment Survey in 2020 at the ward level and nest these estimates within our own geographies.[22] Finally, while we do observe productivity, production factors, and wages at the firm level (as exploited in Section 3.2), we do so within a secure server where we cannot properly invert our very demanding model. For this reason, we rely on labor income estimates for small areas (Middle Layer Super Output Area, provided by the Office for National Statistics in 2020 and based on the Family Resources Survey), which we nest within our geographies to measure wages at the city level. These quantities are key to recovering unobserved endogenous variables that will be the inputs of our counterfactual: amenities and city-industry productivities.

Model inversion In this step, we recover the distribution of city amenities, $\{\bar{a}_c\}_{c\in\{1,\dots,C\}}$, and city-industry productivities, $\left\{\overline{T}_{ic1}\right\}_{c,i}$, that rationalize observed data on wages (w_{c1}) and employment by city-industry in period 1 (L_{ic1}). The following theorem states that, given wages and employment per industry, there is a unique set of fundamentals that rationalize the data.

Theorem 1. *In each period and given the values of structural parameters, trade costs, and data on wages w_{c1} and employment by city-industry L_{ic1}, one can uniquely recover Total Factor Productivity levels $\overline{T}_{ic1}$ (up to scale) and amenities (relative to aggregate welfare, $\bar{a}_c/\overline{U}_1$).*

Proof. See Appendix B.3. □

Intuitively, the uniqueness result of Theorem 1 stems from the fact that the system of equations characterizing the equilibrium—Equations (B.9) to (B.11) in Appendix B.2—can be inverted to recover the values of fundamentals. The output of this procedure is a set of amenities $\mathcal{A} = \{\bar{a}_c\}_c$ and city-industry productivities $\mathcal{T} = \left\{\overline{T}_{ic1}\right\}_{i,c}$ across 435 cities (and 88 industries for the latter).[23]

[22] Employment by 2-digit industry and city in the nineteenth century are available from the quasi-census of Anglican baptism records collected around 1813–1820 and the micro-census records, e.g., in 1881, which we exploit in our empirical analysis.

[23] To implement the inversion procedure, we invert the system of equations (B.9) to (B.11) using an

Estimating long-run productivity spillovers Armed with the distribution of productivities, $\left\{\overline{T}_{ic1}\right\}_{i,c}$, we estimate the last block of our model, i.e., the equation characterizing the evolution of city-industry productivities:

$$\overline{T}_{ic1} = \mathcal{P}_{i1} \cdot L_{c0}^{\rho} \cdot \left[\sum_j \left(\frac{L_{jc0}}{L_{c0}}\right)^2\right]^{-\iota} \cdot L_{ic1}^{\mu} \cdot \varepsilon_{ic1}. \tag{17}$$

Equation (17) is a special case of Equation (12) characterizing the evolution of productivity in the model. In particular, Equation (17) restricts long-run externalities to be a function of two objects: past population (L_{c0}) and the past Herfindahl index, $\sum_j \left(\frac{L_{jc0}}{L_{c0}}\right)^2$. This restriction is motivated by the reduced-form evidence of Section 3.2, which points to the importance of these two variables—especially the latter—in explaining long-run city outcomes.

Equation (17) can be estimated within an empirical setting that mirrors the one described in Section 3.1. Specifically, one can use exogenous land fragmentation ζ_c and the (log) predicted Herfindahl χ_c as instruments for period-0 (log) population and (log) industrial concentration. Equation (17) corresponds to the second stage of this two-stage procedure. Note that we split the term T_{ic1} of Equation (12) into two components in Equation (17): an aggregate industry trend $\mathcal{P}_{i1}$ and a structural error term ε_{ic1}. In effect, the term $\mathcal{P}_{i1}$ can be estimated as an industry fixed effect in Equation (17). Isolating these industry trends allows us to disentangle them from the long-run externalities, in line with the logic discussed at the end of Section 4.3.

In practice, we estimate the following specification,

$$\ln \overline{T}_{ic1} = \eta_i + \rho \ln L_{c0} - \iota \ln \left[\sum_j \left(\frac{L_{jc0}}{L_{c0}}\right)^2\right] + \mu \ln L_{ic1} + \gamma \mathbf{X}_c + \varepsilon_{ic},$$

while instrumenting $\ln L_{c0}$ and $\ln \left[\sum_j \left(\frac{L_{jc0}}{L_{c0}}\right)^2\right]$ by our exogenous predictors $\mathbf{p}_c$, controlling for $\mathbf{X}_c$, used in Table 2, as well as county-level fixed effects and industry-level fixed effects, and clustering standard errors at the level of the closest city as of 2015. Table 5 shows that ρ is not statistically different from 0, consistent with the Gibrat's law. By contrast, ι is statistically different from 0 and quite large: a 20% increase in the Herfindahl index, about half of a standard deviation on average, reduces overall productivity in a city by 15-16% (see column 2 in which we control for Marshall-Arrow-Romer externalities). We provide some visual evidence for this negative relationship in Appendix B.5 and Appendix Figure B2. We also show that industrial concentration is *not* related to the recovered *amenities*—which we will assume to be fixed across periods and across our

algorithm that we detail in Appendix B.5.

Table 5. The structural estimates for long-run externalities.

Productivity $\left(\ln \overline{T}_{ic1}\right)$	(1)	(2)
Herfindahl index ($h_c = \ln \sum_j \left(\frac{L_{jc0}}{L_{c0}}\right)^2$)	-0.675	-0.788
	(0.273)	(0.318)
Employment ($l_c = \ln L_{c0}$)	0.300	-0.234
	(0.389)	(0.409)
Observations	38,280	21,330
F-stat (h_c)	23.26	19.94
F-stat (l_c)	17.80	17.09
Local industrial employment	No	Yes

Notes: A unit of observation is a city/industry where a city is identified following our clustering procedure described in Section 2.2 and an industry is one of the 88 industrial categories (2-digit) present in the Business Register and Employment Survey. In column (2), we control for the (log) local employment in the industry to account for Marshall-Arrow-Romer externalities when such employment is not equal to zero. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). The specifications include the same controls as in column (2) of Table 3, as well as 2-digit industry-level fixed effects.

counterfactual experiments.

5.2 Counterfactuals

We next use the estimated model to conduct counterfactual experiments. First, we assess the role of long-run Jacobs externalities in explaining the distribution of population and income across British cities. Second, we consider a stylized policy experiment in order to shed light on the arbitrage faced by a local planner willing to balance the short-run gains from specialization in a city's comparative advantage(s) against the subsequent productivity losses in the longer run.

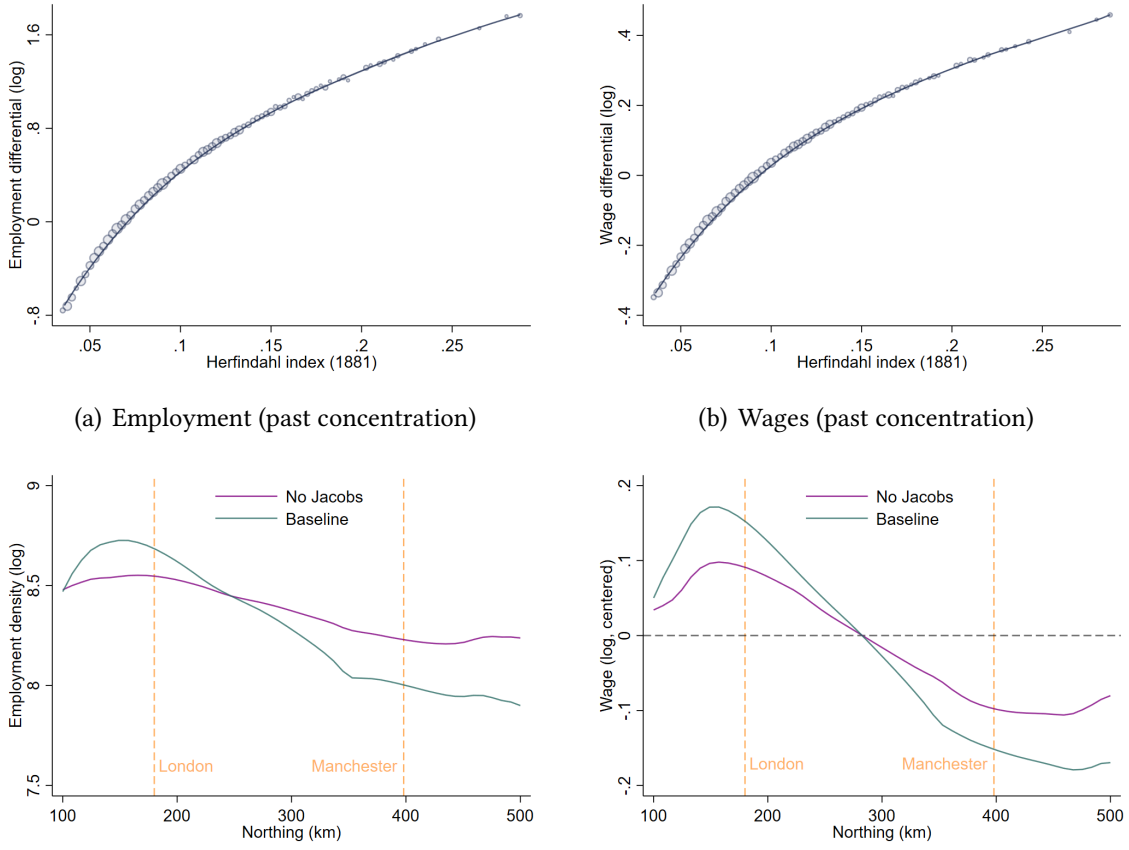
Neutralizing Jacobs externalities In this counterfactual, we offset long-run Jacobs externalities by changing the city-industry productivity levels of period 1 to,

$$\widetilde{\overline{T}}_{ic1} = \overline{T}_{ic1} \left[\sum_j \left(\frac{L_{jc0}}{L_{c0}} \right)^2 \right]^{\iota},$$

which amounts to eliminating the Herfindahl index from Equation (17) and simulating the evolution of an economy where industrial concentration does not negatively affect productivity in the subsequent "generation". Note that we solve for the equilibrium of counterfactual economies by using the procedure laid out in Section 4.2.

How does neutralizing Jacobs externalities affect the distribution of employment and income across cities? In the top panels of Figure 11, we show how employment (panel a) and wages (panel b) change as a function of their 19th-century industrial concentration. Unsurprisingly, we find that cities with a lower degree of initial concentration lose, while cities with a higher degree of specialization gain. The effects are sizable. A city at the 20th percentile of the 19th-century Herfindahl distribution (with a Herfindahl index of 0.06) loses 25% of its employment and 20% of its wages, while a city at the 80th percentile (with a Herfindahl index of 0.13) would markedly grow and gain 12% in wages.

Figure 11. Neutralizing long-run Jacobs externalities.



(a) Employment (past concentration)

(b) Wages (past concentration)

(c) Employment (north/south)

(d) Wages (north/south)

Notes: Panel (a) displays the relationship between industrial concentration in 1881 and the change in total employment from the period-1 data to the counterfactual equilibrium across our 435 cities. Panel (b) displays the same relationship between industrial concentration in 1881 and the change in nominal wages (per capita). Panels (c) and (d) nest cities along the "Northing" level in the British National Grid and display the best non-linear fit between average population density/wage against "Northing". For the sake of exposition, we display the Northing levels of London and Manchester as dashed (orange) lines. The baseline is represented by the green line; the counterfactual is the purple line.

These redistributive effects are substantial, but best illustrated in space or in meaningful measures of (regional) inequalities. One would not care so much about a city A exchanging its advantage and giving it to city B if it leaves spatial inequalities un-

changed. A marker of regional inequalities in Britain is the North/South divide, or how London, its hinterlands, and a few other Southern cities like Bristol, are richer than the great Northern cities of the past (e.g., Liverpool, Manchester, Sheffield, Newcastle or Bradford). We use this north-south divide as a way to benchmark the redistributive effects of our counterfactual experiment. In the bottom panel, we nest cities using their Northing in the British National Grid (kilometers relative to a referential point within the projected coordinate system) and show the average population density and average wage along this latitude. For the sake of exposition, we display two lines corresponding to London (and Bristol) and Manchester (and thus Liverpool or Leeds). One can see that there is indeed a north-south divide—with a productivity gap of about 35% between London and Manchester (panel d). Neutralizing long-run Jacobs externalities does not only reshuffle productivity gains (and employment) across cities; it does so along a (most) relevant north/south divide: the productivity gap is about 60% as large in the counterfactual experiment (20% between London and Manchester). A substantial part of contemporaneous regional inequalities in Britain is the legacy of past industrial concentration.[24]

Another "fundamental" trade-off? The preceding discussion of Jacobs externalities highlights a trade-off faced by local planners. A city may benefit from specializing according to its comparative advantage(s)—an insight at the heart of many place-based industrial policies, such as the European Union's "Smart Specialisation Strategy", innovation clusters in the United States, or the numerous hubs promoted in the United Kingdom.[25] While such specialization may yield short-term gains, these gains might be obscured by long-term scars à la Jacobs.

