California Center for Population Research
University of California - Los Angeles

Agglomeration: A Long-Run Panel Data Approach

W. Walker Hanlon and Antonio Miscio

PWP-CCPR-2016-018

February 22, 2016

California Center for Population Research
On-Line Working Paper Series

Agglomeration: A Long-Run Panel Data Approach*

W. Walker Hanlon
UCLA and NBER

Antonio Miscio
Columbia University

November 20 , 2015

Abstract

This paper studies the sources of agglomeration economies in cities. We begin by incorporating within and cross-industry spillovers into a dynamic spatial equilibrium model in order to obtain a panel data estimating equation. This gives us a framework for measuring a rich set of agglomeration forces while controlling for a variety of potentially confounding effects. We apply this estimation strategy to detailed new data describing the industry composition of 31 English cities from 1851-1911. Our results show that industries grew more rapidly in cities where they had more local suppliers or other occupationally-similar industries. We find no evidence of dynamic within-industry effects, i.e., industries generally did not grow more rapidly in cities in which they were already large. Once we control for these agglomeration forces, we find evidence of strong dynamic congestion forces related to city size. We also show how to construct estimates of the combined strength of the many agglomeration forces in our model. These results suggest a lower bound estimate of the strength of agglomeration forces equivalent to a city-size divergence rate of 1.0-2.3% per decade. JEL Codes: R1, N93, O3

*We thank David Albouy, Pierre-Philippe Combes, Dora Costa, Don Davis, Jonathan Dingel, Gilles Duranton, Glenn Ellison, Ben Faber, Pablo Fajgelbaum, Edward Glaeser, Laurent Gobillon, Richard Hornbeck, Matt Kahn, Petra Moser, Alex Whalley and seminar participants at Columbia, UCLA, Harvard, UC Merced, UC San Diego, the NBER Innovation group, the NBER Urban Economics group, the CURE conference at Brown University, and the Urban Economics Association Annual Conference, for helpful comments and suggestions. Reed Douglas provided excellent research assistance. Funding for this project was provided by a grant from UCLA's Ziman Center for Real Estate.

1 Introduction

What are the key factors driving city growth over the long term? One of the leading answers to this question, dating back to Marshall (1890), is that firms may benefit from proximity to one another through agglomeration economies. While compelling, this explanation raises further questions about the nature of these agglomeration economies. Do firms primarily benefit from proximity to other firms in the same industry, or, as suggested by Jacobs (1969), is proximity to other related industries more important? How do these forces vary across industries? What role does city size play in industry growth? How can we separate all of these features from the fixed locational advantages of cities? These are important questions for our understanding of cities. Their answers also have implications for the design of place-based policies, which can top $80 billion per year in the U.S. and are also widely used in other countries.[1]

Not surprisingly, there is a large body of existing research exploring the nature of agglomeration economies. This study builds on two important strands of this literature.[2] One approach uses long-differences in the growth of city-industries over time and relates them to rough measures of initial conditions in a city, such as an industry's share of city employment or the Herfindahl index over major city-industries (Glaeser *et al.* (1992), Henderson *et al.* (1995)). The main concern with this line of research is that it ignores much of the richness and heterogeneity that are likely to characterize agglomeration economies. A more recent approach allows for a richer set of inter-industry relationships using connection matrices based on input-output flows, labor force similarity, or technology spillovers. These connections are then compared to a cross-section of industry locations (Rosenthal & Strange (2001), Ellison *et al.*

[1] *The New York Times* has constructed a database of incentives awarded by cities, counties and states to attract companies to locate in their area. The database is available at http://www.nytimes.com/interactive/2012/12/01/us/government-incentives.html.

[2] There are several other strands of the agglomeration literature which are less directly related to this paper. One strand focuses on addressing identification issues by comparing outcomes in similar locations, where some locations receive a plausibly exogenous shock to the level of local economic activity (Greenstone *et al.* (2010) and Kline & Moretti (2013)). This approach has the advantage of more cleanly identifying the causal impact of changes in local economic activity, but it may also be less generalizable and more difficult to apply to policy analysis. Thus, we view this line of work as complementary to our approach. Other alternative approaches use individual-level wage data (Glaeser & Mar (2001), Combes *et al.* (2008), Combes *et al.* (2011)) or firm-level data (Dumais *et al.* (2002), Rosenthal & Strange (2003), Combes *et al.* (2012)) to investigate the effects of city size. See Rosenthal & Strange (2004) and Combes & Gobillon (2015) for reviews of this literature.

(2010), Faggio *et al.* (2013)).[3] A limitation of this type of static exercise is that it is more difficult to control for locational fundamentals in cross-sectional regressions.

Our approach builds on these previous studies, but also seeks to address some of the remaining issues facing the literature. Specifically, this study contributes to the existing literature in five ways. First, while this is primarily an empirical paper, we begin by introducing a new dynamic spatial equilibrium model of city-industry growth. This model incorporates a rich set of within- and cross-industry spillover effects, which allows us to ground our study of these agglomeration forces in a theoretically-consistent framework.[4]

Second, motivated by the theory, we introduce a panel-data econometric approach for estimating the magnitude of agglomeration forces.[5] The key feature of our approach is that we are able to estimate the importance of dynamic agglomeration forces related to industry scale, cross-industry connections, and city-size in a unified framework, while dealing with fixed locational fundamentals and time-varying industry-specific shocks. Previous research has examined these elements separately, but we are not aware of existing work that studies all of these effects in a unified way. In addition, the use of panel data offers some well-known advantages relative to the cross-sectional or long-difference methods used in most existing work. However, applying this approach to study agglomeration economies requires overcoming challenges related to identification and correlated errors. Our study makes progress in this direction, allowing us to address some of the identification concerns present in previous work. The approach that we develop can potentially be applied in a wide range of settings in which consistent panels of city-industry employment data can be constructed.

Third, to implement our approach, we construct a rich dataset describing the evolution of city-industry employment over six decades.[6] These new data, which we

[3] These studies are part of a broader literature looking at the impact of inter-industry connections, particularly through input-output linkages, that includes work by Amiti & Cameron (2007) and Lopez & Sudekum (2009).

[4] In a recent handbook chapter, Combes & Gobillon (2015) highlight the need to ground empirical studies of agglomeration economies in theory.

[5] Our panel data approach builds on previous work by Henderson (1997) and Dumais *et al.* (1997). See also Combes (2000) and Dekle (2002). A panel data approach is also used in a recent working paper by Lee (2015) which uses data on U.S. manufacturing industries from 1880-1990 to study static agglomeration forces.

[6] The availability of detailed long-run city-industry data has been a major impediment to previous work on agglomeration economies. The database constructed in this study helps address this defi-

digitized from original sources, cover 31 of the largest English cities (based on 1851 population) for the period 1851-1911. This empirical setting offers several important advantages. One advantage is the very limited level of government regulation and interference in the British economy during this period due to the strong free-market ideology that dominated British policymaking and the small size of the central government.[7] A second important advantage is that we are able to study agglomeration using consistent data over many decades. Studying agglomeration over a long time period is desirable because the time needed to build new housing, factories, and infrastructure means that it may take years for cities to respond to changes in local productivity levels. Our data are also quite detailed; they come from a full census and cover nearly the entire private sector economy, including manufacturing, transportation, retail, and services. A third advantage is that we are able to study a long-established urban system. This contrasts with the U.S., where the open western frontier meant that the U.S. city system was in transition until the middle of the 20th century.[8] Our setting was also characterized by a relatively open economy with high levels of migration into and between cities.[9]

Fourth, we provide new results on the strength of different types of agglomeration and congestion forces for one empirical setting. We find that (1) cross-industry effects were important, and occurred largely through the presence of local suppliers and occupationally similar labor pools, (2) the net effect of within-industry agglomeration forces was generally negative, and (3) city size had a clear negative relationship to city growth. The presence of local buyers appears to have had little positive influence on city-industry growth. We provide a variety of tests examining the robustness of

---

ciency. Recently, other databases of this type have been developed using data from the U.S. County Business Patterns survey by Duranton *et al.* (2014) and from the U.S. Census of Manufacturers by Lee (2015) and others.

[7] This contrasts with modern settings, where the list of confounding factors includes place-based government policies, local land-use regulations such as zoning, environmental policies that vary across locations, local tax incentives, variation in the local burden of national taxation, as well as many other types of regulation. These factors can also affect city growth, making it more difficult to identify and quantify the role of agglomeration forces. To cite some examples, Kline & Moretti (2013) describe the impact of place-based government policies in the U.S. The role of local land use regulations is highlighted by Gyourko *et al.* (2008). Local environmental policies are studied by Henderson (1996) and Chay & Greenstone (2005), among others. Greenstone & Moretti (2003) describe the impact of local tax incentives, while Albouy (2009) describes how federal tax incentives distort urban growth.

[8] See Desmet & Rappaport (Forthcoming). In contrast, Dittmar (2011) finds that Zipf's Law emerged in European cities between 1500 and 1800, well before the beginning of our study period.

[9] See, e.g., Baines (1994) and Long & Ferrie (2004).

these results. For example, we show that our main results are robust to dropping particular cities or particular industries. They are also robust to using an alternative set of matrices measuring cross-industry connections, alternative functional forms for modeling spillovers, or alternative industry definitions. We also show that incorporating cross-city effects, such as market potential or cross-city industry spillovers, has little impact on our results.

Fifth, we introduce an approach for measuring the combined strength of the many cross-industry agglomeration forces represented in our model. This is valuable because it provides a convenient way to assess the aggregate strength of these effects and may be useful for studying how these effects vary in different circumstances. Our results suggest that a lower-bound estimate of the agglomeration forces captured by our empirical model are equivalent to a decadal city-size divergence rate of 1.0-2.3%. To our knowledge this is the first paper to show how the combined strength of these many cross-industry connections can be measured.

The next section presents our theoretical framework, while Section 3 describes the data. The empirical approach is discussed in Section 4. Section 5 presents the main results, while Section 6 examines the impact of city size and shows how this can be used to calculate the aggregate strength of the agglomeration forces in our model. Section 7 concludes.

2 Theory

While this paper is primarily empirical, a theoretical model is useful in disciplining the empirical specification. Grounding our analysis in theory can also help us interpret the results while being transparent about potential concerns.

Our theory focuses on dynamic agglomeration, i.e., localized spillovers that affect technology and thereby influence industry growth rates. In this respect it is related to the endogenous growth literature (e.g., Romer (1986)) and in particular, to the work of Lucas (1988), who emphasized the important role that localized learning in cities was likely to play in economic growth. This is not the only potential agglomeration force that may lie behind our results; alternative models may yield an estimation equation that matches the one we apply. However, because we are interested in dynamic agglomeration, focusing on technology growth is a natural starting point.[10]

The model is dynamic in discrete time. Technology advances over time as a result of two forces. First, firms undertake R&D in order to improve their productivity. Second, some of the new innovations produced by R&D undertaken by one firm spill over to affect other local firms. These spillovers can occur both within and across industries and the extent of the spillovers depends on a matrix of parameters reflecting the strength of within and inter-industry connections. These spillovers are external to firms, so they will not influence the static allocation of economic activity.

At the end of each period, technology diffuses across firms in the same city and industry. This approach, which follows Desmet & Rossi-Hansberg (2014), substantially simplifies the dynamic elements of the model because firm R&D decisions will only affect firm profits in the current period. By simplifying the dynamics in this way, we are able to build a tractable model with a rich set of inter-industry connections.

As is standard in urban theories, we assume that goods are freely traded across locations and workers are free to move between cities. To keep things simple, our baseline model omits some additional features, such as savings and capital investment, or intermediate inputs, that one might want to consider.[11]

We begin by solving the allocation of employment across space in a particular period. We then consider how the allocation in one period affects the evolution of technology and, thus, the allocation of employment in the next period, through knowledge spillovers. Most of the interesting features of the model are on the producer's side, but we begin with a very brief introduction of the consumers.

[10] An alternative approach is to study static agglomeration, i.e., how the level of employment in one industry affects the level of employment in another, or alternatively, how growth in one industry affects growth in another. Some discussion of static vs. dynamic agglomeration forces is provided in Combes & Gobillon (2015). Lee (2015) provides a recent example of a study focusing on static agglomeration forces. He finds that static localized inter-industry spillovers were small and declining in the U.S. across the 20th century. This suggests that static agglomeration forces are unlikely to be behind the growth of cities during this period.

[11] In the Appendix, we explore the impact of adding capital or intermediate goods. In general this does not change the basic estimating equation that we obtain as long as we maintain the assumption of free mobility across locations, though it can change the interpretation of the parameter estimates.

2.1 Consumption

The model is populated by two types of agents, workers and landlords. There is a continuum of workers in the model, each endowed with one unit of labor. Workers have the option of paying a fixed cost, in terms of labor, in order to become entrepreneurs and open up their own firm. The utility function for both workers and landlords is, $U = \sum_{t=0}^{\infty} u_t e^{-\rho t}$ where u_t is utility in period t. There is no saving, so utility is maximized period-by-period.[12] Utility in any period depends on consumption of real estate h_{ct} and a composite of goods G_{ct} according to a Cobb-Douglas utility function:

$$u_t = h_{ct}^{v} G_{ct}^{1-v} \tag{1}$$

where $v \in (0, 1)$. There are i types of goods available, each produced by a separate industry, and consumers have CES preferences over these goods, so,

$$G_{ct} = \left(\sum_i \gamma_i x_{ict}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

where x_{ict} is consumption of type i goods by a consumer in city c, σ is the elasticity of substitution across goods and $\gamma_i > 0$ is a demand shifter for industry i. The corresponding price index, P_t takes the standard form, with the price of each type of good denoted by p_{it}. Note that, with free trade, goods prices are not indexed by c. The index of goods prices is normalized to $P_t = 1$. The price of housing is denoted by q_{ct}. Consumers maximize their utility subject to their budget constraint. This utility maximization problem yields the expected demand equations for goods and real estate.

Workers have access to a time-varying outside option utility v_t^*. We can think of this as the utility offered by remaining in the rural sector or immigrating to another country. In equilibrium, this implies that the indirect utility function of workers must satisfy,

$$V_{ct} = \ln(w_{ct}) - v \ln(q_{ct}) = v_t^* . \tag{2}$$

[12] Adding savings would complicate the model, but as long as capital is mobile across locations it will not alter our basic estimating equation, nor will it influence our empirical results, which are derived from a comparison across locations within a country.

Landlords receive income from land and other local resources. To keep things simple, we think of these landlords as living outside of the cities we study.

2.2 Production

Workers can decide to become entrepreneurs by paying a fixed cost F, denominated in units of labor, to open a firm. The measure of firms in a city-industry is denoted by n_{ict}. We think of firms in a city as being started by workers from that city in the previous period, so if they enter in industry i they begin with the initial technology level available in that industry in that city, denoted $\bar{a}_{icft}$. Firms then invest in R&D to obtain a new technology level, a_{icft}, which is used in production.

Firms compete on perfectly competitive input and output markets. The production function for firm f in industry i and city c is,

$$y_{icft} = a_{icft} L_{icft}^{\alpha_1} H_{icft}^{\alpha_2} R_{icft}^{\beta} E_{icft}^{1-\alpha_1-\alpha_2-\beta}, \tag{3}$$

where a_{icft} is technology, L_{icft} is labor input, R_{icft} is the resource input, H_{icft} is real estate input, E_{icft} is entrepreneurial effort, $\alpha_1 + \alpha_2 + \beta < 1$, and $\alpha_1, \alpha_2, \beta > 0$.