Modeling place-based industrial policy in full is beyond the scope of our framework, which provides a careful representation of geography but a crude characterization of firm interactions. To capture the key trade-off, we introduce a stylized representation in which local policymakers can costlessly redistribute revenue-based total factor productivity across industries in their city according to:

$$\log\left(T^*_{ic1}\right) = \log\left(\overline{T}_{ic1}\right) + k\left[\log\left(\overline{T}_{ic1}\right) - \log\left(\overline{T}_{c1}\right)\right],$$

where $\overline{T}_{ic1}$ denotes productivity in industry i, city c, and period 1, and $\overline{T}_{c1}$ is the unweighted average (log) productivity across all industries in city c. A policy with $k > 0$

[24]The remainder might be partly explained by the distribution of amenities and connectedness; it is however essentially the outcome of location-specific productivities interacted with the life cycles of industries. In short, Northern cities *also* have a comparative advantage in the wrong industries.

[25]There is no direct motivation for such policies in our "frictionless" model that takes sectoral productivities as given. In practice, however, these policies may help overcome coordination failures or relocation frictions.

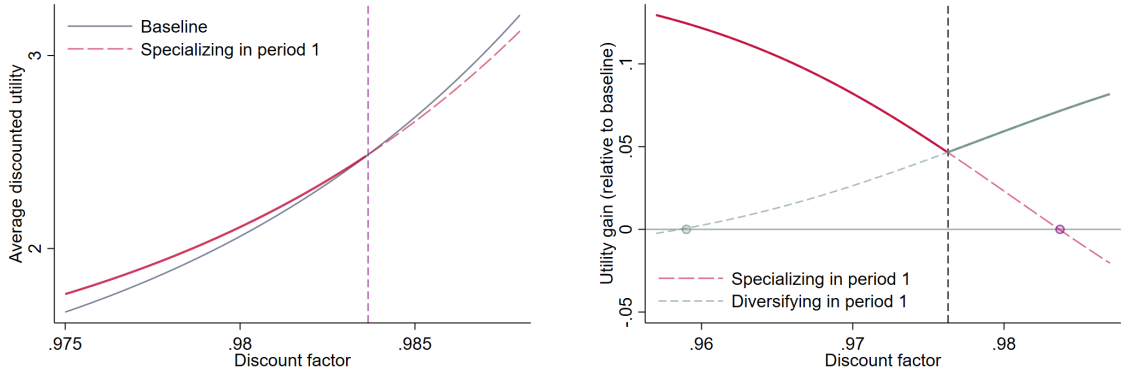
increases the productivity of industries in which the city has a comparative advantage—thus mimicking the effect of policies that promote industrial concentration.

All local planners implement this costless policy, and we evaluate outcomes using data for present-day England and Wales as the initial state. We then simulate the contemporaneous and long-term (period-2) distribution of economic activity—90 years later—keeping other city fundamentals, such as amenities, fixed. Welfare is measured across two generations of workers using expected indirect utilities, U_g, as:

$$\mathcal{V}(\beta) = U_1 + \beta^{90} \cdot U_2,$$

where $\beta \in (0, 1)$ is the planners' annual discount rate, and the economy grows at a constant rate of 2.1% per year (the annual growth rate of the British economy between 1950 and 2015).

Figure 12. Policies affecting local industrial concentration.



(a) Increasing industrial concentration

(b) Specializing/diversifying

Notes: Panel (a) displays the discounted indirect utilities across the two generations of workers, $\mathcal{V}(\beta)$, in the baseline (blue line) and in a counterfactual economy where $k = 0.10$ redistributes productive gains towards the most productive industries within each city c (red line). We highlight the range of discount factors β in which such a policy would generate gains relative to the baseline economy with a vertical dashed line separating the low discount factors where it is desirable (also displaying a plain red line in this desirable region). Panel (b) reports the discounted indirect utilities across the two generations of workers in logarithmic terms relative to the baseline, $\log \mathcal{V}^f(\beta) - \log \mathcal{V}(\beta)$, such that it approximately represents welfare gains in percentage points (i.e., a value of 0.05 would correspond to a welfare increase of about 5%). Panel (b) reports: (i) a counterfactual economy where $k = 0.10$ redistributes productive gains towards the most productive industries (in red); and (ii) a counterfactual economy where $k = -0.10$ redistributes productive gains towards the least productive industries (in green). We highlight the range of discount factors β in which such policies would generate gains relative to the baseline economy by displaying a blue line corresponding to $y = 0$ (and dots in green and in red); the vertical dashed line separates the range of discount factors where the specializing/diversifying are preferred by the planner.

Figure 12 illustrates the trade-off between short-run gains from boosting comparative advantages and long-run losses from the resulting local industrial concentration. Panel (a) shows that a place-based industrial policy fostering industrial concentration across all cities is welfare-enhancing when the planner places relatively little weight on future generations. This could help explain the widespread adoption of such policies,

particularly when planners are relatively impatient. For example, the UK government has recently used social discount rates of 2.5% for projects at over a 75 year horizon—a rate that is typically much higher in the United States. Panel (b) complements this insight by (i) providing a quantification of welfare gains relative to the baseline and (ii) showing the impact of a diversification policy ($k = -0.10$), in which productivities are redistributed away from stronger industries and towards weaker sectors in each city. In line with the previous insight, the diversification policy appeals more to patient planners, while the specialization policy is preferred by impatient ones. Interestingly, *both* policies improve welfare relative to the baseline within a range of discount factors, $\beta \in [0.96, 0.984]$, with minimum gains just under 4.5% at the point where the planner is indifferent between them.

6 Concluding remarks

A large literature discusses the rise and fall of cities, often in light of former industrial localities (e.g., Detroit or Manchester) or factory towns. Another literature focuses on the source of regional inequalities. Our research discusses the inter-temporal allocation of economic activity across space and highlights one factor which shapes (i) the rise and fall of cities and (ii) contemporaneous spatial inequalities: the extent to which (industrial concentration and) long-run externalities à la Jacobs affect future productivity.

Our present research however abstracts from modeling the exact mechanics of Jacobs externalities. Do they operate through the process of technological innovation and the useful transfer of knowledge across industries? Do they operate through the accumulation of human capital (technical aptitudes and entrepreneurial attitudes)? Would we expect them to operate irrespective of the nature of industrial concentration and their associated skills? We leave these questions for future research.

References

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales, "Shift-share designs: Theory and inference," *The Quarterly Journal of Economics*, 2019, *134* (4), 1949–2010.

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf, "The economics of density: Evidence from the Berlin Wall," *Econometrica*, 2015, *83* (6), 2127–2189.

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.

Alvarez-Palau, Eduard J., Dan Bogart, Max Satchell, and Leigh Shaw-Taylor, "Transport and urban growth in the first industrial revolution," Technical Report 2013.

Arribas-Bel, Daniel, M-À Garcia-López, and Elisabet Viladecans-Marsal, "Building (s and) cities: Delineating urban areas with a machine learning algorithm," *Journal of Urban Economics*, 2021, *125*, 103217.

Bairoch, Paul and Gary Goertz, "Factors of urbanisation in the nineteenth century developed countries: a descriptive and econometric analysis," *Urban Studies*, 1986, *23* (4), 285–305.

Baum-Snow, Nathaniel and Lu Han, "The microgeography of housing supply," *Journal of Political Economy*, 2024, *132* (6), 1897–1946.

Bellefon, Marie-Pierre De, Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, and Clément Gorin, "Delineating urban areas using building density," *Journal of Urban Economics*, 2021, *125*, 103226.

Berkes, Enrico, Ruben Gaetani, and Martí Mestieri, "Cities and Technological Waves," Technical Report, mimeo 2021.

Bernard, Andrew B., Jonathan Eaton, J. Bradford Jensen, and Samuel Kortum, "Plants and Productivity in International Trade," *American Economic Review*, 2003, *93* (4), 1268–1290.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-experimental shift-share research designs," *The Review of Economic Studies*, 2022, *89* (1), 181–213.

Brinkman, Jeffrey C, "Congestion, agglomeration, and the structure of cities," *Journal of Urban Economics*, 2016, *94*, 13–31.

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.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and labor market dynamics: General equilibrium analysis of the China trade shock," *Econometrica*, 2019, *87* (3), 741–835.

Caprettini, Bruno and Hans-Joachim Voth, "Rage against the machines: Labor-saving technology and unrest in industrializing England," *American Economic Review: Insights*, 2020, *2* (3), 305–20.

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.

Cobb, Colonel Michael H., *The Railways of Great Britain: A Historical Atlas*, Ian Allen Publishing, 2003.

Coeurdacier, Nicholas, Florian Oswald, and Marc Teignier, "Structural Change, Land Use and Urban Expansion," Technical Report 2022.

Combes, Pierre-Philippe and Laurent Gobillon, "The empirics of agglomeration economies," in "Handbook of regional and urban economics," Vol. 5, Elsevier, 2015, pp. 247–348.

_ **,** _ **, and Yanos Zylberberg**, "Urban economics in a historical perspective: Recovering data with machine learning," *Regional Science and Urban Economics*, 2022, *94*, 103711.

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.

Desmet, Klaus, David Krisztian Nagy, and Esteban Rossi-Hansberg, "The Geography of Development," *Journal of Political Economy*, 2018, *126* (3), 903–983.

Drayton, Elaine, Peter Levell, and David Sturrock, "The Determinants of Local Housing Supply in England," Technical Report, mimeo 2024.

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.

_ **and** _ , "Micro-foundations of urban agglomeration economies," in "Handbook of regional and urban economics," Vol. 4, Elsevier, 2004, pp. 2063–2117.

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 2023.

Faggio, Giulia, Olmo Silva, and William C. Strange, "Heterogeneous agglomeration," *Review of Economics and Statistics*, 2017, *99* (1), 80–94.

Fajgelbaum, Pablo and Stephen J Redding, "Trade, structural transformation, and development: Evidence from Argentina 1869–1914," *Journal of Political Economy*, 2022, *130* (5), 1249–1318.

Fajgelbaum, Pablo D. and Cecile Gaubert, "Place-Based Policies: Lessons from Theory," NBER Working Papers 33517, National Bureau of Economic Research, Inc February 2025.

Feinstein, C. H., *National Income, Expenditure and Output of the United Kingdom, 1855–1965*, Cambridge University Press, 1972.

Franck, Raphaël and Oded Galor, "Flowers of Evil? Industrialization and Long-Run Development," *Journal of Monetary Economics*, 2021, *117*, 108–128.

Galor, Oded and Ömer Özak, "The agricultural origins of time preference," *American Economic Review*, 2016, *106* (10), 3064–3103.

Gaubert, Cecile, Patrick M Kline, Damián Vergara, and Danny Yagan, "Place-Based Redistribution," Working Paper 28337, National Bureau of Economic Research January 2025.

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.

Graham, Daniel J. and Stephen Gibbons, "Quantifying Wider Economic Impacts of Agglomeration for Tranport Appraisal: Existing Evidence and Future Directions," *Economics of Transportation*, 2019, *19*, 100121.

Hanlon, W Walker, "The rise of the engineer: Inventing the professional inventor during the Industrial Revolution," Technical Report, National Bureau of Economic Research 2022.

_ **and Antonio Miscio**, "Agglomeration: A long-run panel data approach," *Journal of Urban Economics*, 2017, *99*, 1–14.

Harari, Mariaflavia, "Cities in bad shape: Urban geometry in India," *American Economic Review*, 2020, *110* (8), 2377–2421.

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.

Head, Keith and Thierry Mayer, "The United States of Europe: A Gravity Model Evaluation of the Four Freedoms," *Journal of Economic Perspectives*, 2021, *35* (2), 23–48.

Heblich, Stephan, Alex Trew, and Yanos Zylberberg, "East-side story: Historical pollution and persistent neighborhood sorting," *Journal of Political Economy*, 2021, *129* (5), 1508–1552.

_ **, 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.

_ **, Stephen J Redding, and Yanos Zylberberg**, "The distributional consequences of trade: Evidence from the Grain Invasion," Technical Report, National Bureau of Economic Research 2024.

Heldring, Leander, James A Robinson, and Sebastian Vollmer, "The economic effects of the English Parliamentary enclosures," Technical Report, National Bureau of Economic Research 2022.

Henderson, J Vernon, "The sizes and types of cities," *The American Economic Review*, 1974, *64* (4), 640–656.

Henderson, Vernon, Ari Kuncoro, and Matt Turner, "Industrial development in cities," *Journal of Political Economy*, 1995, *103* (5), 1067–1090.

Herrendorf, 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 Roderick 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 Death and Life of Great American Cities*, New York: Random House, 1961.

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

Kelly, Morgan, Joel Mokyr, and Cormac Ó Gráda, "The mechanics of the Industrial Revolution," *Journal of Political Economy*, 2023, *131* (1), 59–94.

Kline, Patrick and Enrico Moretti, "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 11 2014, *129* (1), 275–331.

_ **and** _ , "People, Places and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs," *Annual Review of Economics*, 2014, *6* (1), 629–662.

Lin, Jeffrey, "Technological adaptation, cities, and new work," *Review of Economics and Statistics*, 2011, *93* (2), 554–574.

Litvine, Alexis, Arthur Starzec, Rehmana Younis, Yannick Faula, Mickaël Coustaty, Leigh Shaw-Taylor, and Véronique Églin, "An Integrated Methodology for Extracting High-resolution Urban Footprints from Historical Maps," Technical Report 2023.

Matsuyama, Kiminori, "Agricultural productivity, comparative advantage, and economic growth," *Journal of Economic Theory*, 1992, *58* (2), 317–334.

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 Krisztián, "Quantitative economic geography meets history: Questions, answers and challenges," *Regional Science and Urban Economics*, 2022, *94.*

Nagy, Dávid Krisztián, "Hinterlands, city formation and growth: Evidence from the US westward expansion," *Review of Economic Studies*, 2023.

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," *Cliometrica*, 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.

Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Sanderson, Eleanor and Frank Windmeijer, "A weak instrument F-test in linear IV models with multiple endogenous variables," *Journal of econometrics*, 2016, *190* (2), 212–221.

Satchell, Max, Dan Bogart, and Leigh Shaw-Taylor, "Parliamentary Enclosures in England, 1606–1902," Technical Report 2017.

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.

Shaw-Taylor, Leigh, Dan Bogart, and Max Satchell, "The Online Historical Atlas of Transport, Urbanization and Economic Development in England and Wales c.1680-1911," Technical Report 2024.

Simonovska, Ina and Michael E. Waugh, "The Elasticity of Trade: Estimates and Evidence," *Journal of International Economics*, 2014, *92* (1), 34–50.

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.

Thomas, Ryland and Nicholas Dimsdale, "A Millenium of Macroeconomic Data for the UK," 2018.

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.

Wrigley, E. Anthony, "The PST System of Classifying Occupations," Technical Report 2010.

ONLINE APPENDIX—not for publication

A Data appendix

This section provides complements to Sections 2 and 3: we present our data sources, with a detailed description of our map digitization procedure; we discuss the land fragmentation algorithm underlying our exogenous predictor of city growth; we provide descriptive statistics about the rise of cities and their subsequent dynamics; and we discuss a series of robustness checks.

A.1 Data sources and construction

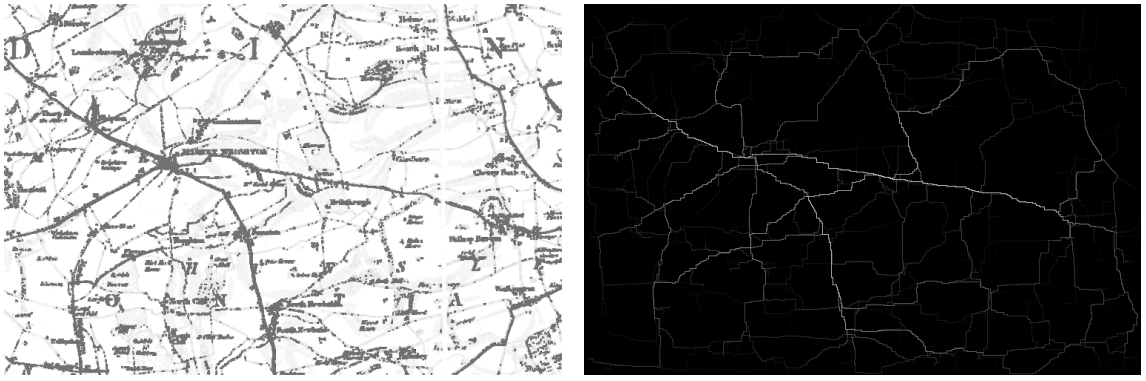
A large collection of county maps and Ordnance drawings We first collect and digitize early county maps from the North of England. Please find below the covered counties, with the map resolution, its date and its author(s): Northumberland, 1-inch map, 1769 (Andrew and Mostyn Armstrong); Durham, 1-inch map, 1768 (Thomas Jefferys and Andrew Armstrong), 1-inch map, 1820 (Christopher Greenwood); Cumberland, 1/2-inch map, 1773 (Joseph Hodskinson and Thomas Donald), 1-inch map, 1823 (Christopher and John Greenwood); Westmorland, 1-inch map, 1771 (Thomas Jefferys); Yorkshire, ca 4/5 inch, 1828 (Christopher Greenwood, et al.); Lancashire, not stated, 1768 (William Yates), 1 inch, 1818 (Christopher Greenwood, et al.), 4/5 inch, 1830 (G. Hennet); Cheshire, 3/4 inch, 1830 (William Swire and W.F. Hutchings), 5/4 inch, 1831 (Andrew Bryant); Derbyshire, 1-inch map, 1767/1791 (Peter Burdett); Nottinghamshire, not stated, 1794 (John Chapman), not stated, 1826 (Christopher and John Greenwood); Lincolnshire, ca. 1 inch, 1828 (Andrew Bryant), ca. 1 inch, 1830 (Christopher and John Greenwood). We carefully project these flat map tiles, which gives us an exhaustive coverage of the North of England, albeit at different dates and resolution between 1790–1830. To cover Southern England and Wales, we rely on about 350 Ordnance Survey drawings that were produced by the Ordnance Survey at the beginning of the nineteenth century and are usually referred to as the Old Series 1-inch maps. We keep the latest series when maps fully or partially overlap, leaving us with about 360 different map tiles produced between 1790–1820 for most of them.

Recognizing built-up from county maps and Ordnance drawings These different maps employ different annotations, symbols, colors, etc. For instance, county maps typically display built-up with dark rectangles, sometimes with dotted areas, but are otherwise quite clear and organized (see Figure 5, for instance). The challenge is then to distinguish built-up from letters or dashed lines representing administrative boundaries. In stark contrast, Ordnance Survey drawings typically display built-up with red marks, but the drawings also contain the delineation of farms and topography (which is

not represented using contour lines of similar altitude, but rather "slope lines", i.e., the orthogonal lines indicating the typical course of water along the slope). In summary, our main object of interest (a building) is displayed across all maps, but with varying symbols. For these reasons, we identify built-up using supervised deep learning with a large training sample.

To recognize built-up and other land uses in historical maps, we develop a convolutional neural network frequently used in the biomedical literature to detect zones of interest within images (a "U-Net").[26] We train the algorithm using a simple manual labeling procedure with research assistants annotating images with contours, and we complement our training sample with previous training data from French historical maps. We add basic transformations (i.e., flipping, zooms, rotations) to each training batch, such as to avoid overfitting and ensure invariance to the orientation and resolution of maps.

Figure A1. Recognizing roads.



(a) "Travel" cost (b) Paths

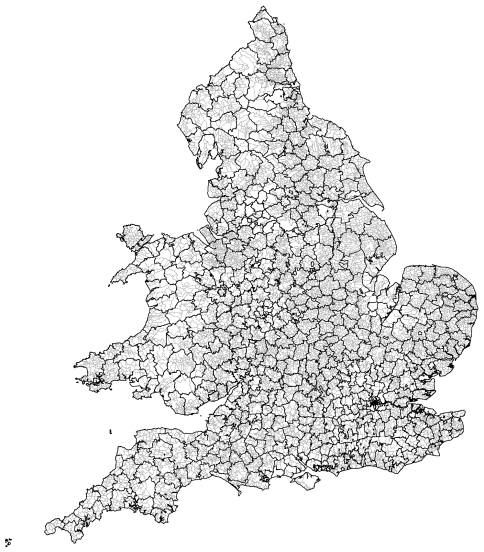
Note: Panel (a) displays the input of the least cost path procedure to detect roads on a historical county map. Panel (b) displays the simulated least cost paths drawn between random departure and arrival points.

To identify roads, we develop a less standard approach: thin, non-straight lines are notoriously difficult to detect using computer vision. Our approach relies on the nature of roads: they are black lines on the map, but, in practice, they are designed to best facilitate transportation across space. In that respect, they markedly differ from all other black lines on the maps, e.g., the text or gradient lines in mountainous terrain. We exploit this discriminatory feature as follows: (i) we create a "stylized travel cost" matrix for each map, penalizing travel outside of darker, black pixels; (ii) we draw many random departure points from black pixels on the image and compute the least cost paths to

[26]Two remarks in order: Litvine et al. (2023) employ a similar approach to extract urban footprints in Britain from these Ordnance Survey drawings and individual county maps; an alternative method is to consider an object-based classification which is standard in remote sensing but less common in map digitization (see, e.g., Combes et al., 2022, for a discussion of these methods).

each other pixels of the image for each draw; (iii) for each draw, we then randomize many arrival points and start drawing the actual minimum paths between departure and arrival pixels.[27]

Figure A2. 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 or poor law unions are displayed with black borders.

Consistent administrative units Our analysis relies on various levels of administration, spanning a long period. Section 2.2 describes how we identify and delineate cities

[27]Our approach relies on the nature of roads: roads are contiguous (impervious) areas which are designed to minimize travel cost between locations of interest. In practice, we can design a procedure to detect black lines, but we would then typically end up with many false positives (i.e., black lines that are not roads). Many of these black lines would, however, not fulfill the previous travel-cost minimization requirement: letters, gradient lines in mountainous terrain or farms are very inefficient patches of black pixels on which to travel between locations of interest. Others may be better suited, e.g., rivers, county boundaries, or railway tracks. We exploit the previous discriminatory feature as follows. We first transform the color map into a gray "travel cost" matrix by (i) filtering out non-gray colors, (ii) thickening black areas to create continuous lines from dashed lines, (iii) simplifying images by interpolating across pixels, and (iv) transforming the gray intensity through a power function in order to best calibrate the "cost" of traveling across pixels of different black intensity (with travel cost over a perfectly black pixel being 1, and travel cost over a perfectly white pixel being as high as possible). The output is the left panel of Figure A1. Second, we draw many random departure points from the subsample of black pixels on the image (as a proxy for the unobserved locations of interest) and we compute the least cost paths to all other pixels of the image for each draw. Roads are already more salient on this matrix. However, roads are designed to be used as "the" minimum cost path between locations: we randomize many arrival points for each departure point and start drawing the actual minimum paths between departure and arrival pixels on a black image. The output is a distribution of many minimum cost paths highlighting the most traveled roads more than the others (right panel of Figure A1). As shown by Figure A1, this algorithm not only identifies roads very well, but also their respective importance within the transportation network.

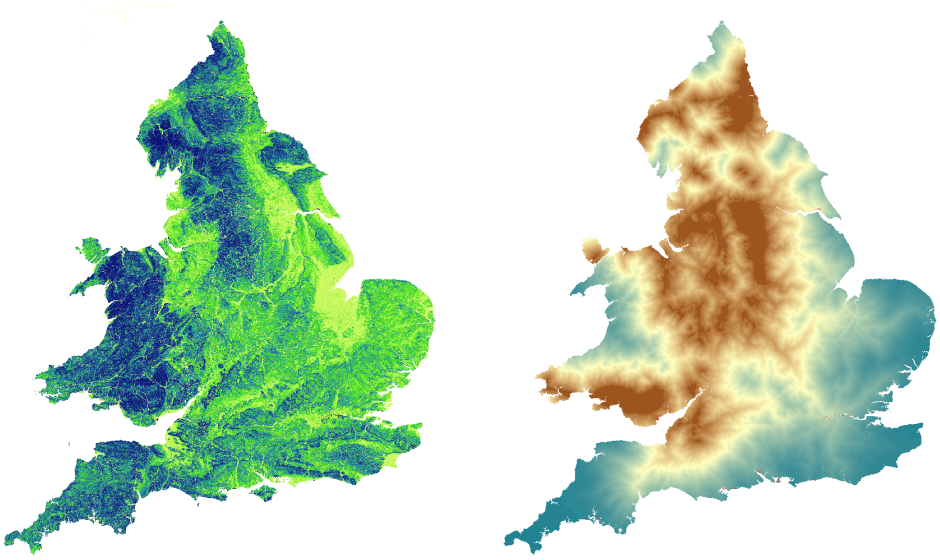
of the nineteenth century, a city shape that we use to nest economic outcomes in the recent period. These cities are the core administrative units of our study.

They are, however, not the only administrative divisions underlying our empirical analysis. First, our main specification (see, e.g., Tables 2 or 3) includes fixed-effects at the level of 48 ceremonial counties whose delineation is set in recent times. Second, in most of the empirical analysis, we report standard errors that are clustered at the level of the closest city as of 2015, a formal status attributed to only 55 cities. In this instance and in the previous one, we use recent geography to better control for spatial auto-correlation within contemporary administrative divisions. Third, while we nest most economic outcomes within the borders of our cities and expanded cities, census data and baptism records are originally provided at the parish level, and parishes are regularly redefined, merged or split over the course of the nineteenth century. To limit the importance of differential measurement error over time, we apply a preliminary "envelope" algorithm which considers the union of the different parishes covering the same points over time. For instance, if a parish is split into two parishes in 1891, we would group the two sub-parishes from 1891 onward to create a consistent, unique parish from 1801 to the current day. This grouping is less relevant at the level of higher administrative divisions, as none of these re-compositions of lowest administrative units significantly affect the allocation of administrative units across higher divisions. The output of the procedure is shown in Figure A2, with about 11,500 consistent parishes across England and Wales.