Entrepreneurial effort is supplied by workers who choose to open a firm. Each entrepreneur has access to only one unit of entrepreneurial effort, so in equilibrium $E_{icft} = 1$ for all firms. This reflects a span-of-control limitation for firm owners.[13] This span-of-control limitation plays an important role in the model; by introducing decreasing returns to scale at the firm level it pins down firm size. As we will see, this implies that growth in city-industry employment is driven by growth in the number of firms.

Labor is the only production input that is mobile across locations.[14] Including real estate in the production function is not central to the model but is done for completeness.

[13]Note that entrepreneurs are not required to trade off entrepreneurial effort against labor effort. Instead, all workers have access to one unit of each type of input, but entrepreneurial effort can only be used by workers that choose to open a firm.

[14]Adding additional mobile inputs, such as capital, would not substantially affect the estimating equation that we obtain.

2.3 Land and natural endowments

Locational fundamentals play a central role in the debate over the determinants of city size, so it is important that they be incorporated into the theory (see, e.g., Davis & Weinstein (2002)). In our model, locational fundamentals are represented by fixed industry-specific city resource endowments, $\bar{R}_{ic}$.[15] In equilibrium, the markets for local resources clear, so $\int_{f=0}^{n_{ict}} R_{icft} df = \bar{R}_{ic}$. Resources play an important role in the model; by introducing decreasing returns at the city-industry level, they allow firms in the same industry to be active in many locations with different technology levels, even when trade is free, labor is mobile, and firms are perfectly competitive. They are also important in the context of the empirical analysis, because they make the impact of locational fundamentals in the estimation strategy explicit.

Real estate, which is used by both workers and firms, represents a congestion force in our theory. We model the price of real estate as an increasing function of the number of workers in a city and the amount of land used by producers:

$$q_{ct} = f\left(L_{ct}, \sum_i \sum_f H_{icft}\right). \tag{4}$$

For our purposes, it is not necessary that we take a stand on the particular functional form of this relationship.

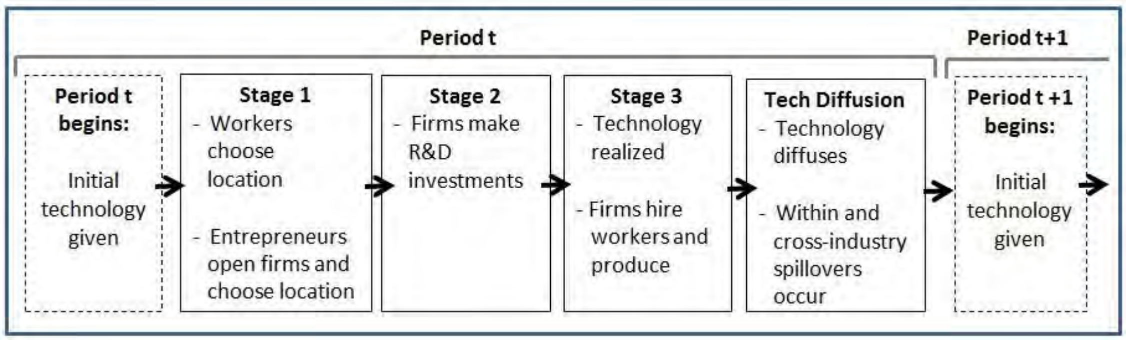
2.4 Timing

Figure 1 describes the timing in the model. At the beginning of each period, firms in the same city-industry share a common and observable technology level denoted $\bar{a}_{icft}$. Given these, workers choose where to locate and whether to pay a fixed cost to become an entrepreneur and open a firm. After workers have moved and firms have opened, firms then choose a level of R&D investment and realize a new technology level a_{icft}. Once technology is realized, firms choose how many workers and other inputs to hire and they produce and sell their outputs. At the end of the period, technology diffusion and technology spillovers occur, leading each firm to attain a

[15] This approach follows Jones (1975) and has recently been used to study the regional effects of international trade by Kovak (2013) and Dix-Carneiro & Kovak (2015).

new technology level $\bar{a}_{icft+1}$. The static portion of the model is solved by starting at Stage 3 and solving backwards.

Figure 1: Model timing



2.5 Production: Stage 3

At the beginning of stage three, the number of workers in a city, L_{ct}, the number of firms in a city-industry, n_{ict}, and the technology level available to each firm a_{icft} have been determined. Given these, firms maximize profits by solving,

$$\max_{L_{icft}, H_{icft}, R_{icft}} p_{it} a_{icft} L_{icft}^{\alpha_1} H_{icft}^{\alpha_2} R_{icft}^{\beta} - w_{ct} L_{icft} - q_{ct} H_{icft} - r_{ict} R_{icft}$$

where w_{ct} is the wage and r_{ict} is the price of local resources. Since entrepreneurs will employ all of the entrepreneurial effort available to them, E_{ict} is not included in this optimization problem. Using the first order conditions for this expression, gross profits – which are the returns to entrepreneurial effort excluding fixed costs of entry and R&D expenditures – are:

$$\pi_{icft} = \left(\frac{w_{ct}^{\alpha_1} q_{ct}^{\alpha_2} r_{ict}^{\beta}}{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2} \beta^{\beta}} \right)^{\frac{-1}{1-\alpha_1-\alpha_2-\beta}} p_{it}^{\frac{1}{1-\alpha_1-\alpha_2-\beta}} a_{icft}^{\frac{1}{1-\alpha_1-\alpha_2-\beta}} (1-\alpha_1-\alpha_2-\beta) \quad (5)$$

Local resource market clearing allows us to solve for the rental rate:

$$r_{ict} = \bar{R}_{ic}^{\frac{-(1-\alpha_1-\alpha_2-\beta)}{1-\alpha_1-\alpha_2}} \left(\frac{w_{ct}^{\alpha_1} q_{ct}^{\alpha_2}}{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2}} \right)^{\frac{-1}{1-\alpha_1-\alpha_2}} \left(p_{it}^{\frac{1}{1-\alpha_1-\alpha_2}} \sum_{f=0}^{n_{ict}} a_{icft}^{\frac{1}{1-\alpha_1-\alpha_2}} \right)^{\frac{1-\alpha_1-\alpha_2-\beta}{1-\alpha_1-\alpha_2}} \quad (6)$$

$$r_{ict} = \bar{R}_{ic}{}^{1-\alpha_1-\alpha_2} \quad \frac{ct \quad ct}{\alpha^{\alpha_1}\alpha^{\alpha_2}\beta^{1-\alpha_1-\alpha_2}} \quad p_{it}^{1-\alpha_1-\alpha_2} \quad {}_{f=0} \; a_{icft}^{1-\alpha_1-\alpha_2-\beta}\, df \qquad (6)$$

2.6 Producers: Stage 2

At the beginning of stage two, workers and firms have already made their location decisions and firms have access to an initial technology level $\bar{a}_{icft}$. Given these, firms must choose how much to invest in R&D to increase their productivity in order to maximize profits. In doing so, they take into account the production decisions that we solved for above.

When firms conduct R&D, they are choosing a technology multiplier $\varphi_{icft} \geq 0$ that increases their initial technology level according to,

$$a_{icft} = (1 + \varphi_{icft})^{\delta}\, \bar{a}_{icft}\,, \tag{7}$$

at a cost $w_{ct}\, C\, \varphi_{icft}$, where C is a parameter that determines the labor cost of innovation.[16] We assume that $\delta < 1 - \alpha_1 - \alpha_2 - \beta$ so that the firm's profit function is concave in the R&D investment level.

Firms choose the innovation investment that maximizes gross profits less R&D expenses,

$$\max_{\varphi_{icft}} (1+\varphi_{icft})^{\frac{\delta}{1-\alpha_1-\alpha_2-\beta}}\, \bar{a}_{icft}^{\frac{1}{1-\alpha_1-\alpha_2-\beta}}\, p_{it}^{\frac{1}{1-\alpha_1-\alpha_2-\beta}} \left(\frac{w_{ct}^{\alpha_1} a_{ct}^{\alpha_2} r_{ict}^{\beta}}{\alpha_1^{\alpha_1}\alpha_2^{\alpha_2}\beta^{\beta}}\right)^{\frac{-1}{1-\alpha_1-\alpha_2-\beta}} (1-\alpha_1-\alpha_2-\beta) - w_{ct} C \varphi_{icft}$$

subject to $\varphi_{icft} \geq 0$. For now, assume that there is an interior solution to this problem, so that $\varphi_{icft} > 0$. In this case, the first order conditions for the firm's problem can be used to obtain the following expression for the firm's R&D decision, where the resource rent is substituted out using Eq. 6:

[16]While we do not allow the R&D cost to vary by industry here, allowing an industry-specific cost parameter would not fundamentally alter our results.

$$(1+\varphi_{icft}) = \left[\left(\frac{\delta}{C}\right)^{1-\alpha_1-\alpha_2-\beta} \bar{a}_{icft} \bar{R}_{ic}^{\frac{\beta(1-\alpha_1-\alpha_2-\beta)}{1-\alpha_1-\alpha_2}} w_{ct}^{\frac{(\alpha_2-1)(1-\alpha_1-\alpha_2-\beta)}{1-\alpha_1-\alpha_2}} p_{it}^{\frac{1-\alpha_1-\alpha_2-\beta}{1-\alpha_1-\alpha_2}} \left(\frac{q_{ct}^{\alpha_2}}{\alpha_1^{\alpha_1}\alpha_2^{\alpha_2}}\right)^{\frac{-(1-\alpha_1-\alpha_2-\beta)}{1-\alpha_1-\alpha_2}} \left(\sum_{f=0}^{n_{ict}} \bar{a}_{icft}^{\frac{1}{1-\alpha_1-\alpha_2-\beta}}\right)^{\frac{-\beta(1-\alpha_1-\alpha_2-\beta)}{1-\alpha_1-\alpha_2}} \right]^{\frac{1}{1-\alpha_1-\alpha_2-\beta-\delta}} \tag{8}$$

Recalling that $1-\delta-\alpha_1-\alpha_2-\beta > 0$, this equation tells us that for an individual firm the optimal level of innovation is increasing in the firm's initial technology level and the city-industry resource endowment. The level of innovation is decreasing in the cost of R&D, the wage level, and the amount of competition the firm faces for local resources, represented by the integral over the technology level of all other firms in the city-industry.

Suppose for now that all firms in an industry start with the same initial technology level $\bar{a}_{icft}$. Later, we will see that this is the case given how the technology diffusion process is modeled.[17] In this case, firms in the industry will face the same R&D optimization problem, which implies that they will all choose the same R&D investment level, which we label φ^*_{ict}. Firms will be aware of the R&D decisions made by other firms and will take this into account when making their own decisions. The R&D investment consistent with these expectations is found by substituting Eq. 7 into Eq. 8 and solving to obtain,

$$(1+\varphi^*_{ict}) = \left[\left(\frac{\delta}{C}\right)^{1-\alpha_1-\alpha_2} \bar{a}_{icft} \bar{R}_{ic}^{\beta} \left(\frac{q_{ct}^{\alpha_2}}{\alpha_1^{\alpha_1}\alpha_2^{\alpha_2}}\right)^{-1} p_{it} w_{ct}^{\alpha_2-1} n_{ict}^{-\beta} \right]^{\frac{1}{1-\alpha_1-\alpha_2-\delta}} \tag{9}$$

This expression tells us that the level of innovation by a firm is increasing in the firm's initial technology level. At the same time it is decreasing in the number of firms in the same city-industry, which implies more competition for fixed city-industry resources.

[17]To keep things simple, and because firm heterogeneity is not central to the exercise undertaken in

this paper, we have decided not to include within-industry firm heterogeneity in the model. However, firm heterogeneity could potentially be incorporated in a more sophisticated version of the model.

2.7 Producers: Stage 1

Next, we consider the entry decisions of firms. Any worker can choose to become an entrepreneur by paying a fixed cost (in terms of labor) of F. Because there is a large supply of potential entrants, *ex post* profits will be driven to zero. Thus, in equilibrium $\pi_{icft} - w_{ct}C\varphi_{icft} = w_{ct}F$. Using this zero profit condition together with Eqs. 5, 6, and 9, we solve for the number of firms in a city-industry:

$$n_{ict} = w_{ct}^{\frac{\alpha_2 - 1}{\beta}} \left(\frac{q_{ct}^{\alpha_2}}{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2}} \right)^{\frac{-1}{\beta}} \left(\frac{1}{n_{it}^{\beta} \bar{a}_t^{\beta}} \right) \left(\frac{1 - \alpha_1 - \alpha_2 - \beta - \delta}{F - C} \right)^{\frac{1-\alpha_1-\alpha_2-\delta}{\beta}} \bar{R}_{ic} \left(\frac{\delta}{C} \right)^{\frac{\delta}{\beta}} \tag{10}$$

This, together with the first-order conditions from the firm's problem in Stage 3, gives city-industry employment:

$$L_{ict} = \alpha_1 w_{ct}^{\frac{\alpha_2 - 1}{\beta}} \left(\frac{q_{ct}^{\alpha_2}}{\alpha_1^{\alpha_1} \alpha_2^{\alpha_2}} \right)^{\frac{-1}{\beta}} \left(\frac{1}{n_{it}^{\beta} \bar{R}_{ict} \bar{a}^{\beta}} \right) \left(\frac{\delta}{C} \right)^{\frac{\delta}{\beta}} \left(\frac{1 - \alpha_1 - \alpha_2 - \beta - \delta}{F - C} \right)^{\frac{1-\alpha_1-\alpha_2-\beta-\delta}{\beta}} \tag{11}$$

Together, Eqs. 10 and 11 imply the following relationship between the number of firms and the number of workers in a city-industry,

$$L_{ict} = \left(\frac{\alpha_1 (F - C)}{1 - \alpha_1 - \alpha_2 - \beta - \delta} \right) n_{ict}. \tag{12}$$

Eq. 12 shows that growth in city-industry employment is driven entirely by growth in the number of firms. Eq. 10 can also be used, together with Eq. 9 to solve for the equilibrium level of R&D in the industry:

$$(1 + \varphi_{ict}^{*}) = \frac{\delta(F - C)}{C(1 - \alpha_1 - \alpha_2 - \beta - \delta)} \tag{13}$$

This expression shows that firms' R&D investments will depend only on model parameters, a useful feature that simplifies the results. So far we have solved the model assuming that $\varphi_{ict}* > 0$. For this to hold, we need, $F > C \left(\frac{1-\alpha_1-\alpha_2-\beta}{\delta} \right)$.[18] For

[18]Note that the expression in parenthesis can be interpreted as the ratio of the gains from additional firms in an industry to the gains from improved technology in the industry. For industries

the remainder of this theory we assume that this condition is satisfied so that we have an interior solution to the firm's R&D optimization problem and R&D occurs in all industries.

2.8 Spillovers and technology diffusion

At the end of a period, after production and consumption have taken place, firms are able to copy technology from other firms in the same industry (diffusion). However, because all firms in a given city-industry end the period with the same technology level, the role of diffusion is simply to rule out strategic behavior. In addition, entrepreneurs may share ideas, and this recombination of ideas can increase their productivity (spillovers). Following Glaeser *et al.* (1992), we write the growth rate in technology at the city-industry level as,

$$\ln\left(\frac{\bar{a}_{ict}}{a_{ict-1}}\right) = S_{ict-1} + f_{ict} \tag{14}$$

where $S_{ict-1} \geq 1$ represent the amount of spillovers available to a city-industry in a period. This can include within-industry effects, cross-industry spillovers, as well as national industry technology growth or city-level aggregate spillovers.