Transportation, travel time, and trading costs Both the derivation of stylized facts and the quantification of our model rely on transportation infrastructure in the nineteenth century and in recent times. Our dynamic characterization of historical transportation relies on transportation infrastructure from the Cambridge Group for the History of Population and Social Structure (including the roads in 1830, density of train lines in 1851, and waterways around 1820, all used to condition the main empirical specification). The data is best described in Alvarez-Palau et al. (2013). The transportation infrastructure is integrated into a transportation network in order to derive a few important controls used, e.g., in Tables 2 or 3: the travel times to the closest city, the closest market town, major ports (London, Liverpool, Plymouth, Portsmouth), and resources (i.e., coal).

The transportation network is also important to derive the trading costs in the quantitative model. We describe this derivation in greater detail in Appendix B.5.

Figure A3. Travel cost and distance to coal across England and Wales, as computed around 1817.



(a) Travel cost (b) Distance to coal

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).

Other geographic variables The derivation of stylized facts and the identification of dynamic Jacobs externalities hinge on numerous geographic factors, either to extract exogenous variation in city growth or to condition the analysis on important confounders. First, we consider the following topographic characteristics: elevation (Open Land Map, 30m resolution), used as a main control and as input in the land fragmentation algorithm together with the derived slope; soil organic carbon content (Open Land Map, 250m resolution), used as a main control; soil bulk density (Open Land Map, 250m resolution), used as a main control; a detailed soil classification (from The National Soil Map and Soil Classification, produced by the National Soil Resources Institute); and a dataset of all rivers and smaller streams in England and Wales (OS Open Rivers). Second, we combine our city identification and delineation (Section 2.2) with geographic characteristics to derive the area/latitude/longitude of the initial city outline, the predicted expanded area in 1881, and the built-up density within 1, 2, 3, 5, 10 kms of the initial city outline. Third, we measure agricultural productivity around our cities by extracting the suitability to grow wheat, oat, grass, and rye from the Global Agro-Ecological Zones produced by the Food and Agriculture Organization of the United Nations—we use the rain-fed, high-input model specification. Fourth, we use a measure of baseline share of arable agriculture (extracted from the Tithe survey, see Heblich et al., 2024) for a similar purpose,

i.e., controlling for initial agricultural production within the city hinterlands. Fifth, we collect information on historical factors possibly anchoring human settlements and the distribution of land ownership across space, most notably the presence of Domesday settlements (1086) extracted from the revision of the Phillimore edition of Domesday Book (1975–1986) and enclosures within 1, 2, 3, 5, 10 kms of the initial city outline (Satchell et al., 2017).

A.2 Land fragmentation

Our land fragmentation algorithm consists of the following steps. In a first step, we extract fine-grained terrain characteristics including elevation, ruggedness, time-invariant soil attributes and water bodies from Google Earth Engine. We then store these characteristics in an "image", i.e., a 2D array where the value at each pixel (x, y) is a vector of attributes $(a_1, \ldots, a_m)$—a standard color image is often stored as a 2D array where the value at each pixel is a triplet RGB. We then normalize the different attributes to be between 0 (minimum) and 1 (maximum). The output is a multi-band raster covering England and Wales at a resolution of 30m (bands: normalized elevation, normalized slope, normalized bulk, discrete soil classes, discrete zones delineated by waterways).

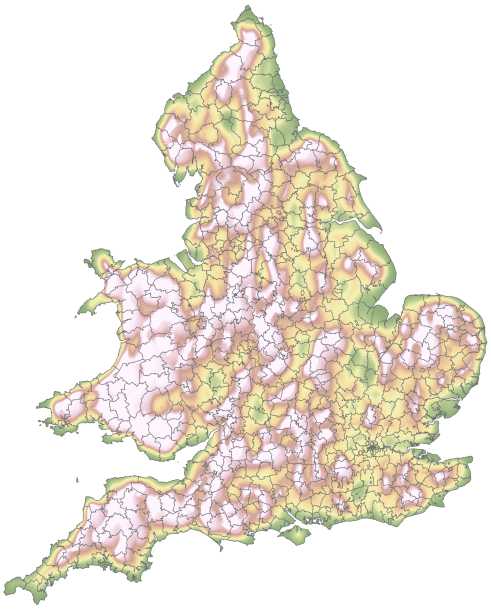
Figure A4. Topography, agricultural parcels and land fragmentation.

Notes: This figure displays the Ordnance Survey drawing around the small village of Lymington—between the New Forest National Park and the Isle of Wight. Red rectangles indicate our detected built-up; dashed lines show the identified gradient breaks in soil conditions and topography.

The second step produces superpixels to isolate contiguous zones of the image that are homogeneous in their attributes. There are several image segmentation algorithms,

e.g., SLIC, Quickshift, Felzenszwalb or watershed.[28] We opt for Quickshift, which allows for a grouping of pixels that is more flexible in their proximity in actual space (i.e., along the physical distance) versus the color space. Applied to this peculiar setting, the algorithm relies on a standard "orthogonal" distance between the n-dimensional vectors that are stored in every pixel as the "distance in the color space" or "value distance." The algorithm maximizes a weighted sum of the two target distances within constituted superpixels, with a weight allocated to physical distance relative to the previously-defined "value distance." We parametrize the algorithm by choosing the scale parameter, the maximum physical distance, and the relative weight between distance in the color-space and physical distance in order to best capture the shape of agricultural parcels. A superpixel with similar values is, here, a reasonably-compact patch of land with homogeneous topography and soil characteristics: a typical agricultural parcel. For illustration, we show in Figure A4 the output of such an algorithm around the small village of Lymington—between the New Forest National Park and the Isle of Wight.

Figure A5. Predicted land fragmentation.



Note: This Figure displays an interpolated, aggregated raster map of predicted land fragmentation. The color scheme is an elevation color scheme: from green (low) to yellow, red and white (high). One can see that mountainous areas are typically more fragmented, but the low-lying hills of Cornwall or Cotswold, or the soil heterogeneity in Norfolk also generates significant land fragmentation.

In a third step, we construct the centroids of all superpixels, and we use these centroids to calculate the predicted density of farms at the fringe of each city—a measure of predicted land fragmentation. The variation in predicted land fragmentation is very local; we do however see aggregate patterns due to overall topography (mountainous

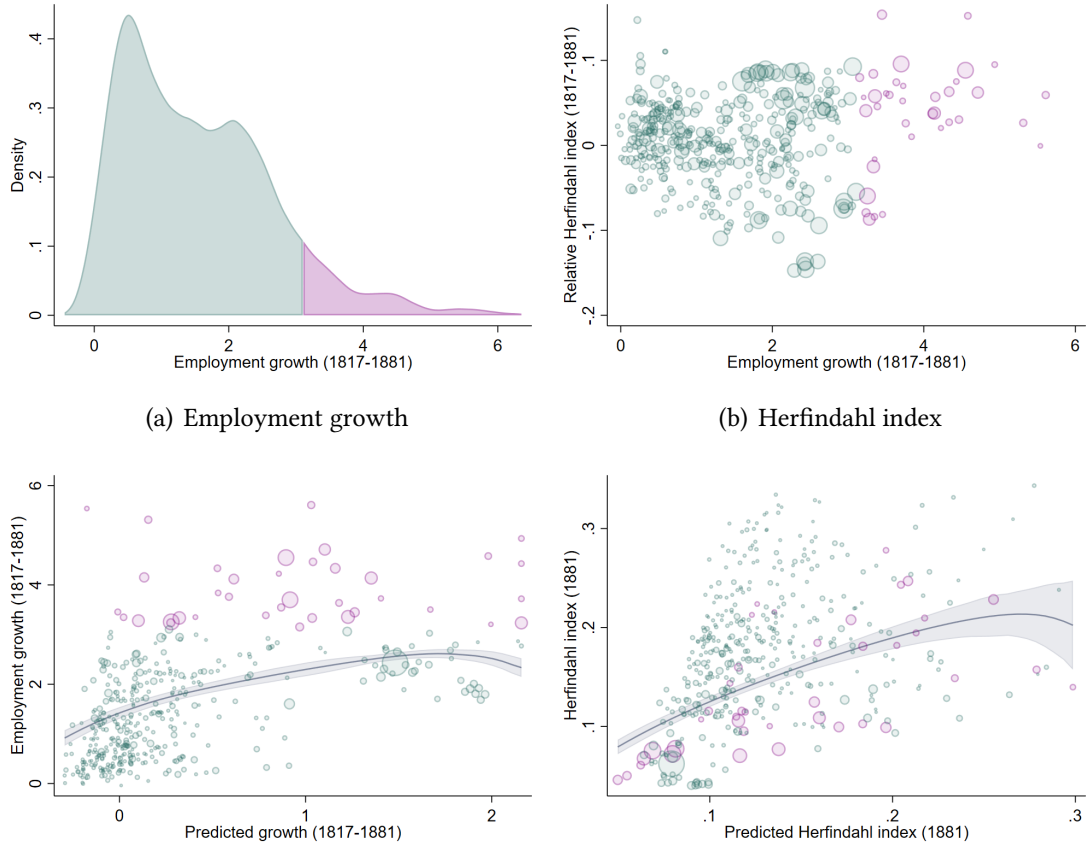
[28] See, e.g., a description of these segmentation algorithms.

terrain in Central England or Wales, hilly terrain in Cornwall) or soil composition (e.g., in Norfolk). We illustrate these aggregate patterns in Figure A5 and validate the measure against actual measures of land ownership in Figure 7 of the paper.

A.3 Descriptive statistics

In this section, we provide complements to Section 2.2 by focusing on the (nature of the) rise of cities.

Figure A6. Growth, industrial concentration, and the role of the initial industrial mix.



(a) Employment growth

(b) Herfindahl index

(c) Predicted employment growth

(d) Predicted Herfindahl index

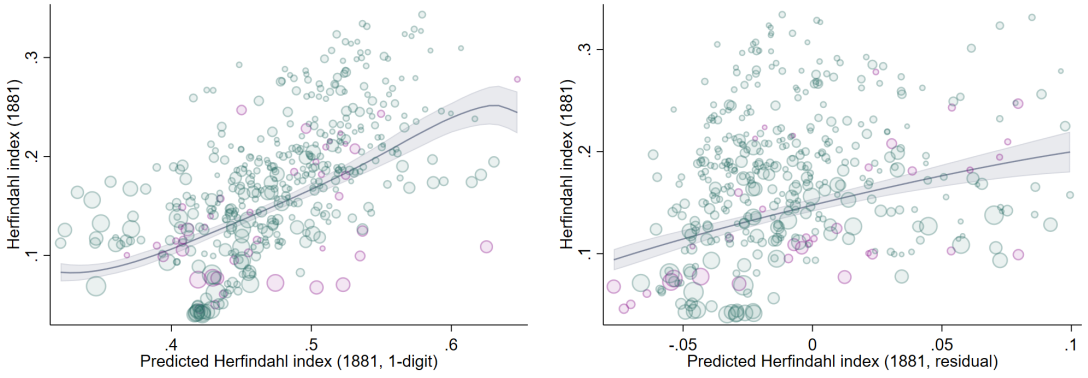
Notes: panel (a) displays the distribution of employment growth, $\frac{L_c^{81}-L_c^{17}}{L_c^{17}}$, within the (larger) boundaries of 435 potential cities. The top decile of growing settlements between 1817–1881 is highlighted in purple. In panel (b), we show the relationship between employment growth and the evolution of industrial concentration, $\sum_j \left(s_{jc}^{81}\right)^2 - \sum_j \left(s_j^{81}\right)^2$, across urban areas—where s_{ic}^{81} is the employment share in industry i and city c, and s_i^{81} is the aggregate employment share in industry i. Panel (c) displays the measure of predicted employment growth, g_c, against the actual employment growth across our 435 potential cities. The top decile of growing settlements are highlighted in purple. The line is a locally weighted regression on all observations, with the weight being the level of employment in 1817. Panel (d) repeats the same exercise with the actual and predicted industrial Herfindahl index, χ_c.

The rise of (different) cities Figure A6 shows the distribution of employment growth and industrial concentration across our urban settlements. We see that the median set-

tlement grows by a factor of about 3. The growth rate for the top decile (in purple) is above 300% (panel a). The large heterogeneity in employment growth is also reflected in a large heterogeneity in industrial concentration. Most of the fastest-growing cities become more specialized (panel b), an effect which is explained by the fact that fast-growing cities are initially specialized across the few key sectors that flourish over the course of the nineteenth century. Importantly, most cities of similar size in 1817 have Herfindahl indices that could differ by up to 0.20 in 1881—equivalent to the difference between a city with employment equally distributed across 3 (2-digit) sectors and a city with employment equally distributed across 10 such sectors.

We illustrate the role of initial industries, as proxies for initial location advantages, in panels (c) and (d) of Figure A6 where we compare the actual employment growth and industrial concentration in 1881 with their predictions, g_c and χ_c, as computed using a "shift-share" design based on initial industries in 1817, described above. We find that the initial industrial mix does predict employment growth within urban areas but also, and more significantly, industrial concentration.

Figure A7. The role of the initial industrial mix—robustness checks.