We can use Eq. 14 to translate the growth in (unobservable) city-industry technology into the growth of (observable) city-industry employment. Using Eq. 2, Eq. 7, Eq. 11, and Eq. 14, we obtain,

$$\begin{aligned} \Delta\ln(L_{ict}) &= \left(\frac{1}{\beta}\right) S_{ict} + \left(\frac{1}{\beta}\right)\Delta\ln(p_{it}) + \left(\frac{\alpha_2 - 1}{\beta}\right)\Delta\ln(\bar{v}_t *) \\ &+ \left(\frac{v(\alpha_2 - 1) - \alpha_2}{\beta}\right)\Delta\ln(q_{ct}) + \zeta + \varepsilon_{ict} \end{aligned} \tag{15}$$

where ζ is a constant function of model parameters. Note that by differencing we have eliminated the local resource endowment from this equation.

As a final step, we need to decide how to model the spillover term. Existing empirical evidence provides little guidance here, so we will opt for a fairly simple

where the inequality above doesn't hold, there will be no innovation.

approach.[19] We model the spillovers benefits to firms in industry i from R&D in local firms in industry k as a function of (1) the amount of new ideas produced in industry k, which is a function of the size of the technology advance made by each firm $(1 + \varphi^*_{kct})$, and the number of firms, n_{kct}, and (2) the usefulness of these ideas to firms in industry i, given by parameter τ_{ki}. Thus, there is a matrix of τ_{ki} parameters representing the usefulness of an idea from industry k to producers in industry i. The diagonal τ_{ii} terms reflects within-industry spillovers.[20] Given this, we write the spillover function as,[21]

$$S_{ict} = \sum_k \tau_{ki} \ln\left(n_{kct}(1+\varphi_{kct})\right) + \xi_{it} + \psi_{ct}.$$

Using Eqs. 12 and 13, this can be rewritten as,

$$S_{ict} = \sum_k \tau_{ki} \ln(L_{kct}) + \xi_{it} + \psi_{ct} + \Gamma$$

where Γ is a constant term. Combining this with Eq. 15 we obtain,

$$\begin{aligned} \ln(L_{ict}) - \ln(L_{ict-1}) &= \left(\frac{1}{\beta}\right)\left[\tau_{ii}\ln(L_{ict}) + \sum_{k\neq i}\tau_{ki}\ln(L_{kct})\right] \\ &+ \left[\ln(p_{it}) - \ln(p_{it-1})\right] + \xi_{it} \\ &- \left[v(\alpha_2 - 1) - \alpha_2\right]\left[\ln(q_{ct}) - \ln(q_{ct-1})\right] + \psi_{ct} \\ &+ (\alpha_2 - 1)\left[\ln(\overline{v}^*_t) - \ln(\overline{v}^*_{t-1})\right] + \tilde{\Gamma} + \epsilon_{ict} \end{aligned} \tag{16}$$

[19]In the empirical analysis we will investigate the robustness of our results to some reasonable alternative formulations.

[20]The intuition behind the within-industry spillovers in this model is that, while all firms achieve the same new technology level after undertaking R&D, this new level need not be achieved in exactly the same way. As a result, it may be possible for firms to achieve further gains by observing the different types of technologies developed by their competitors. However, the potential gains from within-industry spillovers will depend on a number of factors, such as the willingness for firms in an industry to share ideas with their direct local competitors.

[21]Here we are assuming that city-industry resource endowments are such that $n_{kct} \geq 1$. This assumption allows us to express the spillover term in a slightly simpler way, but is not central to our results. If we are worried that n_{kct} can fall below one then we would instead write this as $S_{ict} = \sum_k \tau_{ki} \ln(max(n_{kct}(1+\varphi_{kct}), 0)) + \xi_{it} + \psi_{ct}$.

where the constant terms have been gathered into $\tilde{\Gamma}$. This equation expresses the change in log employment in industry i and location c in terms of (1) within-industry spillovers generated by employment in industry i, (2) cross-industry spillovers, (3) national industry-specific factors that affect industry i in all locations, (4) city-specific factors that affect all industries in a location, and (5) aggregate changes in the outside option of workers that affect all industries in all locations.

This expression for city-industry growth will motivate our empirical specification.[22] One feature that is worth noting here is that we have two factors, city-level aggregate spillovers ψ_{ct} and city congestion costs q_{ct}, both of which vary at the city-year level. Empirically we will not be able to separate these positive and negative effects and so we will only be able to identify their net impact. Similarly, we cannot separate positive and negative effects that vary at the industry-year level.

Note that in the absence of spillovers, and with common technologies across locations, the city size distribution in this model will be determined by the distribution of local resource endowments. Once local technology spillovers are added, city sizes will be determined by a combination of the initial resource endowment and the evolving technology levels. This hybrid of locational fundamentals and increasing returns is consistent with some existing empirical results (e.g., Davis & Weinstein (2002) and Bleakley & Lin (2012)). Once spillovers are included, the dynamics of the system are complex and depend crucially on the matrix of τ_{ki} parameters.[23] Estimating these parameters is the goal of our empirical exercise, which we turn to next.

[22] There are at least two promising alternative theories that may yield an empirical specification similar to the expression generated by our model. One such theory could combine static inter-industry connections, such as pecuniary spillovers through intermediate-goods sales, with changing transport costs. A second alternative combines static agglomeration forces with a friction that results in a slow transition towards a static equilibrium. Our empirical exercises cannot make a sharp distinction between the mechanisms described in our framework and these alternatives, so they should not be interpreted as a direct test of the particular agglomeration mechanism described by the theory.

[23] In addition, the dynamics are likely to depend crucially on city-size congestion forces, which are not fully modeled here. Because the primary goals of this paper are empirical, we leave a full exploration of these dynamics for future work.

3 Data

The main database used in this study was constructed from thousands of pages of original British Census of Population summary reports. The decennial Census data were collected by trained registrars during a relatively short time period, usually a few days in April of each census year. As part of the census, individuals were asked to state their occupation, but the reported occupations correspond more closely to industries than to what we think of as occupations today.[24] A unique feature of this database is that the information is drawn from a full census. Virtually every person in the cities we study provided information on their occupation and all of these answers are reflected in the employment counts in our data.[25]

The database includes 31 cities for which occupation data were reported in each year from 1851-1911, containing 28-34% of the English population over the period we study. The geographic extent of these cities changes over time as the cities grow, a feature that we view as desirable for the purposes of our study.[26] Appendix A.2 provides a list of the cities included in the database, as well as a map showing the location of these cities in England. In general, our analysis industries cover the majority of the working population of the cities, with most of the remainder employed by the government or in agriculture.

The industries in the database span manufacturing, food processing, services and professionals, retail, transportation, construction, mining, and utilities. Because the occupational categories listed in the census reports varied over time, we combined multiple industries in order to construct consistent industry groupings over the study

[24]Examples from 1851 include "Banker", "Glass Manufacture" or "Cotton manufacture". The database does include a few occupations that do not directly correspond to industries, such as "Labourer", "Mechanic", or "Gentleman", but these are a relatively small share of the population. These categories are not included in the analysis. In 1921 the Census office renamed what had previously been called "occupation" to be "industry" and then introduced a new set of data reflecting occupation in the modern sense.

[25]This contrasts with data based on census samples, which often covers 5% or 1% of the available data. We have experimented with data based on a census sample (from the U.S.) and found that, when cutting the data to the city-industry level, sampling error has a substantial effect on the consistency and robustness of the results.

[26]Other studies in the same vein, such as Michaels *et al.* (2013), also use metropolitan boundaries that expand over time. The alternative is working with fixed geographic units. While that may be preferred for some types of work, given the growth that characterizes most of the cities in our sample, using fixed geographic units would mean either that the early observations would include a substantial portion of rural land surrounding the city, or that a substantial portion of city growth would not be part of our sample in the later years. Either of these options is undesirable.

period. This process generates 26 consistent private sector occupation categories.[27] Of these, 23 can be matched to the connections matrices used in the analysis. Table 7 in Appendix A.2 describes the industries included in the database.

A preliminary analysis, using the agglomeration measure from Ellison & Glaeser (1997), suggests that the agglomeration patterns observed in our data are similar to those documented in modern studies (details in Appendix A.2, Tables 8-9). Britain's main manufacturing and export industries, such as Textiles, Metal & Machines, and Shipbuilding, show high levels of geographic agglomeration. Many non-traded services or retail industries, including Merchants, Agents, Etc., Construction, and Shopkeepers, Salesmen, Etc. show low levels of agglomeration. Overall, the median level of industry agglomeration is between 0.02 and 0.027, which is comparable to the levels reported for the modern U.S. economy by Ellison & Glaeser (1997) and somewhat larger than the levels reported for the modern British economy by Faggio *et al.* (2013).[28]

This study also requires a set of matrices measuring the pattern of connections between industries. These measures should reflect the channels through which ideas may flow between industries. Existing literature provides some guidance here. Marshall (1890) suggested that firms may benefit from connections operating through input-output flows, the sharing of labor pools, or other types of technology spillovers. The use of input-output connections is supported by recent literature showing that firms share information with their customers or suppliers.[29] To reflect this channel, we use an input-output table constructed by Thomas (1987) based on the 1907 British Census of Production (Britain's first industrial census).[30] We construct two

[27]Individual categories in the years were combined into industry groups based on (1) the census' occupation classes, and (2) the name of the occupation. Further details of this procedure are available in the Online Appendix.

[28]Using industry data for 459 manufacturing industries at the four-digit level and 50 states, Ellison & Glaeser (1997) calculate a mean agglomeration index of 0.051 and a median of 0.026. For Britain, Faggio *et al.* (2013) calculate industry agglomeration using 94 3-digit manufacturing industries and 84 urban travel-to-work areas. They obtain a mean agglomeration index of 0.027 and a median of 0.009. Kim (1995) calculates an alternative measure of agglomeration for the U.S. during the late 19th and early 20th centuries, but given that he studies only manufacturing industries, and given the substantial differences between his industry definitions and our own, it is difficult to directly compare to his results.

[29]For example, Javorcik (2004) and Kugler (2006) provide evidence that the presence of foreign firms (FDI) affects the productivity of upstream and downstream domestic firms.

[30]For robustness exercises, we have also collected an input-output table for 1841 constructed by Horrell *et al.* (1994) with 12 more aggregated industry categories. See Appendix A.2 for more details.

variables: $IOin_{ij}$, which gives the share of industry i's intermediate inputs that are sourced from industry j, and $IOout_{ij}$ which gives the share of industry i's sales of intermediate goods that are purchased by industry j. One drawback of using these matrices is that they are for intermediate goods; they will not capture the pattern of capital goods flows.

Another channel for knowledge flow is the movement of workers, who may carry ideas between industries.[31] To reflect this channel, we construct two different measures of the similarity of the workforces used by different industries. The first measure is based on the demographic characteristics of workers (their age and gender) from the 1851 Census. These features had an important influence on the types of jobs a worker could hold during the period we study.[32] For any two industries, our demographic-based measure of labor force similarity, EMP_{ij} , is constructed by dividing workers in each industry into these four available bins (male/female and over20/under20) and calculating the correlation in shares across the industries. A second measure of labor-force similarity, based on the occupations found in each industry, is more similar to the measures used in previous studies. This measure is built using U.S. census data from 1880, which reports the occupational breakdown of employment by industry. We map the U.S. industry categories to the categories available in our analysis data. Then, for any two industries our occupation-based measure of labor force similarity, OCC_{ij} is the correlation in the vector of employment shares for each occupation.

4 Empirical approach

The starting point for our analysis is based on Equation 16, which represents the growth rate of a city-industry as a function of within and cross-industry agglomeration effects as well as time-varying city-specific and national industry-specific factors. Rewriting this as a regression equation we have,

$$\Delta \ln(L_{ict+1}) = \tilde{\tau}_{ii} \ln(L_{ict}) + \sum_{k \neq i} \tilde{\tau}_{ki} \ln(L_{kct}) + \theta_{ct} + \chi_{it} + e_{ict} \tag{17}$$

[31] Research by Poole (2013) and Balsvik (2011), using data from Brazil and Norway, respectively, has highlighted this channel of knowledge flow.

[32] For example, textile industries employed substantial amounts of female and child labor, while metal and heavy machinery industry jobs were almost exclusively reserved for adult males.

where Δ is the first difference operator, $\tilde{\tau}_{ii}$ and $\tilde{\tau}_{ki}$ include $1/\beta$, θ_{ct} is a full set of city-year effects and χ_{it} is a full set of industry-year effects. The first term on the right hand side represents within-industry spillovers, while the second term represents cross-industry spillovers.[33]

One issue with Equation 17 is that there are too many parameters for us to credibly estimate given the available data. In order to reduce the number of parameters, we need to put additional structure on the spillover terms. As discussed in the previous section, we follow recent literature in this area, particularly Ellison *et al.* (2010), by parameterizing the connections between industries using the available input-output and labor force similarity matrices:

$$\tilde{\tau}_{ki} = \beta_1 IOin_{ki} + \beta_2 IOout_{ki} + \beta_3 EMP_{ki} + \beta_4 OCC_{ki} \quad \forall\, i, k$$

Substituting this into 17 we obtain:

$$\begin{aligned}\Delta \ln(L_{ict+1}) &= \tilde{\tau}_{ii}\ln(L_{ict}) + \beta_1 \sum_{k \neq i} IOin_{ki}\ln(L_{kct}) + \beta_2 \sum_{k \neq i} IOout_{ki}\ln(L_{kct}) \\ &+ \beta_3 \sum_{k \neq i} EMP_{ki}\ln(L_{kct}) + \beta_4 \sum_{k \neq i} OCC_{ki}\ln(L_{kct}) + \theta_{ct} + \chi_{it} + e_{ict} \qquad (18)\end{aligned}$$

Instead of a large number of parameters measuring spillovers across industries, Equation 18 now contains only four parameters multiplying four (weighted) summations of log employment. Summary statistics for the cross-industry spillover terms are available in Appendix Table 10.

There is a clear parallel between the specification in Equation 18 and the empirical approach used in the convergence literature (Barro & Sala-i Martin (1992)). A central debate in this literature has revolved around the inclusion of fixed effects for the cross-sectional units (see, e.g., Caselli *et al.* (1996)). In our context, the inclusion of such characteristics could help control for location and industry-specific factors that

[33] We purposely omitted the last term of Equation 16, $\Delta \ln(\bar{v}_t^*)$, because although it could be estimated as a year-specific constant, it would be collinear with both the (summation of) industry-year and city-year effects. Moreover, in any given year we also need to drop one of the city or industry dummies in order to avoid collinearity. In all specifications we chose to drop the industry-year dummies associated with the "General services" sector.

affect the growth rate of industry and are correlated with initial employment levels. However, the inclusion of city-industry fixed effects in Equation 18 will introduce a mechanical bias in our estimated coefficients (Hurwicz (1950), Nickell (1981)). This bias is a particular concern in a setting where the time-series is limited. Solutions to these issues have been offered by Arellano & Bond (1991), Blundell & Bond (1998), and others, yet these procedures can also generate biased results, as shown by Hauk Jr. & Wacziarg (2009). In a recent review, Barro (2012) uses data covering 40-plus years and argues (p. 20) that in this setting, "the most reliable estimates of convergence rates come from systems that exclude country fixed effects but include an array of X variables to mitigate the consequence of omitted variables." Our approach essentially follows this advice, but with the additional advantage that we have two cross-sectional dimensions, which allows for the inclusion of flexible controls in the form of time-varying city and industry effects.

There are two issues to address at this point. First, there could be measurement error in L_{ict}. Since this variable appears both on the left and right hand side, this would mechanically generate an attenuation bias in our within-industry spillover estimates. Moreover, since L_{ict} is correlated with the other explanatory variables, such measurement error would also bias the remaining estimates. We deal with measurement error in L_{ict} on the right hand side by instrumenting it with lagged city-industry employment.[34] Under the assumption that the measurement error in any given city-industry pair is *iid* across cities and time, our instrument is $L^{Inst}_{ict} = L_{ict-1} \times g_{i-ct}$, where L_{ict-1} is the lag of L_{ict} and g_{i-ct} is the decennial growth rate in industry i computed using employment levels in all cities *except* city c, as in Bartik (1991).

Second, we are also concerned that there may be omitted variables that affect both the level of employment in industry j and the growth in employment in industry i. Such variables could potentially bias our estimated coefficients on both the cross-industry and (when $j = i$) the within-industry spillovers. For instance, if there is some factor not included in our model which causes growth in two industries i and $k \neq i$ in the same city, a naive estimation would impute such growth to a spillover effect from k to i, thus biasing the estimated spillover upward. Our lagged instrumentation approach can also help us deal with these concerns. Specifically, when using instruments with a one-decade lag to address endogeneity concerns the

[34]This approach is somewhat similar to the approach introduced by Bartik (1991) and has been suggested by Combes *et al.* (2011).

exclusion restriction is that there is not some omitted variable that is correlated with employment in some industry k in period t and affects employment growth in industry i from period $t+1$ to $t+2$. Moreover, the omitted variable cannot affect growth in all industries in a location, else it would be captured by the city-year fixed effect, nor can it affect the growth rate of industry i in all cities.[35] Thus, while our approach does not allow us to rule out all possible confounding factors, it allows us to narrow the set of potential confounding forces relative to most previous work in this area. Now, for the cross-industry case, the summation terms in Equation 18 such as $\sum_{k \neq i} IOin_{ki} \ln(L_{kct})$ are instrumented with $\sum_{k \neq i} IOin_{ki} \ln(L^{Inst}_{kct})$, where L^{Inst}_{kct} is computed as described above.

The estimation is performed using OLS or, when using instruments, two-stage least squares. Correlated errors are a concern in these regressions. Specifically, we are concerned about serial correlation, which Bertrand *et al.* (2004) argue can be a serious concern in panel data regressions, though this is perhaps less of a concern for us given the relatively small time dimension in our data. A second concern is that industries within the same city are likely to have correlated errors. A third concern, highlighted by Conley (1999) and more recently by Barrios *et al.* (2012), is spatial correlation occurring across cities. Here the greatest concern is that error terms may be correlated within the same industry across cities (though the results presented in Appendix A.4.5 suggest that cross-city effects are modest).

To deal with all of these concerns we use multi-dimensional clustered standard errors following work by Cameron *et al.* (2011) and Thompson (2011). We cluster by (1) city-industry, which allows for serial correlation; (2) city-year, which allows for correlated errors across industries in the same city and year; and (3) industry-year, which allows for spatial correlation across cities within the same industry and year. This method relies on asymptotic results based on the dimension with the fewest number of clusters. In our case this is 23 industries $\times$ 6 years = 138, which should be large enough to avoid serious small-sample concerns.

In order to conduct underidentification and weak-instrument tests while clustering standard errors in multiple dimension, we produced a new statistical package following the approach from Kleibergen & Paap (2006). This was necessary because existing statistical packages are unable to calculate these tests correctly when cluster-

[35]The results are not sensitive to the length of the lag used in the instrumentation. We have experimented with two- and three-decade lags and obtained essentially the same results.

ing by more than two dimensions. The procedure used to generate our new statistical package is described in Appendix A.3.2. Our package, which we plan to make publicly available, can accommodate clustering across an arbitrary number of dimensions, which is likely to be useful for future researchers.

Finally, we may be concerned about how well our estimation procedure performs in a data set of the size available in this study. To assess this, we conduct a series of Monte Carlo simulations in which we construct 1000 new data sets with a size and error structure based on the true data, but with known spillover parameter values. We then apply our estimation procedure to these simulated data in order to obtain a distribution of placebo coefficient estimates, which can then be compared to the estimates obtained using the true data. These simulations, which are described in more detail in Appendix A.3.1, suggest that our estimation procedure performs well in datasets with a size and error structure similar to the true data.

To simplify the exposition, we will hereafter collectively refer to the set of regressors $\ln(L_{ict})$ for $i = 1...I$ as the *within* variables. Similarly, with a small abuse of notation the term $\sum_{k \neq i} IOin_{ki} \ln(L_{kct})$ is referred to as $IOin$, and so on for $IOout$, EMP, and OCC. We collectively refer to the latter terms as the *between* regressors since they are the parametrized counterpart of the spillovers across industries.

5 Main results

Our main regression results are based on the specification described in Equation 18. The estimation strategy involves using four measures for the pattern of cross-industry spillovers: forward input-output linkages, backward input-output linkages, and two measures of labor force similarity. We begin our analysis in Table 1 by looking at results that include only one of these at a time. Columns 1-3 include only the forward input-output linkages; Columns 1 presents OLS results; Column 2 presents results with lagged instrumentation on the within terms; and Column 3 uses lagged instrumentation for both the within and between terms. A similar pattern is used for backward input-output linkages in Columns 4-6, the demographic-based labor force similarity measure in Columns 7-9, and the occupation-based labor force similarity measure in Columns 10-12. All of these results include a full set of industry-specific within-industry terms, but these are not reported in Table 1 for space reasons.[36]

These results show strong positive effects operating through forward input-output connections, suggesting that local suppliers play an important role in industry growth. The importance of local suppliers to industry growth is perhaps the clearest and most robust result emerging from our analysis. There is little evidence of positive effects operating through local buyers. The results do provide some evidence that the presence of other industries using similar labor pools may increase growth, particularly when using the more detailed OCC measure. A comparison across columns for each spillover measure shows that the IV results do not differ from the OLS results in a statistically significant way, suggesting that any measurement error or omitted variables concerns addressed by instruments are not generating substantial bias in the OLS results.

Table 2 considers all four channels simultaneously. Columns 1-3 present results in which we estimate a single coefficient on the within-industry terms. Columns 4-6 present results in which we estimate industry-specific within-industry effects. These heterogeneous within-industry coefficients, which are not reported in Table 2, will be explored later. Columns 1 and 4 presents OLS results. In Column 2 and 5 we instrument the within terms. In Column 3 and 6 we use instruments for both the within and between terms. The results are generally similar to those from Table 1; the presence of local suppliers or industries employing a similar labor force both appear to enhance city-industry growth. The presence of local buyers has no positive effect. In Columns 1-3, we can see that the within term is negative, suggesting that on average across all industries employment growth is slower in locations where initial industry employment is already large.

[36]We do not report first-stage results for our instrumental variables regressions because these involve a very large number of first-stage regressions. Instead, for each specification we report the test statistics for the Lagrange Multiplier underidentification test based on Kleibergen & Paap (2006) as well as the test static for weak instruments test based on the Kleibergen-Paap Wald statistic. It is clear from these statistics that weak instruments are not a substantial concern in these specifications.

Table 1: OLS and IV regressions including only one spillover path at a time

	(1)	(2)	(3)	(4)	(5)	(6)
IOin	0.0581***	0.0440***	0.0417***			
	(0.0128)	(0.0113)	(0.0113)			
IOout				-0.0030	-0.0105	-0.0143
				(0.0108)	(0.0112)	(0.0113)
Observations	4,263	3,549	3,549	4,263	3,549	3,549
Estimation	ols	2sls	2sls	ols	2sls	2sls
Instrumented	none	wtn	wtn-btn	none	wtn	wtn-btn
KP under id.		29.27	32.40		23.18	23.37
KP weak id.		77.59	75.24		61.98	61.10

	(7)	(8)	(9)	(10)	(11)	(12)
EMP	0.0009	0.0022*	0.0017			
	(0.0017)	(0.0013)	(0.0014)			
OCC				0.0058**	0.0058*	0.0060*
				(0.0029)	(0.0032)	(0.0032)
Observations	4,263	3,549	3,549	4,263	3,549	3,549
Estimation	ols	2sls	2sls	ols	2sls	2sls
Instrumented	none	wtn	wtn-btn	none	wtn	wtn-btn
KP under id.		25.42	24.53		22.08	21.69
KP weak id.		70.37	64.49		56.4	45.65

Multi-level clustered standard errors by city-industry, city-year, and industry-year in parenthesis. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Heterogeneous *within* terms, city-time and industry-time effects are included in all regressions but not displayed. 2SLS regressions use lagged instruments. Note that the number of observations falls for the instrumented regressions because the instruments require a lagged employment term. Thus, data from 1851 are not available for these regressions. Acronyms: wtn = *within*, btn = *between*. "KP under id." denotes the test statistic for the Lagrange Multiplier underidentification test based on Kleibergen & Paap (2006). "KP weak id." denotes the test statistic for a weak instruments test based on the Kleibergen-Paap Wald statistic.

Based on the results from Column 6 of Table 2, our preferred specification, a one standard deviation increase (3.21) in the presence of local suppliers (the *IOin* channel) increases city-industry growth by 20%. Turning to the occupational similarity channel, *OCC*, a one standard deviation increase in the presence of occupationally-similar local industries (25.70) leads to a 17% increase in city industry growth when using the results from Column 6 of Table 2. Thus, both of these channels appear to exert a substantial positive effect on city-industry growth.

Table 2: Results with all cross-industry spillover channels

	(1)	(2)	(3)	(4)	(5)	(6)
IOin	0.0216***	0.0187***	0.0161**	0.0758***	0.0623***	0.0622***
	(0.0072)	(0.0067)	(0.0069)	(0.0170)	(0.0158)	(0.0159)
IOout	0.0043	0.0004	-0.0007	-0.0051	-0.0144	-0.0178
	(0.0047)	(0.0051)	(0.0050)	(0.0095)	(0.0105)	(0.0109)
EMP	0.0002	0.0001	-0.0001	0.0001	0.0020*	0.0016
	(0.0005)	(0.0005)	(0.0005)	(0.0016)	(0.0012)	(0.0014)
OCC	0.0020	0.0013	0.0012	0.0087***	0.0070**	0.0068**
	(0.0013)	(0.0013)	(0.0013)	(0.0029)	(0.0032)	(0.0032)
Within	-0.0316***	-0.0225*	-0.0211*			
	(0.0114)	(0.0122)	(0.0124)			
Observations	4,263	3,554	3,549	4,263	3,549	3,549
Estimation	ols	2sls	2sls	ols	2sls	2sls
Instrumented	none	wtn	wtn-btn	none	wtn	wtn-btn
Within terms	homog	homog	homog	heter	heter	heter
KP under id.		22.79	23.99		28.05	29.87
KP weak id.		5508.28	1082.03		76.09	50.72

Multi-level clustered standard errors by city-industry, city-year, and industry-year in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Heterogeneous regressors *within* are included in Columns 4-6 but not displayed. City-time and industry-time effects are included in all regressions but not displayed. 2SLS regressions use lagged instruments. Note that the number of observations falls for the instrumented regressions because the instruments require a lagged employment term. Thus, data from 1851 are not available for these regressions. Acronyms: wtn = *within*, btn = *between*. "KP under id." denotes the test statistic for the Lagrange Multiplier underidentification test based on Kleibergen & Paap (2006). "KP weak id." denotes the test statistic for a weak instruments test based on the Kleibergen-Paap Wald statistic.

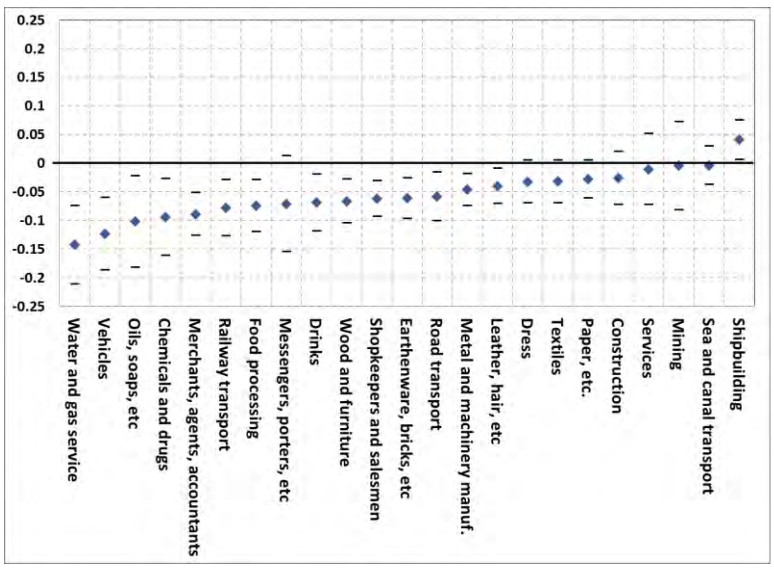
Our analysis can also help us understand the strength of within-industry spillovers, reflected in the $\ln(L_{ict})$ term in Equation 17.[37] When analyzing these results, it is important to keep in mind that they reflect the *net* effect of within-industry agglomeration forces, which may be generated through a balance between agglomeration forces and negative forces such as competition or mean-reversion due to the diffusion of technologies across cities. We cannot identify the strength of local within-industry agglomeration forces independent of counteracting forces. However, it is the net strength of these forces, which we are able to estimate, that is relevant for understanding the contribution of within-industry agglomeration forces to city growth.

We have already seen, in Table 2 Columns 1-3, that the average within-industry effect across all industries is negative. These results are consistent with negative

[37] In a static context these are often referred to as localization economies.

dynamic within-industry effects, perhaps linked to the unwillingness of firms to share new ideas with their direct competitors. However, the fact that our results change substantially once we allow for heterogeneous within-industry effects, as in Columns 4-6 of Table 2, suggests that these are likely to vary substantially across industries. We explore these heterogeneous within-industry effects in Figure 2, which presents coefficients and 95% confidence intervals for industry-specific within-industry spillover coefficients from regressions corresponding to Column 6 of Table 2.

Figure 2: Strength of within-industry effects by industry



Results correspond to the regression described in Column 6 of Table 2. This figure displays coefficient estimates and 95% confidence intervals based on standard errors clustered by city-industry, city-year, and industry-year. The regression includes a full set of city-year and industry-year effects as well as *between* terms. Both the within and between terms are instrumented using one-decade lags.

In only one industry, shipbuilding, do we observe any evidence of positive within-industry effects. This industry was characterized by increasing returns and strong patterns of geographic concentration. All other industries exhibit slower growth in locations where initial industry employment was large, after controlling for other forces. Within-industry agglomeration benefits, it would appear, are more the exception than the rule.