(a) Herfindahl index (1881, 1-digit)

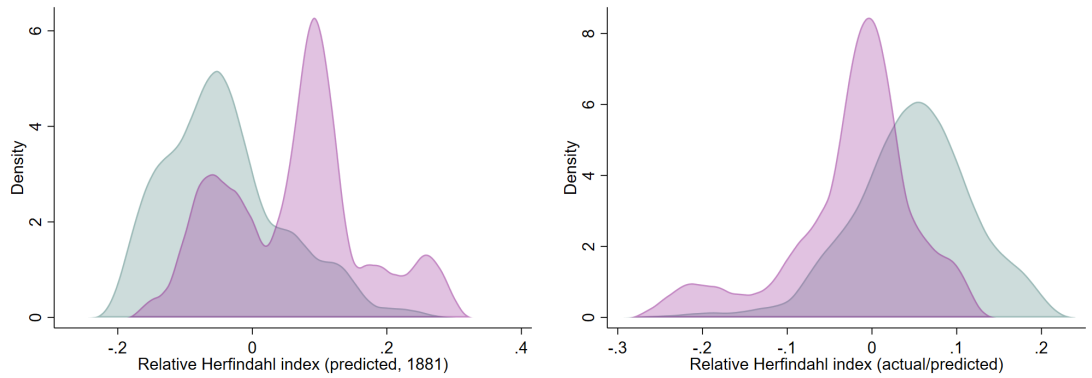
(b) Herfindahl index (1881, residual)

Notes: The left panel displays the measure of predicted industrial concentration computed at the 1-digit level against the actual industrial concentration. The top decile of growing settlements are highlighted in purple. The line is a locally weighted regression on all observations. The right panel repeats the same exercise with a residualized measure of predicted industrial concentration that is (i) computed at the 2-digit level and (ii) cleaned for the role of the previous 1-digit measure of predicted industrial concentration. The slope is about 60% to that of Figure A6 (panel d).

Both our actual and predicted indices of industrial concentration are based on 2-digit industries. We now shed additional light on the role of 1-digit sectors in cities' changing industrial composition as opposed to the more granular variation within these 1-digit sectors. In Figure A6, we display the relationship between the 1881 Herfindahl index of industrial concentration computed at the 2-digit level and a predicted measure, also computed at the 2-digit level. In Figure A7, we decompose this relationship into (i) the

part explained by predicted industrial concentration at the 1-digit level (panel a) and (ii) the residual of the 1-digit prediction, thus exploiting within-sector variation only (panel b). We can see that the relationship remains strong in panel (b); the slope is about 60% to that of Figure A6.

Figure A8. Specialization of cities—prediction and residual Herfindahl.



(a) Predicted Herfindahl index

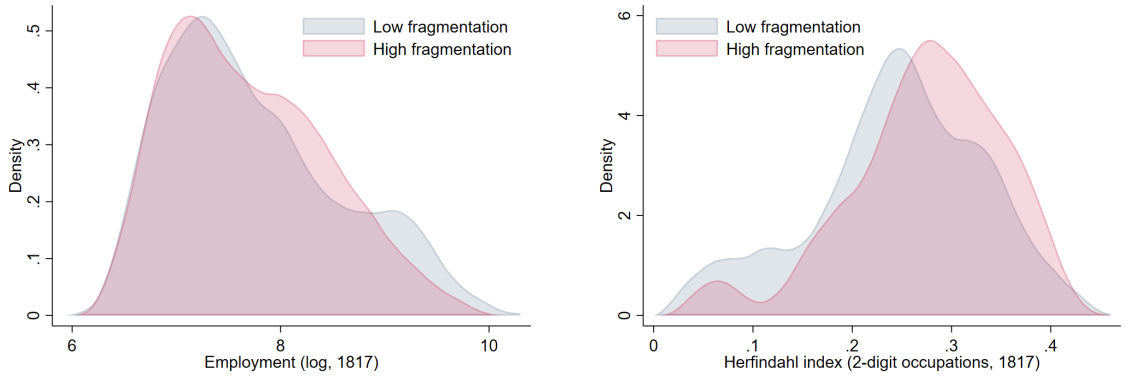
(b) Residual Herfindahl index

Notes: This Figure shows the distribution of predicted industrial concentration χ_c (see Section 2.2, left panel) and residual industrial concentration $h_c - \chi_c$ (right panel). The top decile of growing settlements are highlighted in purple.

The previous evidence shows that fast-growing cities become more specialized. We better qualify this observation in Figure A8 where we decompose industrial concentration h_c (as shown in Figure A6) into predicted industrial concentration χ_c (see Section 2.2) and residual industrial concentration $h_c - \chi_c$. We then plot the distribution of these two objects for the top quartile in terms of employment growth during 1817–1881 (in purple) and the rest (in teal). We see that the specialization of fast-growing cities is entirely predicted from their initial industry mix: fast-growing cities are initially specialized across a few key sectors that flourish over the course of the nineteenth century.

Exogenous variation in city size To isolate exogenous variation in late nineteenth century population, we use a proxy for land ownership fragmentation in the immediate fringe of urban settlements. In this appendix, we provide a balance test by illustrating the correlation between settlement characteristics at the beginning of the nineteenth century and predicted land fragmentation in these settlements' immediate fringe. We find that settlements with above- (red) and below-median (blue) predicted land fragmentation are very similar in population density and industrial concentration in 1817 (see Figure A9).

Figure A9. A balance test for urban settlements with different land fragmentation in their immediate fringe.

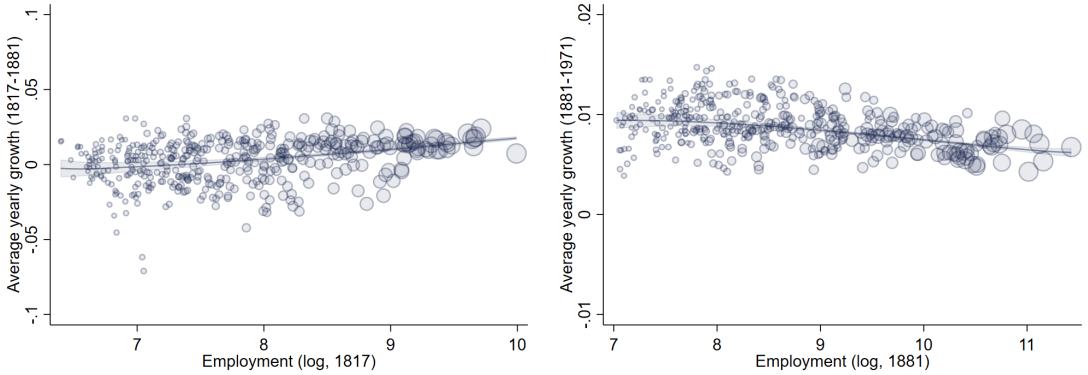


(a) Employment (1817)

(b) Herfindahl index (1817)

Notes: This Figure displays the distribution of population density (left panel) and industrial concentration (calculated as a Herfindahl index across 2-digit occupations, right panel) for cities with above- (red) and below-median (blue) predicted land fragmentation in their immediate fringe.

Figure A10. Gibrat's law.



(a) Gibrat's law, 1817–1881

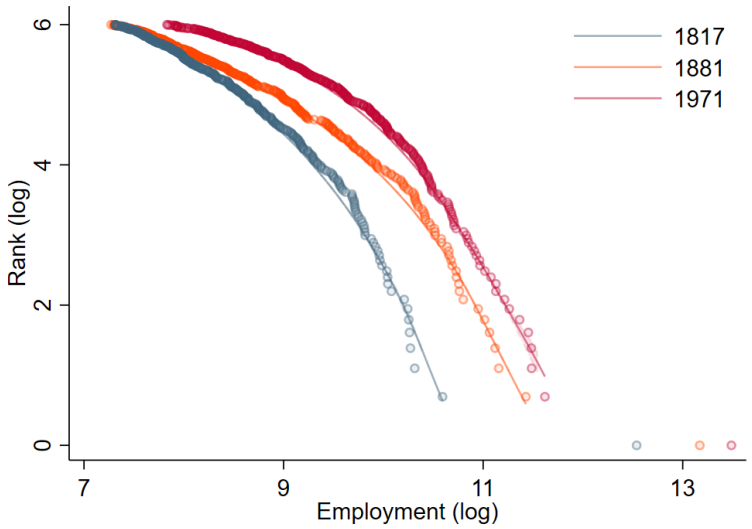
(b) Gibrat's law, 1881–1971

Notes: This Figure shows the relationship between cities' (log) initial employment and subsequent yearly growth. Panel (a) conducts the analysis for the period 1817–1881, while panel (b) conducts it for the period 1881–1971.

City population and city growth This section presents additional evidence about the distribution of city populations and growth over time. In Figure A10, we display the relationship between cities' average yearly growth and initial employment. Panel (a) shows this relationship in the nineteenth century. Initially large cities grow slightly faster during this period. Panel (b) repeats the analysis for the twentieth century. Over this period, the relationship becomes flat, even slightly negative: that is, large cities lose their advantage in terms of growth. This is in line with the empirical findings of Sec-

tion 3.2: larger population does not confer a significant advantage in terms of long-run growth. It is also in line with the urban literature documenting a typically flat relationship between city size and growth (Gibrat's Law).

Figure A11. Zipf's law.



Notes: This Figure shows the relationship between cities' (log) rank in 1817 (blue), 1881 (orange) and 1971 (purple) and their (log) employment. The estimated slope is -1.20 in 1817, -0.95 in 1881, and -1.03 in 1971.

We also provide support for another empirical regularity characterizing the distribution of city size: the Zipf law relating the rank of cities to their size. In Figure A11, we plot the relationship between the (log) rank and the (log) employment across our cities in 1817 (blue), 1881 (orange) and 1971 (purple). One can see that the relationship is close to linear in all years; and the estimated slope is -1.20 in 1817, -0.95 in 1881, and -1.03 in 1971. In summary, we find that the Zipf's law of city size holds reasonably well over the course of two centuries for Great British cities.

A.4 Robustness checks

In this section, we provide a sensitivity analysis for our main finding of Section 3.2.

Identification hypothesis In our baseline specification, we provide evidence that land fragmentation affects urban growth even when controlling for competing mechanisms, more specifically: general agricultural productivity (Coeurdacier et al., 2022); and the exposure to the repeal of the Corn Laws (Heblich et al., 2024). Table A1 reports a sensitivity analysis around our baseline specification (Table 3, column 2). In column (1), we consider a more parsimonious specification with fewer (geographic) controls. In col-

Table A1. The long-run effect of industrial concentration and population—sensitivity to a parsimonious specification and to additional geographic controls.

Wage (2020)	(1)	(2)	(3)	(4)	(5)
Herfindahl index (1881, h_c)	-2.852	-2.794	-2.293	-2.624	-2.624
	(1.020)	(1.078)	(0.911)	(0.889)	(0.833)
Employment (1881, l_c)	-0.278	-0.338	-0.232	-0.309	-0.219
	(0.169)	(0.226)	(0.204)	(0.176)	(0.200)
Observations	428	428	422	428	428
F-stat (h_c)	14.16	9.00	13.85	14.89	13.61
F-stat (l_c)	11.98	8.74	12.12	14.83	11.75
Controls	Parsimonious	Longitude	Agr. prod.	Land value	Enclosures

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). The dependent variable is the (log) wage in 2020 (labor income estimates for small areas based on the Family Resources Survey). Column (1) reports the estimates of a more parsimonious specification with: dummies corresponding to deciles of employment in 1817, deciles of industrial concentration in 1817, and deciles of agricultural employment share in 1817; (log) population density in 1821; the industry-based shift-share, g_c, described in Section 2.2; a shift-share control based on employment shares in 1881 and aggregate employment growth between 1881–1971; (log) area of the initial city outline; average elevation; slope, bulk density, and latitude; travel time to the closest city, to each major port (London, Liverpool, Plymouth), to coal; density of roads (1830), train lines (1851), and waterways; share of arable agriculture (Tithe survey); suitability to grow wheat, grass; share of heavy soil (NSRI); and the predicted land ownership fragmentation within the contours of the settlement around 1800 and within a 10-kilometer buffer. The following columns use the baseline controls (Table 3), but: column (2) adds a control for longitude; column (3) adds controls for agricultural productivity within the expanded city fringe (caloric potential, and bulk density); column (4) adds a control for land prices in 1815 (see Heblich et al., 2024); and column (5) adds a control for the presence of agricultural buildings in 1750, the number of Domesday buildings and the incidence of parliamentary enclosures within the contours of the settlement around 1800 and within 1, 2, 3, 5, 10 kms (Heldring et al., 2022). The two instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2 and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

umn (2), we control for longitude. In column (3), we control for the average Caloric Suitability Index (Galor and Özak, 2016) and for carbon content—both calculated within the expanded city fringe. In column (4), we control for land prices in 1815 (see Heblich et al., 2024). In column (5), we control for the presence of agricultural buildings in 1750, the number of Domesday buildings and the role of local institutions through the prevalence of parliamentary enclosures (Heldring et al., 2022). Across these specifications, the estimates for the Herfindahl index (1881) and employment (1881) are stable.