The results presented so far describe coefficients generated using all industries, where each industry is given equal weight. We may be concerned that these results are being driven primarily by smaller industries or smaller cities. To check this, we have also calculate weighted regressions, where the set of observations for each city-industry is weighted based on employment in that city-industry in 1851.[38] Note that this puts a lot of weight on the effects observed in a few very large cities. The results are presented in Table 3. These weighted regressions continue to highlight the important role played by local suppliers. Thus, this result is not driven by smaller industries or cities. However, we no longer observe positive results associated with the occupational similarity measure. This suggests that the positive impact of local industries employing similar workers observed in Table 2 is being driven by smaller industries or smaller cities, an interesting result in itself.

Table 3: Weighted regression results with all cross-industry spillover channels

	(1)	(2)	(3)	(4)	(5)	(6)
IOin	0.0161	0.0224**	0.0259***	0.0299**	0.0314***	0.0361***
	(0.0105)	(0.0096)	(0.0100)	(0.0137)	(0.0120)	(0.0126)
IOout	-0.0039	-0.0082	-0.0086	-0.0038	-0.0111	-0.0110
	(0.0057)	(0.0065)	(0.0065)	(0.0130)	(0.0138)	(0.0143)
EMP	0.0003	0.0002	0.0003	0.0002	0.0008	0.0007
	(0.0003)	(0.0003)	(0.0003)	(0.0009)	(0.0008)	(0.0009)
OCC	-0.0007	-0.0006	-0.0002	-0.0023	-0.0023	-0.0015
	(0.0015)	(0.0015)	(0.0015)	(0.0028)	(0.0029)	(0.0028)
Within	-0.0127	-0.0120	-0.0122			
	(0.0111)	(0.0116)	(0.0115)			
Observations	4,253	3,544	3,541	4,253	3,541	3,541
Estimation	ols	2sls	2sls	ols	2sls	2sls
Instrumented	none	wtn	wtn-btn	none	wtn	wtn-btn
wtn	homog	homog	homog	heter	heter	heter
KP under id.		23	24.21		28.03	29.82
KP weak id.		5261.37	1026.95		75.83	50.28

Multi-level clustered standard errors by city-industry, city-year, and industry-year in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. Heterogeneous regressors *within* are included in Columns 4-6 but not displayed. City-year and industry-year effects are included in all regressions but not displayed. 2SLS regressions use lagged instruments. Note that the number of observations falls for the instrumented regressions in columns 3-6 because the instruments require a lagged employment term. Thus, data from 1851 are not available for these regressions. Acronyms: wtn = *within*, btn = *between*. "KP under id." denotes the test statistic for the Lagrange Multiplier underidentification test based on Kleibergen & Paap (2006). "KP weak id." denotes the test statistic for a weak instruments test based on the Kleibergen-Paap Wald statistic. Weights for each city-industry observation are based on employment in the city-industry in 1851.

[38] Specifically, this is done by weighting the importance of each city-industry observation based on the number of workers it represented in 1851. We base the weights for all years on initial city-industry employment to avoid the potential for endogeneity in the city-industry weights.

We have also investigated the robustness of our results to dropping individual industries or individual cities from the analysis database (see Appendix A.4.1). These exercises show that the significance of the estimates on the *IOin* and *OCC* channels are robust to dropping any city or any industry. However, the estimated coefficient and confidence levels for the *IOout* coefficient is sensitive to the exclusion of particular industries. Specifically, when shipbuilding is excluded we observe that the *IOout* coefficient becomes positive but not statistically significant.[39] This suggests that in general the presence of local buyers may have a mild positive effect on industry growth.

In addition, we have explored the sensitivity of our results to using alternative functional forms to represent the relationship between spillovers and technological progress. In Appendix A.4.3 we present results using alternative concave functional relationships such as a square root or fifth root. Our findings are not sensitive to these alternatives.

We have also explored the robustness of our results to the use of alternative connections matrices. In particular, in Appendix A.4.4 we present results obtained while using the less detailed input-output table constructed by Horrell *et al.* (1994), which covers 12 more aggregated industry categories in 1841. This alternative input-output matrix delivers similar results to those shown in our main regression tables.

The results discussed so far reveal average patterns across all industries. An additional advantage of our empirical approach is that it is also possible to estimate industry-specific coefficients in order to look for (1) heterogeneity in the industries that benefit from each type of inter-industry connection or (2) heterogeneity in the industries that produce each type of inter-industry connections. In Appendix A.4.2, we estimate industry-specific coefficients for both spillover-benefiting and spillover-producing industries and then compare them to a set of available industry characteristics such as firm size, export and final goods sales shares, and labor or intermediate cost shares. With only 23 estimated industry coefficients we cannot draw strong conclusions from these relationships. However, our results do suggest several interesting patterns. The only clear result is that industries that benefit from or produce spillovers through the OCC channel tend to have a higher labor cost to sales ratio, a finding that seems very reasonable. We also observe a consistent negative relationship

[39] Shipbuilding stands out relative to the other industries because it is particularly reliant on local geography.

between firm size and all types of inter-industry connections. While this relationship is not statistically significant, it is consistent across all spillover types and it fits well with previous work highlighting the importance of inter-industry connections for smaller firms (e.g., Chinitz (1961)).

We can also look at how the estimated industry-specific within-industry coefficients are related to industry characteristics. This is done in Appendix A.4.2. With such a small number of industry coefficients we cannot draw strong conclusions from these results. However, we do observe some evidence that within-industry connections are more important in industries with larger firm sizes, which contrasts with the consistent negative relationship that we observe between firm size and cross-industry spillovers.

While the analysis described above focuses on spillovers occurring within-cities, we have also explored the possibility that there may be important cross-city effects. To explore cross-city effects, we have run additional regressions including variables measuring market size as well as cross-industry spillovers occurring across cities. Our results, reported in Appendix A.4.5, suggest that cross-city effects are much weaker than within-city forces. This makes sense given that we think that the shape of cities reflects the rapidly decaying strength of local agglomeration forces. We also find that accounting for cross-city effects has little impact on our estimates of the strength of within-city agglomeration forces.

6 Strength of the agglomeration forces

In this section we examine the relationship between city size and city-industry growth and show how our city-year effects can be used to construct a summary measure of the aggregate strength of the many cross-industry agglomeration forces present in our model. In standard urban models, the impact of agglomeration forces is balanced by congestion forces related to city size, operating through channels such as higher housing prices or greater commute times. In our model, we have been largely agnostic about the form of the congestion forces, which will be captured primarily by the city-time effects. Thus, examining these estimated city-time coefficients offers an opportunity for assessing the *net* impact of dynamic congestion or agglomeration force related to overall city size.[40] Also, the difference between these estimated city-time effects and actual city growth rates must be due to the impact of the agglomeration forces in the estimation equation. As a result, comparing the estimated city-time effects to actual city growth rates allows us to quantify the combined strength of the many cross-industry agglomeration forces captured by our measures.

To make this comparison more concrete, consider the graphs in Figure 3. The dark blue diamond symbols in each graph describe, for each decade starting in 1861, the relationship between the actual growth rate of city working population and the log of city population at the beginning of the decade. The slopes of the fitted lines for these series fluctuate close to zero, suggesting that on average Gibrat's Law holds for the cities in our data.

We want to compare the relationship between city size and city growth in the actual data, as shown by the dark blue diamonds in Figure 3, to the relationship between these variables obtained while controlling for within and cross-industry agglomeration forces. This can be done using the estimated city-time effects represented by θ_{ct} in Eq. 18. The red squares in Figure 3 describe the relationship between the estimated city-year coefficients for each decade, θ_{ct}, and the log of city population at the beginning of each decade. In essence, these are showing us the relationship between city size and city growth after controlling for national industry growth trends and the agglomeration forces included in our model. We can draw two lessons from these graphs. First, in all years the fitted lines based on the θ_{ct} terms slope downward more steeply than the fitted lines for actual city growth. This suggests that, once we control for cross-industry agglomeration forces, city size is negatively related to city growth, consistent with the idea that there are dynamic city-size congestion forces. Second, the difference between the slopes of the two fitted lines can be interpreted as the aggregate effect of the various agglomeration forces in our model averaged across cities. Put simply, if we can add up the strength of the convergence force in any period and compare it to the actual pattern of city growth, then the difference must be equal to the strength of the agglomeration forces. Third, the patterns described in Figure 3 appear to be close to linear in logs, suggesting that these forces do not differ dramatically across different city sizes.

The strength of these effects can be quantified in terms of the implied convergence

[40]These results will reflect only the net impact of city size, including both congestion and agglomeration forces.

rate following the approach of Barro & Sala-i Martin (1992). To do so, we run the following regressions:

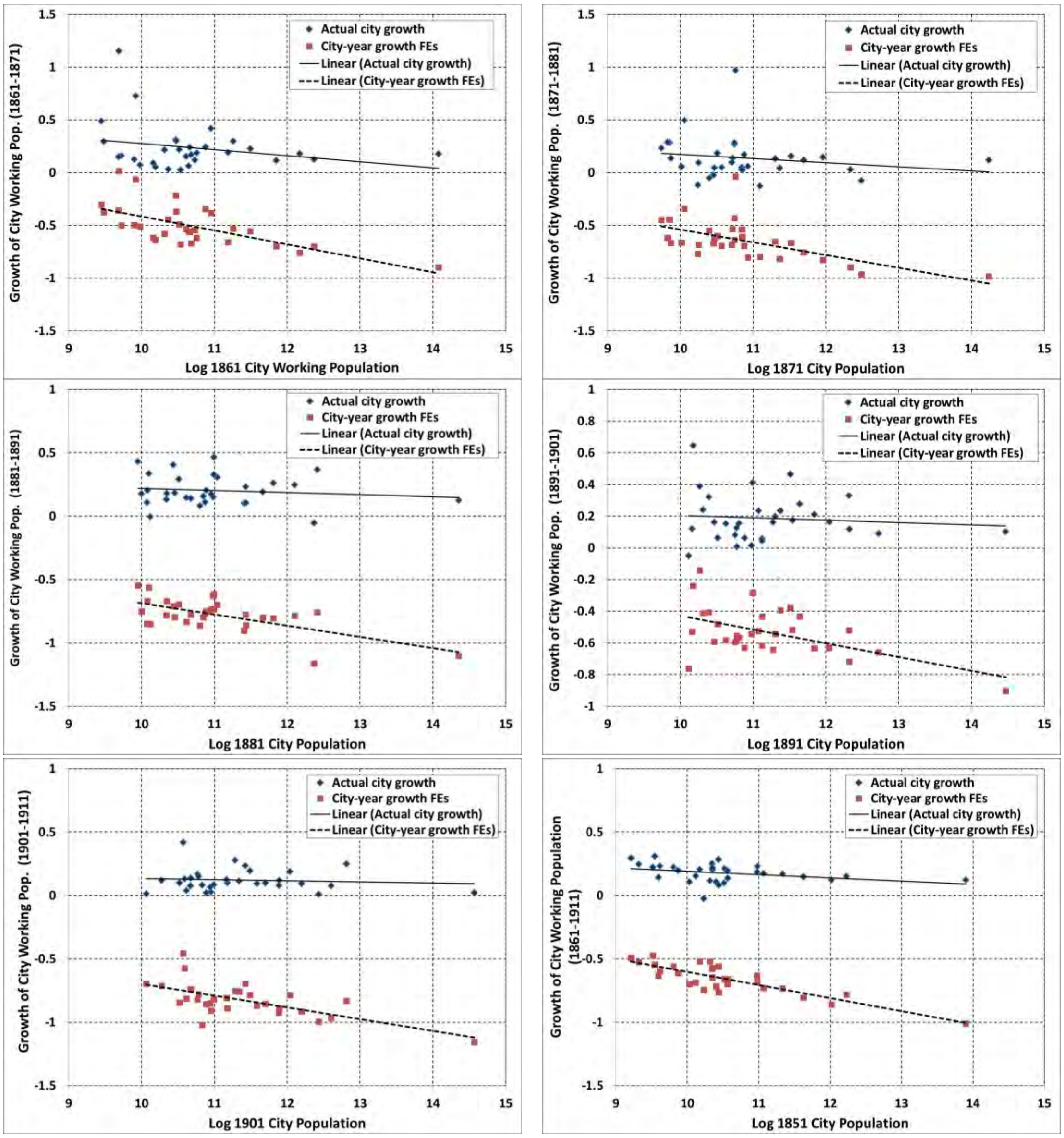
$$\theta_{ct} = a_0 + a_1 \ln(WORKpop_{ct}) + f_{ct} \tag{19}$$

$$GrowthWORKpop_{ct} = b_0 + b_1 \ln(WORKpop_{ct}) + f_{ct} \tag{20}$$

where θ_{ct} is the estimated city-time effect for the decade from t to $t+1$, $WORKpop_{ct}$ is the working population of the city in year t, and $GrowthWORKpop_{ct}$ is the actual growth rate of the city from t to $t+1$. These regressions are run separately for each decade from 1861 to 1911. Convergence rates can be calculated using the estimated a_1 and b_1 coefficients.

The results are presented in the top panel of Table 4. The two left-hand columns describe the results from Equation 19 and the annualized city-size divergence rate implied by these estimates. The next two columns describe similar results based on Equation 20. The difference between these two city-size divergence rates, given in the right-hand column, describes the aggregate strength of the agglomeration force reflected in the cross-industry terms. These results suggest that the strength of city agglomeration forces, in terms of the implied divergence rate, was 7.5-8.9% per decade. In the bottom panel of Table 4 we calculate similar results except that the θ_{ct} terms are obtained using regressions in which each observation is weighted based on the employment in each city-industry in 1851. These results suggest a weaker agglomeration force, equal to an implied divergence rate of 1.0-2.3% per decade. The difference between these two results suggests that the agglomeration forces we capture may have played a more important role for small industries.

Figure 3: City size and city growth



Solid lines: Fitted lines comparing actual city growth over a decade to the log of city size at the beginning of the decade. **Dotted lines:** Fitted lines comparing estimated coefficients from city-time effects for each decade to the log of city size at the beginning of the decade. **Blue diamonds:** Plot the actual city growth over a decade against the log of city population at the beginning of the decade. **Red squares:** Plot the estimated city-time coefficients over the same decade (the θ_{ct} terms estimated using Eq. 18) against the log of city population at the beginning of the decade. The bottom right-hand panel compares the log of city population in 1851 to the average of city growth rates over the entire 1861-1911 period and the average of city-time fixed effects across the entire 1861-1911 period.

Table 4: Measuring the aggregate strength of the agglomeration forces

Results based on unweighted regressions

	Results based on θct		Results for actual city growth		Aggregate strength of agglomeration force (implied divergence rate per decade)
	Estimated city-size coefficient	Implied divergence Beta	Estimated city-size coefficient	Implied divergence Beta	
1861-1871	-0.133	14.27%	-0.058	6.00%	8.28%
1871-1881	-0.121	12.90%	-0.039	4.02%	8.88%
1881-1891	-0.089	9.31%	-0.017	1.66%	7.65%
1891-1901	-0.087	9.08%	-0.015	1.50%	7.58%
1901-1911	-0.093	9.72%	-0.009	0.90%	8.82%

Results based on regressions weighted by city-industry size in 1851

	Results based on θct		Results for actual city growth		Aggregate strength of agglomeration force (implied divergence rate per decade)
	Estimated city-size coefficient	Implied divergence Beta	Estimated city-size coefficient	Implied divergence Beta	
1861-1871	-0.068	7.06%	-0.058	6.00%	1.07%
1871-1881	-0.059	6.03%	-0.039	4.02%	2.01%
1881-1891	-0.039	3.99%	-0.017	1.66%	2.32%
1891-1901	-0.034	3.42%	-0.015	1.50%	1.92%
1901-1911	-0.024	2.40%	-0.009	0.90%	1.50%

Column 1 presents the a_1 coefficients from estimating Equation 19 for each decade (cross-sectional regressions). Column 2 presents the decadal convergence rates implied by these coefficients. Column 3 presents the b_1 coefficients from estimating Equation 20 and Column 4 presents the decadal divergence rates implied by these coefficients. Column 5 gives the aggregate strength of the divergence force due to the agglomeration economies, which is equal to the difference between the decadal divergence coefficients in Columns 2 and 4. The results in the top panel are based on city-time effects estimated from unweighted regressions while the results in the bottom panel are based on city-time effects estimated using weighted regressions based on city-industry employment in 1851.

There are some caveats to keep in mind when assessing these results. First, there are likely to be agglomeration forces not captured by our estimation. These omitted agglomeration forces may be partially reflected in the city-year fixed effects, which would lead us to understate the strength of the agglomeration forces. Second, some congestion forces may also be captured by our cross-industry terms. Similarly, there may be some agglomeration forces captured by the within-industry terms, which will also not be reflected in our results. Thus, the strength of the cross-industry agglomeration force measured here is likely to be a lower bound.

We may be concerned that the results described in Table 4 are driven in part by the inclusion of industry-time effects in the regressions used to obtain the θ_{ct} terms. We explore this possibility in Appendix A.4.6 by comparing the relationship between our estimated θ_{ct} terms and city-time effects estimated while controlling for industry-year effects. Since the only difference between these specifications is the presence of the within and cross-industry agglomeration forces, we can be sure that these are driving any differential results. Estimates obtained using this method are very similar, but slightly larger, than those described in Table 4.

We can use a similar exercise to estimate the aggregate strength of the convergence force due to within-industry effects. We begin by estimating,

$$\Delta \ln(L_{ict+1}) = \tilde{\gamma}_{ii} \ln(L_{ict}) + \theta_{ct}^{WITHIN} + \chi_{it} + \epsilon_{ict}. \tag{21}$$

Next, we use the values of θ_{ct}^{WITHIN} to estimate,

$$\theta_{ct}^{WITHIN} = d_0 + d_1 \ln(WORKpop_{ct}) + f_{ct}. \tag{22}$$

We then calculate the convergence force associated with the within-industry terms using the same approach that we used previously. Table 5 describes the results. The negative measured divergence force in this table highlights that within-industry effects, on net, act as a convergence force. The strength of this force is sensitive to whether the regressions are weighted, which suggests that the negative within-industry employment effects tend to be stronger for smaller industries.

Table 5: Measuring the aggregate strength of the convergence force associated with the within-industry effects

Results based on unweighted regressions

	Results based on θ^{WITHIN}		Results for actual city growtl		Difference: aggregate strength of agglomeration force (implied divergence
	Estimated city-size coefficient	Implied divergence Beta	Estimated city-size coefficient	Implied divergence Beta	
1861-1871	-0.004	0.04%	-0.058	0.60%	-0.56%
1871-1881	0.006	-0.06%	-0.039	0.40%	-0.47%
1881-1891	0.037	-0.36%	-0.017	0.17%	-0.53%
1891-1901	0.039	-0.39%	-0.015	0.15%	-0.54%
1901-1911	0.035	-0.34%	-0.009	0.09%	-0.43%

Results based on regressions weighted by city-industry size in 1851

	Results based on θ^{WITHIN}		Results for actual city growtl		Difference: aggregate strength of agglomeration force (implied divergence
	Estimated city-size coefficient	Implied divergence Beta	Estimated city-size coefficient	Implied divergence Beta	
1861-1871	-0.042	0.43%	-0.058	0.60%	-0.17%
1871-1881	-0.032	0.32%	-0.039	0.40%	-0.08%
1881-1891	-0.013	0.13%	-0.017	0.17%	-0.04%
1891-1901	-0.008	0.08%	-0.015	0.15%	-0.07%
1901-1911	0.003	-0.03%	-0.009	0.09%	-0.12%

Column 1 presents the d_1 coefficients from estimating Equation 22 for each decade (cross-sectional regressions). Column 2 presents the decadal divergence rates implied by these coefficients. Column 3 presents the b_1 coefficients from estimating Equation 20 and Column 4 presents the decadal divergence rates implied by these coefficients. Column 5 gives the aggregate strength of the divergence force due to the agglomeration economies, which is equal to the difference between the decadal convergence coefficients. The negative values in Column 5 indicate that within-industry effects are, on net, a source of convergence across cities. The results in the top panel are based on city-time effects estimated from unweighted regressions while the results in the bottom panel are based on city-time effects estimated using weighted regressions based on city-industry employment in 1851.

7 Conclusion

In the introduction, we posed a number of questions about the nature of localized agglomeration forces. The main contribution of this study is to provide a theoretically grounded empirical approach that can be used to address these questions and the detailed city-industry panel data needed to implement it.

We can now provide some answers for the particular empirical setting that we study. First, we find evidence that cross-industry agglomeration economies were more important than within-industry agglomeration forces for generating city employment growth. Within-industry effects are, on net, generally negative. This suggests that local clusters of firms working in the same industry, which have attracted substantial attention, are unlikely to deliver dynamic benefits. Second, our results suggest that industries grow more rapidly when they co-locate with their suppliers or with other industries that use occupationally-similar workforces. This result is in line with arguments made by Jacobs (1969), as well as recent empirical findings. We document a clear negative relationship between city size and city growth that appears once we account for agglomeration forces related to a city's industrial composition. This suggests that Gibrat's law is generated by a balance between agglomeration and dispersion forces. A lower bound estimate of the overall strength of the agglomeration forces captured by our approach, in terms of the implied annual divergence rate in city size, is around 1.0-2.3% per decade, though we find evidence that the effect on smaller industries and smaller cities is likely to be substantially larger.

One of the most striking features of our results is how similar they look to some of the existing findings in the literature, most of which are based on modern U.S. or European data. In particular, the ordering of importance for the different spillover channels – with input-output paths showing the strongest effects, followed by occupational similarity – looks very similar to the ordering obtained by Ellison *et al.* (2010). This provides suggestive evidence that there may be substantial persistence in the importance of these agglomeration economies over time and across space. Understanding how the patterns of within-industry and inter-industry connections evolve over time is one avenue for future research.

The techniques introduced in this paper can be applied in any setting where sufficiently rich long-run city-industry panel data can be constructed. Recent work has made progress in constructing data of this type for the U.S. in both the modern and historical period. Applying our approach to these emerging data sets is another promising avenue for future work.

References

Albouy, David. 2009. The Unequal Geographic Burden of Federal Taxation. *Journal of Political Economy*, **117**(4), pp. 635–667.

Amiti, Mary, & Cameron, Lisa. 2007. Economic Geography and Wages. *Review of Economics and Statistics*, **89**(1), pp. 15–29.

Arellano, Manuel, & Bond, Stephen. 1991. Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, **58**(2), 277–297.

Baines, Dudley. 1994. Population, Migration and Regional Development, 1870-1939. *In:* Floud, R, & McCloskey, D (eds), *The Economic History of Britain Since 1700*, second edition edn. Cambridge University Press.

Balsvik, Ragnhild. 2011. Is labor mobility a channel for spillovers from multinationals? Evidence from Norwegian manufacturing. *Review of Economics and Statistics*, **93**(1), 285–297.

Barrios, Thomas, Diamond, Rebecca, Imbens, Guido, & Kolesar, Michal and. 2012. Clustering, Spatial Correlations and Randomized Inference. *Journal of the American Statistical Association*, **107**(498), 578–591.

Barro, Robert J. 2012 (August). *Convergence and Modernization Revisited.* NBER Working Paper No. 18295.

Barro, Robert J., & Sala-i Martin, Xavier. 1992. Convergence. *Journal of Political Economy*, **100**(2), pp. 223–251.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Bertrand, M, Duflo, E, & Mullainathan, S. 2004. How Much Should We Trust Differences-in-Differences Estimates? *The Quarterly Journal of Economics*, **119**(1), pp. 249–275.

Bleakley, Hoyt, & Lin, Jeffrey. 2012. Portage: Path Dependence and Increasing Returns in U.S. History. *Quarterly Journal of Economics*, **127**(May), 587–644.

Blundell, Richard, & Bond, Stephen. 1998. Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics*, **87**(1), 115 – 143.

Cameron, A. C., Gelbach, J.B., & Milller, D. L. 2011. Robust Inference with Multi-Way Clustering. *Journal of Business and Economic Statistics*, **29**(2), 238–249.

Caselli, Francesco, Esquivel, Gerardo, & Lefort, Fernando. 1996. Reopening the Convergence Debate: A New Look at Cross-Country Growth Empirics. *Journal of Economic Growth*, **1**(3), pp. 363–389.

Chay, Kenneth Y., & Greenstone, Michael. 2005. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, **113**(2), pp. 376–424.

Chinitz, Benjamin. 1961. Contrasts in Agglomeration: New York and Pittsburgh. *The American Economic Review*, **51**(2), pp. 279–289.

Combes, Pierre-Philippe. 2000. Economic Structure and Local Growth: France, 1984-1993. *Journal of Urban Economics*, **47**(3), 329 – 355.

Combes, Pierre-Philippe, & Gobillon, Laurent. 2015. The Empirics of Agglomeration Economies. *Pages pp. 247–348 of:* Duranton, Gilles, Henderson, Vernon, & Strange, William (eds), *Handbook* *of Urban and Regional Economics*, vol. 5. Elsevier.

Combes, Pierre-Philippe, Duranton, Gilles, & Gobillon, Laurent. 2008. Spatial Wage Disparities: Sorting Matters!. *Journal of Urban Economics*, **63**(2), 723 – 742.

Combes, Pierre-Philippe, Duranton, Gilles, & Gobillon, Laurent. 2011. The Identification of Agglomeration Economies. *Journal of Economic Geography*, **11**(2), 253 – 266.

Combes, Pierre-Philippe, Duranton, Gilles, Gobillon, Laurent, Puga, Diego, & Roux, Sebastien. 2012. The Productivity Advantages of Large Cities: Distinguishing Agglomeration from Firm Selection. *Econometrica*, **80**(6), pp. 2543–2594.

Conley, T.G. 1999. GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics*, **92**(1), 1 – 45.

Crafts, Nicholas, & Mulatu, Abay. 2006. How Did the Location of Industry Respond to Falling Transport Costs in Britain before World War I? *The Journal of Economic History*, **66**(3), pp. 575–607.

Davis, Donald R, & Weinstein, David E. 2002. Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, **92**(5), 1269–1289.

Dekle, Robert. 2002. Industrial Concentration and Regional Growth: Evidence from the Prefectures. *The Review of Economics and Statistics*, **84**(2), pp. 310–315.

Desmet, Klaus, & Rappaport, Jordan. Forthcoming. The Settlement of the United States, 1800-2000: The Long Transition Towards Gibrat's Law. *Journal of Urban Economics*.

Desmet, Klaus, & Rossi-Hansberg, Esteban. 2014. Spatial Development. *American Economic Review*, **104**(4), 1211–43.

Dittmar, Jeremiah E. 2011 (August). *Cities, Markets, and Growth: The Emergence of Zipf's Law*. Working Paper.

Dix-Carneiro, Rafael, & Kovak, Brian. 2015 (January). *Trade Reform and Regional Dynamics: Evidence from 25 Years of Brazilian Matched Employer-Employee Data*. NBER Working Paper No. 20908.

Dumais, Guy, Ellison, Glenn, & Glaeser, Edward L. 1997 (November). *Geographic Concentration as a Dynamic Process*. NBER Working Paper No. 6720.

Dumais, Guy, Ellison, Glenn, & Glaeser, Edward L. 2002. Geographic Concentration as a Dynamic Process. *The Review of Economics and Statistics*, **84**(2), pp. 193–204.

Duranton, Gilles, Morrow, Peter M., & Turner, Matthew A. 2014. Roads and Trade: Evidence from the US. *The Review of Economic Studies*, **81**(2), 681–724.

Ellison, G., Glaeser, E., & Kerr, W. 2010. What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*, **100**(3), pp. 1195–1213.

Ellison, Glenn, & Glaeser, Edward L. 1997. Geographic concentration in US manufacturing industries: A dartboard approach. *Journal of Political Economy*, **105**(5), 889–927.

Faggio, Guilia, Silva, Olmo, & Strange, William. 2013 (October). *Heterogeneous Agglomeration*. Mimeo.

Glaeser, Edward L., & Mar, David C. 2001. Cities and Skills. *Journal of Labor Economics*, **19**(2), pp. 316–342.

Glaeser, Edward L, Kallal, Hedi D, Scheinkman, Jose A, & Shleifer, Andrei. 1992. Growth in Cities.

Journal of Political Economy, **100**(6), 1126–1152.

Greenstone, Michael, & Moretti, Enrico. 2003 (July). *Bidding for Industrial Plants: Does Winning a 'Million Dollar Plant' Increase Welfare?* NBER Working Paper No. 9844.

Greenstone, Michael, Hornbeck, Richard, & Moretti, Enrico. 2010. Identifying Agglomeration Spillovers: Evidence from Million Dollar Plants. *Journal of Political Economy*, **118**(3), pp. 536–598.

Gyourko, Joseph, Saiz, Albert, & Summers, Anita. 2008. A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index. *Urban Studies*, **45**(3), 693–729.

Hanson, Gordon H. 2005. Market potential, increasing returns and geographic concentration. *Journal of International Economics*, **67**(1), 1 – 24.

Hauk Jr., William R., & Wacziarg, Romain. 2009. A Monte Carlo Study of Growth Regressions. *Journal of Economic Growth*, **14**(2), pp. 103–147.

Henderson, J. Vernon. 1996. Effects of Air Quality Regulation. *The American Economic Review*, **86**(4), pp. 789–813.

Henderson, J Vernon. 1997. Externalities and Industrial Development. *Journal of Urban Economics*, **42**(3), 449–470.

Henderson, V, Kuncoro, A, & Turner, M. 1995. Industrial Development in Cities. *Journal of Political Economy*, **103**(5), 1067–1090.

Horrell, Sara, Humphries, Jane, & Weale, Martin. 1994. An Input-Output Table for 1841. *Economic History Review*, **47**(3), 545–566.

Hurwicz, Leonid. 1950. Least-Squares Bias in Time Series. *In:* Koopmans, Tjalling C. (ed), *Statistical Inference in Dynamic Economic Models*. New York: Wiley.

Jacks, David S., Meissner, Christopher M., & Novy, Dennis. 2008. Trade Costs, 1870-2000. *The American Economic Review*, **98**(2), pp. 529–534.

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

Javorcik, Beata S. 2004. Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages. *The American Economic Review*, **94**(3), 605–627.

Jones, Ronald W. 1975. Income Distribution and Effective Protection in a Multicommodity Trade Model. *Journal of Economic Theory*, **11**(1), 1 – 15.