Our argument is that the industrial concentration effect operates through intertemporal externalities à la Jacobs, i.e., through the (lack of) diversity of geographically proximate industries, rather than secular trends in agricultural employment or manufacturing. In Table A2, we show that our findings are orthogonal to the process of structural transformation and the reallocation of labor across wider sectors (e.g., between agriculture and manufacturing, or towards services). More specifically, we add to our baseline specification: a Herfindahl index predicted as in Section 2.2, but with 1-digit sectors

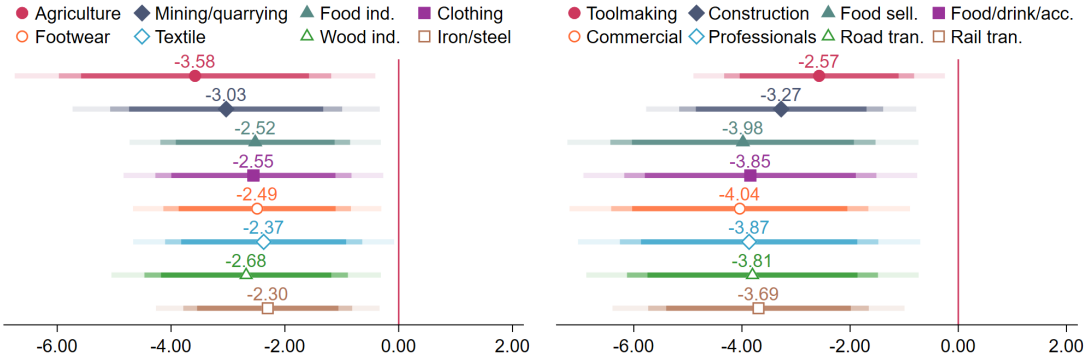
Table A2. The long-run effect of industrial concentration and population—sensitivity to industry controls.

Wage (2020)	(1)	(2)	(3)	(4)	(5)
Herfindahl index (1881, h_c)	-3.111	-2.824	-2.639	-2.688	-2.912
	(1.060)	(0.908)	(0.903)	(0.788)	(0.961)
Employment (1881, l_c)	-0.357	-0.223	-0.309	-0.225	-0.204
	(0.191)	(0.190)	(0.178)	(0.183)	(0.204)
Observations	428	428	428	428	428
F-stat (h_c)	12.43	16.82	16.41	14.03	11.56
F-stat (l_c)	12.98	14.65	13.69	10.33	8.84
Additional controls	Herf. (1d)	Agri.	Service	Manu.	All

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). The dependent variable is the (log) wage in 2020 (labor income estimates for small areas based on the Family Resources Survey). The specifications include the same controls as in Table 3. The two instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2 and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c. Column (1) adds a control for a Herfindahl index predicted as in Section 2.2, but with 1-digit sectors (agriculture, manufacturing, services, transport, public); column (2) adds a control for the agricultural employment share in 1881; column (3) adds a control for the service employment share in 1881; column (4) adds a control for the manufacturing employment share in 1881; and column (5) adds all previous controls. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

(agriculture, manufacturing, services, transport, public); a control for the agricultural employment share in 1881; a control for the service employment share in 1881; a control for the manufacturing share in 1881; and all these controls altogether.

Figure A12. The long-run effect of industrial concentration—dropping industries one by one.



Notes: This Figure shows the estimated effect of 1881 industrial concentration on the 2020 (log) wage in specifications in which we drop each of the 16 major nineteenth-century industries one by one from the actual and predicted Herfindahl indices.

Next, we examine whether our main empirical finding is driven by any one of the *major* nineteenth-century industries. To this end, we re-run the baseline specification of Table 3, column (2), after dropping each of the 16 major industries, one by one. Fig-

ure A12 shows that the estimated effect of industrial concentration on the long-run wage remains negative and stable in all of these specifications. This suggests that none of the major industries are driving our headline empirical result.

Table A3. The long-run effect of industrial concentration and population—alternative instruments (and indices) for industrial concentration.

Wage (2020)	(1)	(2)	(3)	(4)
Herfindahl index (1881, h_c)	-2.583	-3.443	-2.463	-1.551
	(0.788)	(1.572)	(0.894)	(0.791)
Employment (1881, l_c)	-0.311	-0.263	-0.340	-0.406
	(0.182)	(0.164)	(0.176)	(0.212)
Observations	428	428	428	428
F-stat (h_c)	17.56	9.26	14.13	31.23
F-stat (l_c)	14.60	10.94	14.80	12.33
Instrument (χ_c)	Leave-out	Shift-share	Buffer	Linkages

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). The dependent variable is the (log) wage in 2020 (labor income estimates for small areas based on the Family Resources Survey). The specifications include the same controls as in Table 3. The two instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2 and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c. Column (1) relies on an instrument for industrial concentration (χ_c) using a leave-out approach for the shifts (sectoral growth outside of a buffer of 10 kilometers around the city); column (2) uses an instrument for industrial concentration (χ_c) in a purer shift-share design $(\sum_i (s_{ic} g_i)^2 / \sum_i s_{ic}^2)$; column (3) constructs the baseline instrument for industrial concentration using a buffer of 10 kilometers around the city (industrial concentration, h_c, is then also constructed within this buffer); and column (4) constructs a weighted Herfindahl index where cross-sectoral weights are parametrized on the occupational transitions for employed males in the micro-censuses of 1851 and 1861. In column (4), we add the weighted Herfindahl index in 1817 as a control. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

Alternative indices and instruments of industrial concentration In Table A3, we probe the robustness of our results to alternative instruments and indices of industrial concentration: using a leave-out approach for the shifts (sectoral growth outside of a 10-km buffer around the city, column 1); using a purer shift-share design $\sum_j \left(s_{jc} g_j\right)^2 / \sum_j s_{jc}^2$ in column (2); using a buffer of 10 kilometers around the city to define local industrial concentration and local predicted industrial concentration (column 3); and using Herfindahl indices (actual and predicted) with cross-sectoral weights in column (4). More specifically, letting α_{ij} denote the symmetric transition probability between sectors i and j for employed males linked through the micro-censuses in 1851 and in 1861, one can define a Herfindahl index as $h_c = \sum_{i,j} \alpha_{ij} s_{ic} s_{jc}$, both in 1881 and in 1817, or based on predicted shares for the instrument. Intuitively, this measure coincides with our baseline measure when the matrix $\mathbf{A} = \left(\alpha_{ij}\right)_{i,j}$ is the identity matrix. The main take-away messages are that the negative effect of industrial concentration resists these alternative

specifications for the indices/instruments of industrial concentration.

Table A4. The long-run effect of industrial concentration and population—alternative specifications for industrial concentration.

Wage (2020)	(1)	(2)	(3)	(4)
Herfindahl index (1881, h_c)	-0.496	0.132	-0.576	-4.464
	(0.158)	(0.042)	(0.657)	(1.325)
Employment (1881, l_c)	-0.354	-0.393	-0.414	-0.103
	(0.195)	(0.204)	(0.300)	(0.209)
Observations	428	428	428	428
F-stat (h_c)	17.74	21.65	4.74	12.94
F-stat (l_c)	13.57	12.87	6.38	12.76
Specification (h_c, χ_c)	Logarithm	Inverse	Distance	Baseline
Additional controls	–	–	–	Active ind.

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). The dependent variable is the (log) wage in 2020 (labor income estimates for small areas based on the Family Resources Survey). The specifications include the same controls as in Table 3. The two instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2 and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c. Column (1) uses the logarithms of the Herfindahl index (1881, h_c) and its instrument; column (2) transforms the Herfindahl index (1881, h_c) and its instrument through an inverse ($f(x) = 1/(0.05 + x)$); column (3) uses instead an indicator of distance to the average national portfolio ($\sum_j |s_{jc} - s_j|$); and column (4) uses the baseline specification but controls for the number of active industries in 1817 and in 1881. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

Table A4 further probes the robustness of our estimates to the exact specification, e.g., using the logarithms of the Herfindahl index (1881, h_c) or its inverse (as discussed in Combes and Gobillon, 2015). We perform these robustness checks in column (1) and (2): a 50% higher Herfindahl index is associated with a 25% lower wage; and a 1.8 higher value for the index of industrial diversity (a standard deviation in the inverse transformation of the Herfindahl index) is associated with a 26% higher wage. Column (3) considers an alternative measure of local industrial "specialization", measuring the industrial shares against those of the country-wide portfolio: $h_c = \sum_j |s_{jc} - s_j|$ (and its associated instrument based on predicted shares). A 0.1 higher specialization index—about one standard deviation—is associated with a 6% lower wage, but the estimate is noisy due to a weak first stage. Finally, column (4) uses the baseline specification but controls for the number of active industries in 1817 and in 1881.

Spatial spillovers and the effect of specialization in nearby cities While our quantitative model allows for spatial linkages between cities in terms of factor mobility and the trade of varieties produced in each sector, it does not allow for Jacobs externalities to operate across cities. In Table A5, we provide support for this hypothesis by adding

Table A5. The long-run effect of industrial concentration and population—spatial spillovers.

Wage (2020)	(1)	(2)	(3)
Herfindahl index (1881, h_c)	-2.954	-2.902	-2.716
	(1.252)	(1.250)	(1.130)
Herfindahl index in nearby cities (1881, h_c^d)	0.406	0.568	0.173
	(1.098)	(1.227)	(0.942)
Employment (1881, l_c)	-0.309	-0.311	-0.319
	(0.179)	(0.187)	(0.229)
Observations	428	428	428
F-stat (h_c)	31.71	13.03	9.89
F-stat (h_c^d)	48.15	14.38	11.10
F-stat (l_c)	15.79	12.60	9.54
Nearby instrument (χ_c^d)	Buffer-10	Buffer-20	Buffer-40

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). The dependent variable is the (log) wage in 2020 (labor income estimates for small areas based on the Family Resources Survey). The specifications include the same controls as in Table 3. The three instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2, the "shift-share" predictor of industrial concentration using buffers of 10, 20, and 40 kilometers around the city, and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c. The variables h_c^d are industrial concentration indices using buffers of 10, 20, and 40 kilometers around the city. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

a measure of industrial concentration within 10, 20, and 40 kilometers around each city, and instrumenting the latter with a "shift-share" predictor of industrial concentration based on initial employment within the same buffers. We find that Jacobs externalities operate within 10 kilometers around each city—we have 435 cities in our baseline distributed over more than 100,000 square kilometers such that a limited number of those cities fall into 10-kilometer buffers around other cities.

Treatment heterogeneity In Table A6, we document treatment heterogeneity in the incidence of services (in 1881, column 1), manufacturing (in 1881, column 2), transportation (in 1881, column 3), and the success of the industry portfolio during the twentieth century (a shift-share control based on employment shares in 1881 and aggregate employment growth between 1881–1971, column 4). The interacted variables are standardized such that the estimate in front of the interaction can be understood as the extent to which one standard deviation in the interacted variable affects the treatment effect. We mostly find that cities that are service-intensive are slightly less prone to a specialization curse, and conversely, manufacturing intensity increases the treatment effect. This heterogeneity in treatment remains however very limited.

In Table A7, we look at the interacted effect of industrial concentration and size: we

Table A6. The long-run effect of industrial concentration and population—treatment heterogeneity.

Interaction with ...	Serv. (1881)	Manu. (1881)	Tra. (1881)	Manu. (1971)
Wage (2020)	(1)	(2)	(3)	(4)
Herfindahl index (1881, h_c)	-2.805	-3.281	-2.956	-2.615
	(1.027)	(1.270)	(0.925)	(0.856)
Employment (1881, l_c)	-0.338	-0.314	-0.299	-0.263
	(0.192)	(0.243)	(0.176)	(0.169)
Herfindahl (1881) × Int.	-0.171	-1.056	-0.464	-0.325
	(0.348)	(0.819)	(0.294)	(0.474)
Observations	428	428	428	428
F-stat (h_c)	15.96	6.03	20.22	18.36
F-stat (l_c)	13.11	5.03	14.29	17.55
F-stat (Int.)	48.82	5.18	62.56	42.46

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). The dependent variable is the (log) wage in 2020 (labor income estimates for small areas based on the Family Resources Survey). The specifications include the same controls as in Table 3 (in addition to the interacted variable). The three instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2, the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c and its interaction with the different interacted variables. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

find that larger cities, e.g., 50% as large, are able to mitigate the industrial concentration effect—which becomes around 25% smaller, whether we consider the share of unskilled workers in 1971 or the 2020 (log) wage. Again, this heterogeneity in treatment is limited.

Sensitivity analysis We consider a sensitivity analysis around our baseline specification changing: the land fragmentation instrument, computed with a fixed buffer (1 kilometer), dropping bulk density from the algorithm delineating parcels, using SLIC rather than Quickshift to segment space into potential agricultural parcels (see Figure A5); with different inference to account for spatial correlation (clustering at the level of registration districts or poor law unions, at the level of counties, at the level of Travel-to-Work Area, see Figure A13); with different cut-offs to define urban settlements (4000, 5000, 6000 inhabitants around 1820, see Figure A13).