Kim, Sukkoo. 1995. Expansion of Markets and the Geographic Distribution of Economic Activities - The Trends in US Regional Manufacturing Structure, 1860-1987. *Quarterly Journal of Economics* , **110**(4), 881–908.

Kleibergen, Frank, & Paap, Richard. 2006. Generalized Reduced Rank Tests Using the Singular Value Decomposition. *Journal of Econometrics*, **133**, 97–126.

Kleibergen, F. ans Schaffer, M.E. 2010. *Rankest: Module for Testing the Rank of a Matrix using the Kleibergen-Paap rk Statistic.* Available at: http://ideas.repec.org/c/boc/bocode/s456865.html.

Kline, Patrick, & Moretti, Enrico. 2013. Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics.*

Kovak, Brian K. 2013. Regional Effects of Trade Reform: What is the Correct Measure of Liberalization? *American Economic Review*, **103**(5), 1960–1976.

Krugman, Paul. 1991. Increasing Returns and Economic Geography. *Journal of Political Economy*, **99**(3), pp. 483–499.

Kugler, Maurice. 2006. Spillovers from Foreign Direct Investment: Within or Between Industries? *Journal of Development Economics*, **80**(2), 444–477.

Lee, Jamie. 2015 (January). *Measuring Agglomeration: Products, People and Ideas in U.S. Manufacturing, 1880-1990.* Mimeo.

Long, Jason, & Ferrie, Joseph. 2004 (March). *Geographic and Occupational Mobility in Britain and the U.S., 1850-1881.* Working Paper.

Lopez, R, & Sudekum, J. 2009. Vertical Industry Relations, Spillovers, and Productivity: Evidence from Chilean Plants. *Journal of Regional Science*, **49**(4), pp. 721–747.

Lucas, Robert E. 1988. On the Mechanics of Economic Development. *Journal of Monetary Economics*, **22**(1), 3–42.

Marshall, Alfred. 1890. *Principles of Economics.* New York: Macmillan and Co.

Michaels, Guy, Rauch, Ferdinand, & Redding, Stephen. 2013 (January). *Task Specialization in U.S. Cities from 1880-2000.* NBER Working Paper No. 18715.

Nickell, Stephen. 1981. Biases in Dynamic Models with Fixed Effects. *Econometrica*, **49**(6), pp. 1417–1426.

Poole, Jennifer P. 2013. Knowledge Transfers from Multinational to Domestic Firms: Evidence from Worker Mobility. *The Review of Economics and Statistics*, **95**(2).

Romer, PM. 1986. Increasing Returns and Long-run Growth. *Journal of Political Economy*, **94**(5), 1002–1037.

Rosenthal, S., & Strange, W. 2004. Evidence on the Nature and Sources of Agglomeration Economies. *In:* Henderson, JV., & Thisse, JF. (eds), *Handbook of Regional and Urban Economics.* Elsevier.

Rosenthal, Stuart S, & Strange, William C. 2001. The Determinants of Agglomeration. *Journal of Urban Economics* , **50**(2), 191–229.

Rosenthal, Stuart S., & Strange, William C. 2003. Geography, Industrial Organization, and Agglomeration. *The Review of Economics and Statistics*, **85**(2), pp. 377–393.

Thomas, Mark. 1987. *An Input-Output Approach to the British Economy, 1890-1914.* Ph.D. thesis, Oxford University.

Thompson, Samuel B. 2011. Simple Formulas for Standard Errors that Cluster by both Firm and Time. *Journal of Financial Economics*, **99**(1), 1 – 10.

A Appendix

A.1 Theory appendix

This appendix explores several additional factors that are not included in the model provided in the main text. We begin by discussing the implications of allowing variation in the costs of innovation or firm entry across industries. Next, we consider allowing production function parameters to vary across industries. We then look at incorporating capital into the model. Finally, we consider the implications of including intermediate goods and the closely related issue of incorporating trade costs into the model.

Variation in Industry Innovation or Entry Cost Parameters

Suppose that the innovation costs or entry costs are allowed to vary by industry, so that we have F_i and C_i. This will affect the rate of firm entry and R&D, which will affect the size of city-industry employment. However, as long as these cost parameters are fixed over time and both denominated in the same units (labor), they will be differenced out when we obtain the main regression specification. Thus, our empirical approach will be robust to this modification. Note that holding these parameters fixed over time does not imply that the costs of entry or innovation is fixed over time, since that cost will also depend on the wage, which will vary over time and across locations. However, it does imply that the relative cost of entry and R&D is constant over time and across locations, even if it varies across industries.

Variation in Industry Production Function Parameters

Suppose that we allow the production function parameters to vary by industry, so that they are now all indexed by i. In this case, the coefficient on the spillover term from industry k to industry i, which in our baseline model is τ_{ki}/β is now given by τ_{ki}/β_i. What this tells us is that the way in which employment in industry k is translated into employment growth in industry i will now depend on the importance of local industry-specific resources in the production function. This is because local resources are the factors that tie industries to particular locations. Industries in which local resources are relatively unimportant (low β) should exhibit large estimated spillover coefficients because employment will be more able to respond to changing technology levels by shifting across locations.

When we estimate single cross-industry spillover terms we will be estimating the average impact across industries with potentially varying β parameters. It is also possible to estimate industry-specific spillover benefit terms in our framework, which is done in Appendix A.4.2.

Incorporating Capital

Suppose that we incorporate capital as an input into production, so that the production function is,

$$y_{icft} = a_{icft} L_{icft}^{\alpha_1} H_{icft}^{\alpha_2} K_{icft}^{\alpha_3} R_{icft}^{\beta} E_{icft}^{1-\alpha_1-\alpha_2-\beta}.$$

If there is a national capital market then this implies a national price of capital, which we denote κ_t. In this case, introducing capital into the model would simply add an additional time-varying national factor to the estimating equation. This will be absorbed by the fixed effects and would not affect our results.

Alternatively, it may be the case that capital markets are more local. In this case, the price of capital will be κ_{ct}. When industries share a common set of production function parameters this will affect all industries in a similar way. As a result, it will be absorbed into the city-time effects and will not impact our results.

However, if we also allow industries to be more or less capital intensive, then variation in the local price of capital may have heterogeneous effects on city-industry growth. Under these circumstances, the source of capital will become important. If capital is related to city size, for example because capital availability depends on total local savings which scales with city size, then this will introduce an industry-specific city size effect. It is possible to incorporate industry-specific city size effects into our regression framework. We have experimented with doing so and it does not substantially alter our results. In order for capital to be one channel behind our within and cross-industry spillover estimates, we need three things to be true: (1) capital is local, (2) capital intensity varies across industries, and (3) the accumulation of savings depends on the local composition of industries. If all of these factors are in place, then local capital channels may be a dynamic agglomeration force in our setting. However, capital was fairly mobile across regions in Britain during the period we study, suggesting that local capital accumulation is unlikely to be behind our results, though this may have been a more important factor during the earlier first Industrial Revolution period.

Intermediate Goods and Trade Costs

Next, we discuss the implications of incorporating intermediate goods and trade costs into the theory. We begin by introducing intermediate inputs while maintaining the assumption of free trade. We then consider the implications of allowing non-zero trade costs.

Suppose that each firm uses a basket of intermediate inputs denoted by I_{icft} in production, with a production function parameter ϕ. Let the set of intermediate inputs used in production vary across industries, but for simplicity, we assume that within an industry all firms use these inputs in fixed proportions. Let Z be an input-output matrix with elements z_{ij} such that I_{it} units of intermediate input to industry i require $I_{it}z_{ij}$ units of output from industry j (i.e., the production function for intermediates is Leontief). Total intermediate demand for the output from industry j is then $x^{IO}_{jt} = \sum_i I_{it}z_{ij}$. With costless trade, each industry will face a national intermediate good input price in each period, which we denote d_{it}. The resulting firm optimization problem in period three is,

$$\max_{L_{icft}, R_{icft}} p_{it}a_{icft}L^{\alpha_1}_{icft}H^{\alpha_2}_{icft}I^{\phi}_{icft}R^{\beta}_{icft} - w_{ct}L_{icft} - q_{ct}H_{icft} - d_{it}I_{icft} - r_{ict}R_{icft}$$

with $1 - \alpha_1 - \alpha_2 - \phi - \beta > 0$

With free trade, this will yield a regression specification that is very similar to the one obtained in the main text:

$$
\begin{aligned}
\ln(L_{ict}) - \ln(L_{ict-1}) = {} & \left(\frac{1}{\beta}\right)\left[\tau_{ii}\ln(L_{ict}) + \sum_{k \neq i} \tau_{ki}\ln(L_{kct})\right] \\
& - \phi\left[\ln(d_{it}) - \ln(d_{it-1})\right] \\
& + \left[\ln(p_{it}) - \ln(p_{it-1})\right] + \xi_{it} \\
& - [\nu(\alpha_2 - 1) - \alpha_2]\left[\ln(q_{ct}) - \ln(q_{ct-1})\right] + \psi_{ct} \\
& + (\alpha_2 - 1)\left[\ln(\bar{v}_t^*) - \ln(\bar{v}_{t-1}^*)\right] + \epsilon_{ict}
\end{aligned}
\tag{23}
$$

Thus, under the assumption of free trade across locations, the introduction of intermediate inputs will not impact our results because the impact of changing input prices will be absorbed in the time-varying industry effects.

We can use Eq. 23 to explore the impact of introducing trade costs into the model in a partial equilibrium way. Allowing non-zero trade costs will affect this equation in two ways. First, output prices will vary at the local level, so p_{it} will become p_{ict}. Second, intermediate input prices will also vary locally, so d_{it} becomes d_{ict}. With trade costs, both the input and the output prices faced by firms in industry i can vary across cities.

To consider the impact of trade costs, suppose for now that we turn off all spillover channels, so $S_{ict} = 0$ and,

$$
\begin{aligned}
\ln(L_{ict}) - \ln(L_{ict-1}) = {} & - \phi\left[\ln(d_{it}) - \ln(d_{it-1})\right] \\
& + \left[\ln(p_{it}) - \ln(p_{it-1})\right] \\
& - [\nu(\alpha_2 - 1) - \alpha_2]\left[\ln(q_{ct}) - \ln(q_{ct-1})\right] \\
& + (\alpha_2 - 1)\left[\ln(\bar{v}_t^*) - \ln(\bar{v}_{t-1}^*)\right] + \epsilon_{ict}
\end{aligned}
\tag{24}
$$

Now, focusing on the input prices side, suppose that there are two cities, A and B, and that City A has more industry i suppliers than city B so that the cost of intermediate inputs to industry i is lower in City A than in City B. This implies that employment in industry i will be larger in City A than in City B in some initial period:

this is static agglomeration a la Krugman (1991). A similar effect can be generated through output price channels. However, as we roll the model forward, Equation 24 shows that, absent other changes, industry i will *not* grow faster in City A than in City B. In the absence of other effects, input-output connections alone cannot act as a dynamic agglomeration force.

Where input-output connections can generate dynamic agglomeration patterns is by transmitting the effects of other changes, such as falling transport costs. However, falling trade costs cannot be a sustained force of dynamic agglomeration since trade costs are bounded below by zero. Moreover, trade costs were fairly stable over at least part of the period we study, while urbanization continued apace.[41] This pattern suggests that input-output connections and trade costs can be an important static agglomeration force, but these forces are unlikely to generate the dynamic agglomeration patters studied here.

In a world of static inter-industry agglomeration forces, the growth in industry i must be driven by growth in industry j, rather than the level of industry j. But this raises questions about the causes of the initial growth in industry j. Ultimately, a world of static agglomeration forces is a world of exogenous city-industry growth. In contrast, dynamic agglomeration offers an explanation for city industry growth, just as endogenous growth theory offers an explanation for aggregate growth.

[41]Crafts & Mulatu (2006) conclude that, "falling transport costs had only weak effects on the location of industry in the period 1870 to 1911." Jacks *et al.* (2008) find a rapid fall in external trade costs prior to 1880, with a much slower decline thereafter.

A.2 Data appendix

Table 6: Cities in the primary analysis database

City	Population in 1851	Working population in 1851	Workers in analysis industries in 1851
Bath	54,240	27,623	23,609
Birmingham	232,841	111,992	101,546
Blackburn	46,536	26,211	24,458
Bolton	61,171	31,211	28,885
Bradford	103,778	58,408	55,223
Brighton	69,673	32,949	27,954
Bristol	137,328	64,025	54,962
Derby	40,609	19,299	16,787
Gateshead	25,568	18,058	8,562
Halifax	33,582	18,058	16,488
Huddersfield	30,880	13,922	12,465
Kingston-upon-Hull	84,690	36,983	31,513
Ipswich	32,914	14,660	11,996
Leeds	172,270	83,570	7 4,959
Leicester	60,496	31,140	28,481
Liverpool	375,955	165,300	142,197
London	2,362,236	1,088,285	930,797
Manchester	401,321	204,688	187,000
Newcastle-upon-Tyne	87,784	38,564	33,271
Northampton	26,657	13,626	12,062
Norwich	68,195	34,114	29,710
Nottingham	57,407	33,967	31,106
Oldham	72,357	38,853	35,958
Portsmouth	72,096	31,345	19,039
Preston	69,542	36,864	33,085
Sheffield	135,310	58,551	53,472
South Shields	28,974	11,114	10,028
Southampton	35,305	14,999	12,215
Stockport	53,835	30,128	27,836
Sunderland	63,897	24,779	21,639
Wolverhampton	49,985	22,727	19,851

Figure 4: Map showing the location of cities in the analysis database

Table 7: Industries in the primary analysis database with 1851 employment

Manufacturing		**Services and Professional**	
Chemicals & drugs	18,514	Professionals*	40,733
Clothing, shoes, etc.	328,669	General services	460,885
Instruments & jewelry*	31,048	Merchant, agent, accountant, etc.	58,172
Earthenware & bricks	19,580	Messenger, porter, etc.	72,155
Leather & hair goods	26,737	Shopkeeper, salesmen, etc.	27,232
Metal & Machines	167,052		
Oil, soap, etc.	12,188		
Paper and publishing	42,578	**Transportation services**	
Shipbuilding	14,498	Railway transport	10,699
Textiles	315,646	Road transport	35,207
Vehicles	9,021	Sea & canal transport	66,360
Wood & furniture	69,648		
Food, etc.		**Others industries**	
Food processing	113,610	Construction	137,056
Spiritous drinks, etc.	8,179	Mining	24,505
Tobacconists*	3,224	Water & gas services	3,914

Industries marked with a * are available in the database but are not used in the baseline analysis because they cannot be linked to categories in the 1907 British input-output table.