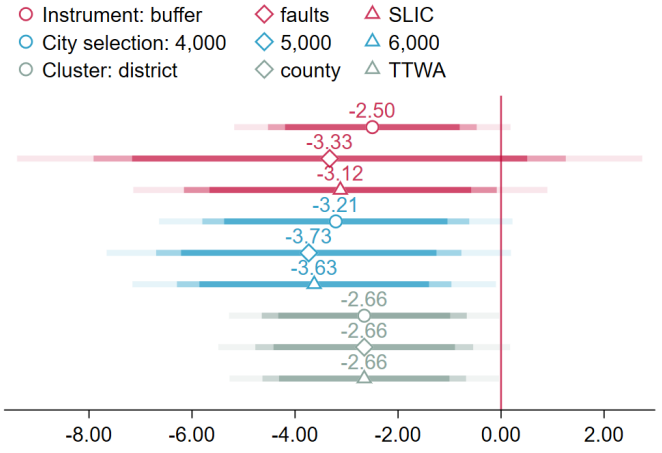
Finally, our baseline specification uses an outcome in 1971, before the deregulation policies implemented by Margaret Thatcher. In Table A8, we conduct a sensitivity analysis with outcomes computed in 1981 and 1991 and we show similar standardized effects over time, with a slight inflection after 1991.

Table A7. The long-run effect of industrial concentration, population, and their interaction.

	Unskilled employment (1971)	Wage (2020)
	(1)	(2)
Herfindahl index (1881, h_c)	0.447	-3.979
	(0.197)	(1.462)
Employment (1881, l_c)	0.060	-0.668
	(0.049)	(0.339)
Herfindahl × Employment	-0.157	2.663
	(0.262)	(2.072)
Observations	428	428
F-stat (h_c)	14.12	14.12
F-stat (l_c)	13.67	13.67
F-stat (Interaction)	7.49	7.49

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). In column (1), the dependent variable is the share of unskilled workers in 1971 (following the definition used in Heblich et al., 2021); in column (2), the dependent variable is a measure of (log) wages in 2020. The main endogenous variables are the Herfindahl index of industrial concentration in 1881 (h_c), (log) employment in 1881 (l_c), and their interaction. The set of baseline controls are the same as in Table 2. The three instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2, the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c, and their interaction. We also control for a shift-share control based on employment shares in 1881 and aggregate employment growth between 1881–1971, to clean for sector-specific trends. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

Figure A13. The long-run effect of industrial concentration—instrument, city selection, and clustering.



Notes: This Figure shows the estimated effect of 1881 industrial concentration on the 2020 (log) wage in the following robustness checks: we consider a sensitivity analysis around our baseline specification changing the land fragmentation instrument, computed with a fixed buffer (1 kilometer, "buffer"), dropping bulk density from the algorithm delineating parcels ("faults"), using SLIC rather than Quickshift to segment space into potential agricultural parcels ("SLIC"); with different cut-offs to define urban settlements (4000, 5000, 6000 inhabitants around 1820); with different inference to account for spatial correlation (clustering at the level of registration districts or poor law unions, at the level of counties, at the level of Travel-to-Work Area.

Table A8. The long-run effect of population and industrial concentration—sensitivity to other years.

	1971	1981	1991
Unskilled employment	(1)	(2)	(3)
Herfindahl index (1881, h_c)	0.369	0.570	1.025
	(0.148)	(0.215)	(0.315)
	[9.47]	[10.18]	[12.37]
Employment (1881, l_c)	0.038	0.065	0.099
	(0.030)	(0.042)	(0.066)
	[0.98]	[1.16]	[1.20]
Observations	428	428	428
F-stat (h_c)	15.30	15.30	15.30
F-stat (l_c)	14.94	14.94	14.94

Notes: A unit of observation is a cluster of settlements around 1790-1820—identified following our clustering procedure described in Section 2.2. Standard errors are reported between parentheses and are clustered at the level of the closest city as of 2015 (there are 55 cities with a formal city status in England). Standardized effects are reported in square brackets. The dependent variables are the shares of unskilled workers in 1971, 1981 and 1991 (following the definition used in Heblich et al., 2021). The specifications include the same controls as in column (2) of Table 3. The two instruments are the "shift-share" predictor of industrial concentration (χ_c) defined in Section 2.2 and the land ownership fragmentation in the immediate fringe of urban settlements, ζ_c. F-statistics are derived using the weak instrument F-test proposed in Sanderson and Windmeijer (2016).

B Theory appendix

B.1 A model with infinitely-lived workers

In this appendix, we present a version of the model in which workers are infinitely lived and make forward-looking decisions about their locations, subject to migration costs across periods.

We assume that a worker m who resided in city c_{t-1} at the end of period $t-1$ chooses her city c_t for the next period to maximize her total discounted stream of future utilities. This implies that her value function is given by,

$$v^m_{c_{t-1},t} = \max_{c_t} \left\{ \max_{i_t} \ln U^m_{i_t,c_t,t} - \ln m_{c_{t-1},c_t} + \beta \mathbf{E}_t \left[v^m_{c_t,t+1} \right] \right\},$$

where $U^m_{i_t,c_t,t}$ is the amenity-adjusted real income in period t, given by Equation (2), and m_{c_{t-1},c_t} is the cost of moving from city c_{t-1} to city c_t. The last term, $\mathbf{E}_t \left[v^m_{c_t,t+1} \right]$, is the worker's continuation value, which is discounted by a factor $\beta \in [0,1)$. The continuation value includes an expectation because the future taste shocks for cities, given by Equation (5), are not yet realized in period t. The model presented in Section 4 can be viewed as a special case of this more general model, with $\beta = 0$ and $m_{cd} = 0$ for all cities c and d.

Given the Fréchet distribution of idiosyncratic city tastes, it can be shown that the expected value of a city c at time t, $V_{ct} = e^{\mathbf{E}_t \left[v^m_{c,t+1} \right]}$, evolves according to the equation,

$$V_{ct} = \sum_{d=1}^{C} \left(\bar{a}_d \frac{w_{d,t+1} + R_{d,t+1}}{P_{d,t+1}} L^{-\lambda}_{d,t+1} m^{-1}_{cd} V^{\beta}_{d,t+1} \right)^{1/\eta}, \tag{B.1}$$

while the fraction of workers who choose to move from city c to city d between periods $t-1$ and t equals,

$$\mu_{c,d,t-1} = \frac{\left(\bar{a}_d \frac{w_{dt}+R_{dt}}{P_{dt}} L^{-\lambda}_{dt} m^{-1}_{cd} V^{\beta}_{dt} \right)^{1/\eta}}{\sum_{d'=1}^{C} \left(\bar{a}_{d'} \frac{w_{d',t}+R_{d',t}}{P_{d',t}} L^{-\lambda}_{d',t} m^{-1}_{c,d'} V^{\beta}_{d',t} \right)^{1/\eta}}.$$

As a consequence, the population of a city d evolves according to the equation,

$$L_{dt} = \sum_{c=1}^{C} \frac{\left(\bar{a}_d \frac{w_{dt}+R_{dt}}{P_{dt}} L^{-\lambda}_{dt} m^{-1}_{cd} V^{\beta}_{dt} \right)^{1/\eta}}{\sum_{d'=1}^{C} \left(\bar{a}_{d'} \frac{w_{d',t}+R_{d',t}}{P_{d',t}} L^{-\lambda}_{d',t} m^{-1}_{c,d'} V^{\beta}_{d',t} \right)^{1/\eta}} L_{c,t-1}. \tag{B.2}$$

Equations (B.1) and (B.2), together with Equations (9), (11) and (12), determine the equi-

librium of the model with infinitely-lived workers. Note that Equation (12), which we use to estimate the strength of dynamic externalities, is unchanged in this more general model.

B.2 Derivation of Equation (13)

Given that workers are freely mobile across industries, nominal wages equalize across industries within each city:

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

Plugging this result into Equation (8), 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 we obtain,

$$w_{ct} + R_{ct} = \frac{1}{\gamma} w_{ct}. \tag{B.3}$$

Also, from Equation (7), we get

$$r_{ct} = \left(\frac{1-\gamma}{\gamma}\right)^{1/\zeta_{ct}} w_{ct}^{1/\zeta_{ct}} L_{ct}^{1/\zeta_{ct}}. \tag{B.4}$$

Given perfect competition, the price of each variety c in industry i at the factory gate is equal to its marginal cost of production in equilibrium,

$$p_{ict} = \mathcal{T}_{ict}^{-1} w_{ct}^{\gamma} r_{ct}^{1-\gamma} = \left(\frac{1-\gamma}{\gamma}\right)^{\frac{1-\gamma}{\zeta_{ct}}} \mathcal{T}_{ict}^{-1} L_{ct}^{\frac{1-\gamma}{\zeta_{ct}}} w_{ct}^{\gamma+\frac{1-\gamma}{\zeta_{ct}}}, \tag{B.5}$$

where we used Equation (B.4). As a result, we can write the price index of industry i, Equation (10), as,

$$P_{idt} = \left[\sum_{c=1}^{C} \left(\frac{1-\gamma}{\gamma}\right)^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} \mathcal{T}_{ict}^{\epsilon-1} L_{ct}^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} w_{ct}^{-\left(\gamma+\frac{1-\gamma}{\zeta_{ct}}\right)(\epsilon-1)} \tau_{icdt}^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}, \tag{B.6}$$

and market clearing condition (9) as,

$$w_{ct}L_{ict} = \left(\frac{1-\gamma}{\gamma}\right)^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} \tilde{T}_{ict}^{\epsilon-1} L_{ct}^{-\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} w_{ct}^{-\left(\gamma+\frac{1-\gamma}{\zeta_{ct}}\right)(\epsilon-1)} \sum_{d=1}^{C} P_{dt}^{\sigma-1} P_{idt}^{\epsilon-\sigma} w_{dt} L_{dt} \tau_{icdt}^{1-\epsilon}, \tag{B.7}$$

where we also use Equation (B.3).

By the Fréchet distribution of idiosyncratic city tastes, the probability that a worker

chooses city c equals,

$$Pr\left[U_{ct}^{m} \geq U_{dt}^{m},\ \forall d\right] = \frac{\left(\bar{a}_c \frac{w_{ct}+R_{ct}}{P_{ct}} L_{ct}^{-\lambda}\right)^{1/\eta}}{\sum_{d=1}^{C}\left(\bar{a}_d \frac{w_{dt}+R_{dt}}{P_{dt}} L_{dt}^{-\lambda}\right)^{1/\eta}} = \frac{\left(\frac{1}{\gamma}\bar{a}_c \frac{w_{ct}}{P_{ct}} L_{ct}^{-\lambda}\right)^{1/\eta}}{\sum_{d=1}^{C}\left(\frac{1}{\gamma}\bar{a}_d \frac{w_{dt}}{P_{dt}} L_{dt}^{-\lambda}\right)^{1/\eta}}.$$

In equilibrium, the fraction of workers choosing to live in c becomes equal to this probability:

$$\frac{L_{ct}}{\bar{L}} = \frac{\left(\frac{1}{\gamma}\bar{a}_c \frac{w_{ct}}{P_{ct}} L_{ct}^{-\lambda}\right)^{1/\eta}}{\sum_{d=1}^{C}\left(\frac{1}{\gamma}\bar{a}_d \frac{w_{dt}}{P_{dt}} L_{dt}^{-\lambda}\right)^{1/\eta}},$$

from which:

$$P_{ct} = \left(\gamma \overline{U}_t\right)^{-1} \bar{a}_c w_{ct} L_{ct}^{-(\lambda+\eta)}, \tag{B.8}$$

where:

$$\overline{U}_t = \left[\frac{\sum_{d=1}^{C}\left(\frac{1}{\gamma}\bar{a}_d \frac{w_{dt}}{P_{dt}} L_{dt}^{-\lambda}\right)^{1/\eta}}{\bar{L}}\right]^{\eta}.$$

Plugging this result into Equations (11) and (B.7) and rearranging Equation (B.6), we obtain the following system of equations:

$$P_{ict}^{1-\epsilon} = \sum_{d=1}^{C}\left(\frac{1-\gamma}{\gamma}\right)^{-\frac{1-\gamma}{\zeta_{dt}}(\epsilon-1)} \overline{T}_{idt}^{\epsilon-1} L_{dt}^{-\left(\frac{1-\gamma}{\zeta_{dt}}-\alpha\right)(\epsilon-1)} w_{dt}^{-\left(\gamma+\frac{1-\gamma}{\zeta_{dt}}\right)(\epsilon-1)} \tau_{idct}^{1-\epsilon} \tag{B.9}$$

$$\bar{a}_c^{1-\sigma} w_{ct}^{1-\sigma} L_{ct}^{(\lambda+\eta)(\sigma-1)} = \left(\gamma \overline{U}_t\right)^{1-\sigma} \sum_{i=1}^{I} P_{ict}^{1-\sigma} \tag{B.10}$$

$$\begin{aligned}&\left(\frac{1-\gamma}{\gamma}\right)^{\frac{1-\gamma}{\zeta_{ct}}(\epsilon-1)} w_{ct}^{1+\left(\gamma+\frac{1-\gamma}{\zeta_{ct}}\right)(\epsilon-1)} L_{ct}^{1+\left(\frac{1-\gamma}{\zeta_{ct}}-\alpha\right)(\epsilon-1)} = \\ &\left(\gamma \overline{U}_t\right)^{1-\sigma} \sum_{i=1}^{I}\sum_{d=1}^{C} \overline{T}_{ict}^{\epsilon-1} P_{idt}^{\epsilon-\sigma} \bar{a}_d^{\sigma-1} w_{dt}^{\sigma} L_{dt}^{1-(\lambda+\eta)(\sigma-1)} \tau_{icdt}^{1-\epsilon}\end{aligned} \tag{B.11}$$

where:

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

is the part of Total Factor Productivity that is exogenous in period t. Applying the change in variables in Equation (15) and recalling that $w_{ict} = w_{ct}$, Equation (13) immediately follows from Equations (B.9), (B.10) and (B.11).

B.3 Proof of Theorem 1

Rearranging the period-1 version of Equations (B.9) to (B.11) yields:

$$P_{ic1}^{1-\epsilon} = \sum_{d=1}^{C} \left(\frac{1-\gamma}{\gamma}\right)^{-\frac{1-\gamma}{\xi_{d1}}(\epsilon-1)} \overline{T}_{id1}^{\epsilon-1} L_{d1}^{-\left(\frac{1-\gamma}{\xi_{d1}}-\alpha\right)(\epsilon-1)} w_{d1}^{-\left(\gamma+\frac{1-\gamma}{\xi_{d1}}\right)(\epsilon-1)} \tau_{idc1}^{1-\epsilon} \tag{B.12}$$

$$\left(\frac{\bar{a}_c}{\overline{U}_1}\right)^{1-\sigma} w_{c1}^{1-\sigma} L_{c1}^{(\lambda+\eta)(\sigma-1)} = \gamma^{1-\sigma} \sum_{i=1}^{I} P_{ic1}^{1-\sigma} \tag{B.13}$$

$$\begin{aligned} &\left(\frac{1-\gamma}{\gamma}\right)^{\frac{1-\gamma}{\xi_{c1}}(\epsilon-1)} w_{c1}^{1+\left(\gamma+\frac{1-\gamma}{\xi_{c1}}\right)(\epsilon-1)} L_{c1}^{\left(\frac{1-\gamma}{\xi_{c1}}-\alpha\right)(\epsilon-1)} \overline{T}_{ic1}^{1-\epsilon} L_{ic1} = \\ &\gamma^{1-\sigma} \sum_{d=1}^{C} P_{id1}^{\epsilon-\sigma} \left(\frac{\bar{a}_d}{\overline{U}_1}\right)^{\sigma-1} w_{d1}^{\sigma} L_{d1}^{1-(\lambda+\eta)(\sigma-1)} \tau_{icd1}^{1-\epsilon} \end{aligned} \tag{B.14}$$

from which we obtain the following system of $3IC$ equations:

$$\begin{aligned} \widehat{x}_{ic}^1 &= \sum_{j=1}^{I} \sum_{d=1}^{C} \left(\widehat{x}_{jd}^3\right)^{-1} \widehat{K}_{icjd}^1 \\ \widehat{x}_{ic}^2 &= \sum_{j=1}^{I} \sum_{d=1}^{C} \left(\widehat{x}_{jd}^1\right)^{\frac{\sigma-1}{\epsilon-1}} \widehat{K}_{icjd}^2 \\ \widehat{x}_{ic}^3 &= \sum_{j=1}^{I} \sum_{d=1}^{C} \left(\widehat{x}_{jd}^1\right)^{-\frac{\epsilon-\sigma}{\epsilon-1}} \left(\hat{x}_{jd}^2\right)^{-1} \widehat{K}_{icjd}^3 \end{aligned}$$

where $\widehat{x}_{ic}^1 = P_{ic1}^{1-\epsilon}$, $\widehat{x}_{ic}^2 = \left(\frac{\bar{a}_c}{\overline{U}_1}\right)^{1-\sigma}$, and $\widehat{x}_{ic}^3 = \overline{T}_{ic1}^{1-\epsilon}$.

The solution to this system exists and is unique if the largest eigenvalue of matrix $\mathbf{A}$ is less or equal to one in absolute value (Allen et al., 2020), where $\mathbf{A}$ is:

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ \left|\frac{\sigma-1}{\epsilon-1}\right| & 0 & 0 \\ \left|\frac{\epsilon-\sigma}{\epsilon-1}\right| & 1 & 0 \end{bmatrix}.$$

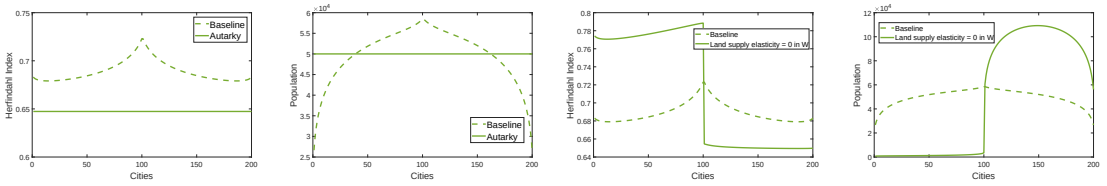
One can easily verify that, under the assumption $\epsilon > \sigma > 1$, the largest eigenvalue of $\mathbf{A}$ is equal to 1.

B.4 Complements to the linear economy

This section provides complements to our illustrative exercise of Section 4.3. More specifically, the top panels of Figure B1 presents the spatial distribution of industrial concentration and population in a world in which cities are in autarky. In particular,

panel (a) sheds some light on the absolute and relative impact of trade on industrial concentration. The bottom panels of Figure B1 presents the same outcomes in a world in which the land supply elasticity of Western cities is decreased to zero.

Figure B1. The linear economy under autarky in period 0.



(a) Concentration (b) Employment (c) Concentration (d) Employment

Notes: The values for the structural parameters are set as follows: $\alpha = 0.06$; $\gamma = 0.65$; $\epsilon = 5$; $\phi = 0.25$; $\sigma = 4$; $\bar{L} = 10,000,000$; $\zeta_{ct} = 2$.

B.5 Complements to the model parameterization and estimation

Trading costs over time To compute trading costs over time, we follow Alvarez-Palau et al. (2013) and combine the main transport modes at the time. Specifically, we combine information on: (i) the road network that was represented by turnpike roads; (ii) waterways consisting of rivers, canals and coastal routes; and (iii) the railway network and railway stations observed in 1880 (Cobb, 2003) in a multi-modal transport network and calculate the least cost path between all city-centroids of our sample. We connect the different transport modes and city-centroids with direct linear routes. Straight-line distances between points are calculated using the Manhattan distance metric, which reflects the reality that roads typically do not follow direct, 'as-the-crow-flies' paths. We consider the following straight-line connections:

- Every city-centroid located within a 15km radius is interconnected through direct linear routes. These direct links bypass the need for detours along the actual road network for short distances.
- Every city-centroid located within 10km from the nearest river is interconnected through linear routes.
- Every city-centroid located within 10km from the nearest turnpike road to the east, south, west and north is interconnected through linear routes.
- Every city-centroid located within 10km from the nearest port is interconnected through linear routes.

- Every city-centroid located within 10km from the nearest 1880 railway station is interconnected through linear routes.
- Every seaport located within 10km from the nearest river is connected through linear routes.
- Every seaport located within 10km from the nearest turnpike road is interconnected through linear routes.
- Every seaport located within 10km from the nearest railway station is interconnected through linear routes.
- Every railway station located within 10km from the nearest river is interconnected through linear routes.
- Every railway station located within 10km from the nearest turnpike road is interconnected through linear routes.

All other distances are calculated along the specific transport network path. To translate distances into costs, we consider the following transport costs in ascending order: (i) cost of shipping goods on coastal waterways: 0.003 per mile; (ii) cost of shipping goods on rail: 0.027 per mile; (iii) cost of shipping goods on rivers or canals: 0.036 per mile; (iv) cost of shipping goods on roads: 0.213 per mile. In addition, we assume fixed transshipment costs between road and railway, waterways, or coastal routes of 1.895. Lastly, we assume that transport between the city-centroid and the nearest sea or river port, railway station or turnpike road is not distance dependent, i.e., we assume that actual production locations within cities are directly connected to relevant transport modes. This is on line with reality where we see that canals meander through Manchester, connecting different production locations, and we also see railways branching out to industrial districts.

Numerical procedure to invert the model The challenge involved in solving Equations (B.12) to (B.14) is that the system is homogenous of degree one in unobserved fundamentals. While iteration can be proven to work in similar systems if the degree of homogeneity is below one (Allen et al., 2020), this is no longer necessarily true if it is equal to one.

To circumvent this issue, we borrow from Desmet et al. (2018) and approximate the system in the following way:

$$P_{ic1}^{1-\epsilon} = \sum_{d=1}^{C} \left(\frac{1-\gamma}{\gamma} \right)^{-\frac{1-\gamma}{\xi_{d1}}(\epsilon-1)} \overline{T}_{id1}^{(\epsilon-1)(1-\delta)} L_{d1}^{-\left(\frac{1-\gamma}{\xi_{d1}}-\alpha\right)(\epsilon-1)} w_{d1}^{-\left(\gamma+\frac{1-\gamma}{\xi_{d1}}\right)(\epsilon-1)} \tau_{idc1}^{1-\epsilon} \tag{B.15}$$

$$\left(\frac{\bar{a}_c}{\overline{U}_1}\right)^{1-\sigma} w_{c1}^{1-\sigma} L_{c1}^{(\lambda+\eta)(\sigma-1)} = \gamma^{1-\sigma} \sum_{i=1}^{I} P_{ic1}^{1-\sigma} \tag{B.16}$$

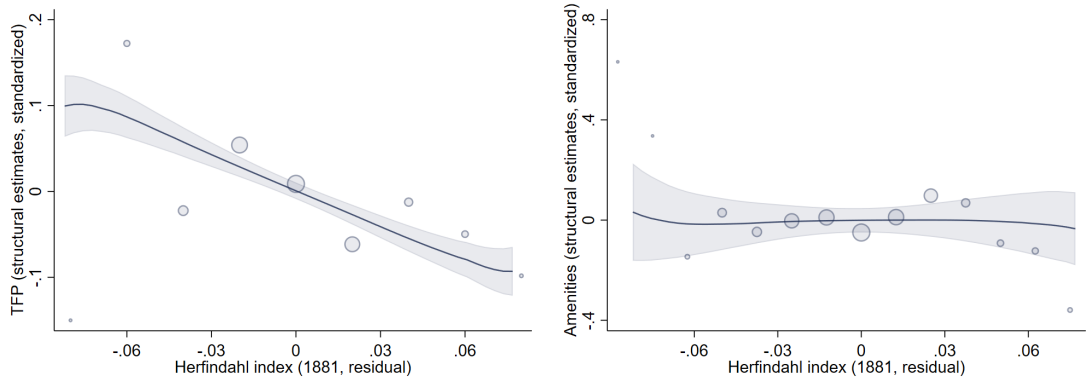
$$\left(\frac{1-\gamma}{\gamma}\right)^{\frac{1-\gamma}{\xi_{c1}}(\epsilon-1)} w_{c1}^{1+\left(\gamma+\frac{1-\gamma}{\xi_{c1}}\right)(\epsilon-1)} L_{c1}^{\left(\frac{1-\gamma}{\xi_{c1}}-\alpha\right)(\epsilon-1)} \overline{T}_{ic1}^{1-\epsilon} L_{ic1} = \gamma^{1-\sigma} \sum_{d=1}^{C} P_{id1}^{\epsilon-\sigma-\delta} \left(\frac{\bar{a}_d}{\overline{U}_1}\right)^{\sigma-1} w_{d1}^{\sigma} L_{d1}^{1-(\lambda+\eta)(\sigma-1)} \tau_{icd1}^{1-\epsilon} \tag{B.17}$$

where the only difference relative to Equations (B.12) to (B.14) is that the parameter δ appears in Equations (B.15) and (B.17). As $\delta \to 0$, this system becomes identical to (B.12) to (B.14).

Intuitively, for any small but positive δ, the system (B.15) and (B.17) is homogenous to a degree below one. Therefore, it can be solved by iteration. In practice, we start with a $\delta = 0.128$ and gradually lower δ to get closer and closer to the original system. We stop when δ is low enough that the solution of the model under the recovered amenities $\{\bar{a}_c\}_c$ and city-industry productivities $\left\{\overline{T}_{ic1}\right\}_{i,c}$, computed as in Section 4.2, is sufficiently close to the data.

The structural estimates for dynamic spillovers Section 5.1 describes (i) how we recover the distribution of city amenities $\bar{a}_c$ and city-industry productivities $\overline{T}_{ic}$ around 2020 from wages w_c and employment by city-industry L_{ic} and (ii) the empirical estimation of Equation (17) (see, e.g., Table 5).

Figure B2. The structural estimates for dynamic spillovers—an illustration.



(a) Productivity (b) Amenity

Notes: Panel (a) displays the relationship between industrial concentration in 1881 and standardized productivities $\overline{T}_{ic}$ across our 435 cities and 88 industries; both measures are cleaned for all controls used in Table 2 and industry fixed effects. Panel (b) displays the relationship between industrial concentration in 1881 and standardized amenities $\bar{a}_c$; the x-axis and y-axis variables are then cleaned for all controls used in Table 2.

In this Appendix, we provide an illustration of the effect of dynamic spillovers in

the spirit of Figure 8: we residualize city-industry (standardized) productivities $\overline{T}_{ic}$ and (standardized) amenities $\bar{a}_c$, and we show their dependence on industrial concentration in 1881. One can see that industrial concentration is negatively associated with future city productivity, but not with the estimated city amenities—consistent with the assumptions underlying our counterfactual experiments.