Table 8: Industry agglomeration patterns based on the Ellison & Glaeser index

Industry	**1851**	**1861**	**1871**	**1881**	**1891**	**1901**	**1911**
Textiles	0.166	0.178	0.182	0.195	0.182	0.159	0.150
Metal & machines	0.094	0.089	0.086	0.093	0.087	0.079	0.088
Shipbuilding	0.081	0.079	0.121	0.142	0.197	0.252	0.254
Paper & publishing	0.076	0.067	0.067	0.049	0.045	0.035	0.025
Mining related	0.075	0.108	0.106	0.151	0.166	0.166	0.165
Professionals	0.070	0.047	0.056	0.050	0.043	0.030	0.036
Sea & canal transport	0.056	0.057	0.058	0.087	0.067	0.080	0.077
Instruments & jewelry	0.053	0.054	0.051	0.050	0.037	0.024	0.023
Oil, soap, etc.	0.028	0.017	0.008	0.007	0.008	0.024	0.042
Road transport	0.027	0.024	0.028	0.012	0.021	0.010	0.010
Leather, hair, etc.	0.025	0.020	0.025	0.026	0.023	0.023	0.039
General services	0.021	0.021	0.017	0.024	0.020	0.025	0.026
Vehicles	0.021	0.012	0.006	0.005	0.033	0.049	0.053
Tobacco	0.020	0.022	0.010	-0.008	-0.008	0.011	0.016
Earthenware & bricks	0.018	0.026	0.032	0.024	0.015	0.009	0.007
Wood & furniture	0.016	0.016	0.022	0.019	0.015	0.009	0.010
Chemicals & Drugs	0.013	0.000	0.007	0.008	-0.004	0.001	0.004
Drinks	0.012	0.011	0.010	0.017	0.008	0.001	-0.002
Food processing	0.009	0.004	0.003	0.001	0.001	0.001	0.001
Clothing, shoes, etc.	0.006	0.007	0.006	0.005	0.008	0.007	0.008
Shopkeepers, salesmen, etc.	0.006	0.001	-0.003	-0.003	-0.004	-0.004	-0.004
Construction	0.004	0.004	0.002	0.003	0.002	0.001	0.002
Merchants, agents, etc.	-0.040	-0.048	-0.041	-0.050	-0.049	-0.054	-0.060
Median	0.021	0.021	0.022	0.019	0.020	0.023	0.023
Mean	0.037	0.035	0.037	0.039	0.040	0.041	0.042

This table reports industry agglomeration in each year based on the index from Ellison & Glaeser (1997). This approach adjusts for the size of plants in an industry using an industry Herfindahl index. We construct these Herfindahl indices using the firm size data reported in the 1851 Census and apply the same Herfindahl for all years, since firm-size data are not reported in later Censuses. This may introduce bias for some industries, such as shipbuilding, where evidence suggests that the average size of firms increased substantially over the study period. Some analysis industries are not included in this table due to lack of firm size data.

Table 9: Industry agglomeration patterns excluding London

Industry	1851	1861	1871	1881	1891	1901	1911
Shipbuilding	0.204	0.216	0.204	0.197	0.208	0.249	0.219
Sea & canal transport	0.196	0.235	0.229	0.268	0.252	0.245	0.211
Instruments & jewelry	0.162	0.181	0.153	0.173	0.101	0.063	0.052
Oil, soap, etc.	0.081	0.053	0.036	0.031	0.033	0.070	0.068
Metal & machines	0.063	0.046	0.040	0.041	0.037	0.033	0.031
Textiles	0.049	0.050	0.048	0.053	0.051	0.053	0.052
Chemicals & Drugs	0.032	0.013	0.018	0.012	0.001	0.014	0.023
Mining related	0.028	0.035	0.019	0.025	0.030	0.031	0.033
Leather, hair, etc.	0.021	0.023	0.034	0.030	0.030	0.027	0.024
Earthenware & bricks	0.020	0.013	0.011	0.011	0.014	0.015	0.014
Vehicles	0.018	0.029	0.018	0.018	0.056	0.061	0.094
Road transport	0.014	0.017	0.010	0.011	0.006	0.003	0.002
Drinks	0.011	0.011	0.012	0.009	0.006	0.004	0.001
Tobacco	0.011	0.005	-0.001	0.021	0.021	0.045	0.078
Clothing, shoes, etc.	0.010	0.008	0.010	0.016	0.026	0.019	0.018
Shopkeepers, salesmen, etc.	0.008	0.003	0.002	0.006	0.002	0.000	-0.003
Food processing	0.007	0.004	0.003	0.002	0.002	0.003	0.003
Wood & furniture	0.007	0.006	0.006	0.004	0.004	0.004	0.004
Paper & publishing	0.006	0.005	0.004	0.005	0.005	0.004	0.006
General services	0.004	0.001	0.000	-0.001	-0.001	-0.001	-0.002
Construction	0.004	0.002	0.002	0.002	0.001	0.001	0.001
Professionals	-0.005	-0.006	-0.006	-0.007	-0.007	-0.008	-0.007
Merchants, agents, etc.	-0.041	-0.044	-0.046	-0.060	-0.065	-0.066	-0.066
Median	0.014	0.013	0.011	0.012	0.014	0.015	0.018
Mean	0.040	0.039	0.035	0.038	0.035	0.038	0.037

This table reports industry agglomeration in each year based on the index from Ellison & Glaeser (1997). This approach adjusts for the size of plants in an industry using an industry Herfindahl index. We construct these Herfindahl indices using the firm size data reported in the 1851 Census and apply the same Herfindahl for all years, since firm-size data are not reported in later Censuses. Some analysis industries are not included in this table due to lack of firm size data.

In addition to the data sets described in the main text, we have collected additional information on a variety of other industry and city characteristics. The 1851 Census of Population was particularly detailed, and provides information on firm sizes in each industry at the national level. From the 1907 input-output table, we have measures of the share of industry output that is sold directly to households, as well as the share exported abroad. The 1907 Census of Production provides us with information on the total wage bill of each industry and the value of output for each industry. These are used to construct, for each industry, estimates of the ratio of labor cost to total sales and, together with the input-output table, the ratio of intermediate cost to total sales. Finally, we collect data on the distance between cities (as the crow flies) from Google Maps, which we will use when considering cross-city effects in Section A.4.5.

Table 10: Summary statistics for the cross-industry spillover terms

	Obs.	Mean	SD	Min	Max
Main analysis matrices and industry categories (1851-1911)					
$\sum_{k \neq i} IOin_{ki} \ln(L_{kct})$	4,263	9.31	3.21	2.11	21.86
$\sum_{k \neq i} IOout_{ki} \ln(L_{kct})$	4,263	8.80	6.26	0.00	42.77
$\sum_{k \neq i} EMP_{ki} \ln(L_{kct})$	4,263	100.8	42.51	-92.52	191.50
$\sum_{k \neq i} OCC_{ki} \ln(L_{kct})$	4,263	36.25	25.70	-1.10	111.10
Alternative matrices and aggregated industry categories (1851-1911)					
$\sum_{k \neq i} IOin1841_{ki} \ln(L_{kct})$	2,232	2.87	2.85	0.00	12.10
$\sum_{k \neq i} IOout1841_{ki} \ln(L_{kct})$	2,232	3.98	3.88	0.00	11.77
$\sum_{k \neq i} EMP_{ki} \ln(L_{kct})$	2,232	50.22	24.98	-29.45	95.33
$\sum_{k \neq i} OCC_{ki} \ln(L_{kct})$	2,232	24.90	16.60	-0.66	67.22
Cross-city connection measures (1861-1911)					
$\sum_{k \neq i} IOin_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt})$	3,549	237.56	72.05	65.98	389.92
$\sum_{k \neq i} IOout_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt})$	3,549	223.59	152.47	0.00	741.70
$\sum_{k \neq i} EMP_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt})$	3,549	2,570.20	987.25	-1,606.76	3,417.66
$\sum_{k \neq i} OCC_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt})$	3,549	926.17	631.58	-19.62	1,995.84
MP_{ct}	3,549	15.70	0.24	14.76	16.07

Note: We report cross-city summary statistics for 1861-1911 because we only report instrumented cross-city regression results in the main text, which means that 1851 is used only to construct lagged values. For the others, we report summary statistics using the full 1851-1911 period since we report both OLS and instrumented results.

A.3 Empirical approach appendix

A.3.1 Monte Carlo simulations

We use Monte Carlo simulations to assess how well our estimation strategy performs in datasets displaying the size and characteristics of our data. The basic idea is to generate datasets that mimic our real data, but obtained from a data generating process (DGP) with known parameter values. We then apply our estimation strategy to these placebo data sets, recover parameter estimates, and compare them to the estimates obtained in the true data. This allows us to assess the ability of our estimation strategy to obtain unbiased results and accurate confidence intervals.

We begin by estimating our baseline regression specification, Eq. 18, in order to obtain a set of industry-year effects ($\hat{\varphi}_{it}$), city-year ($\hat{\theta}_{ct}$) effects, and estimated residuals $\hat{\varepsilon}_{cit}$. These ingredients will be used to simulate new datasets in which the city-year and industry-year effects are held constant at the estimated values, and the error terms are drawn from a multivariate Normal distribution whose parameters are computed using the estimated residuals.

Step 1 – constructing the simulated error term

We want to generate a simulated error vector that displays correlation within the city-year (CY), industry-year (IY) and city-industry (CI) dimensions but is uncorrelated across these dimensions. In other words, we need to draw entire vectors of errors ε_{cit} from a multivariate distribution whose covariance matrix Ω has zeros if two observations do not share any cluster, and non-zeros if they share at least a cluster. We follow Cameron *et al.* (2011) and construct such multi-clustered covariance matrix Ω as the sum of four single-clustered covariance matrices.[42]

$$\Omega = \Omega^{CY} + \Omega^{IY} + \Omega^{CI} - 2\Omega^{CIT}$$

[42] Following Cameron *et al.* (2011)'s notation, with three non-nested dimensions of clustering (denoted by A, B, C) the correct formula to compute a multi-clustered covariance matrix is $\Omega^{ABC} = \Omega^{A} + \Omega^{B} + \Omega^{C} - \Omega^{A\cap C} - \Omega^{A\cap B} - \Omega^{B\cap C} + \Omega^{A\cap B\cap C}$ where, for instance, the entries of Ω^{A} are non-zero if two observations share the same cluster along a single dimension A, while the entries of $\Omega^{A\cap B}$ are non-zero if two observations share the same cluster defined by the intersection of A and B. In our application, notice that $\Omega^{CY\cap IY} = \Omega^{CY\cap CI} = \Omega^{IY\cap CI} = \Omega^{CIT}$, therefore the formula above collapses to four distinct terms only.

Notice that if we sort the observations by a given dimension of clustering x, Ω^x has a block diagonal structure. For example, Ω^{CY} consists of blocks of zeros if the corresponding observations are not in the same city-year cluster, and blocks along the diagonal with elements potentially different from zero if the corresponding observations are from the same city-year pair. We denoted these non-zero submatrices by W^{CY} and assume that they are identical across clusters. Therefore the typical element of W^{CY} is $\sigma^{ij} = cov(\varepsilon_{\bar{c}i\bar{t}}, \varepsilon_{\bar{c}j\bar{t}}) \neq 0$.

We use the estimated residuals $\hat{\varepsilon}_{cit}$ from the baseline specification to construct the elements of each submatrix W^x. For instance, taking any two industries i and j, we set $\hat{\sigma}^{ij} = \frac{1}{\#CY} \sum_{CY} \hat{\varepsilon}_{cit}\hat{\varepsilon}_{cjt}$, where $\#CY$ is the number of different city-year pairs. We compute the elements of Ω^{IY} and Ω^{CI} in the same way. We take a different approach to compute the elements of Ω^{CIY} since each cluster has only one observation, i.e. there's a single observation for each triplet city-industry-year. All the diagonal elements of Ω^{CIY} are set to the mean squared residual, i.e. $\hat{\sigma}^{cit} = \hat{\sigma} = \frac{1}{N} \sum_{CIY} \hat{\varepsilon}^2_{cit}$, where N is the number of observations. The off-diagonal elements of Ω^{CIY} are zeros.[43]

We draw 1,000 vectors of error terms from the multivariate distribution $N(0, \Omega)$ and rescale each vector so that it has exactly the same mean (zero) and variance as the original residuals. The result of this procedure is a simulated error term $\tilde{\varepsilon}^{SIM}$ that displays correlated errors along the city-year, industry-year and city-industry dimensions with a variance matching that of the original estimated error term.

Part 2: Simulating the data

The next step in our procedure involves simulating a new set of data with the same dimensions as the original data and with known within-industry and cross-industry spillover parameters.

In order to generate a simulated growth rate for the first period we begin with the level of initial city-industry employment from the data and use Eq. 18 to compute a simulated employment growth rate for each city-industry. So, for example, if we let $\beta_1 = 0.05$ and all other β terms and τ_{ii} terms to zero then growth rate of employment in city c and industry i is:

[43] As noted in Cameron *et al.* (2011), multi-clustered covariance matrices are not guaranteed to be positive semidefinite. When that happens, as in our case, such Ω cannot be used by a random number generator. Our solution is to replace Ω with the nearest positive semidefinite matrix computed using Matlab routine nearestSPD.

$$\tilde{g}_{ic1} = 0.05 \sum_{k=i} IOin_{ki} \ln(L_{kc0}) + \hat{\varphi}_{i1} + \hat{\theta}_{c1} + \tilde{\varepsilon}_{ic1}^{CY-IY-CI} \tag{25}$$

where $IOin_{ki}$ is the actual input-output weight observed in the data. The shifters $\hat{\varphi}_{it}$ and $\hat{\theta}_{ct}$ are kept constant across simulations at the values estimated in the initial regression.

We use this simulated growth rate to obtain L_{kc1}, the level of city-industry employment in the following period, which is then fed back into Eq. 25 to obtain L_{kc2}, and so on. We repeat the process until we generate a level of employment for each city-industry-year triplet observed in the data. This procedure delivers a simulated dataset that by construction has the desired clustered error structure and the same number of observations as the original data.

Step 3: Results

We follow this procedure to generate 1,000 datasets that look like the true data, but that are generated using a data generating process with known τ_{ii} and β parameters. We apply our estimation strategy (as in Table 2 Column 6) to each of these data sets and obtain a distribution of estimated τ and β parameters.

Figure 5 displays the mean, 90% and 95% confidence intervals for the distribution of estimated parameters when β_1 is set to 0.05 and all the other spillover parameters are set to zero. As an example, we also plot the distribution of estimated coefficients for $IOin$ and $wtn1$. We can see that our estimators are asymptotically normal and unbiased.

We also perform a second Monte Carlo exercise in which we set all β and τ parameters to zero and then compare the distribution of estimated coefficients coming out of this counterfactual DGP with the estimates obtained using the real dataset. This allows us to asses the likelihood of observing the real dataset and the corresponding estimates under the null hypothesis that all parameters are zeros. This method provides us with an alternative way to do hypothesis testing that does not rely on our multi-dimensional clustered standard errors.

Figure 6 plots the distribution of estimated IOin parameters obtained using the 1000 simulated data sets, as well as the estimate obtained from the true data. These results suggests that obtaining the point estimate for $IOin$ of 0.0622 that we got from the true data (Table 2, Column 6) is extremely unlikely when the true parameter value is zero. The implied p-value is 0.000 and the coefficient is significantly different from zero at the 1% level.

Table 11 presents the similar results for all the other coefficients of interest and confirms the significance levels of our baseline results from Column 6 of Table 2. This is reassuring because one may wonder whether our dataset is sufficiently large to consistently estimate all the parameters of interest, especially given that the observations are potentially correlated across multiple dimensions.

Discussion

These monte carlo results can help us assess how well our approach performs on simulated data sets sharing the same size and variance as the data used in our main analysis. However, this procedure comes with obvious limitations. In particular, we are assuming that the model is correctly specified and that the error terms are clustered in a particular way. Thus, this simulation cannot be used to assess how well our procedure performs under alternative data generating processes or when standard errors display alternative clustering patterns.

Figure 5: Simulated results when $\beta_1 = 0.05$ and all other spillover parameters are zero

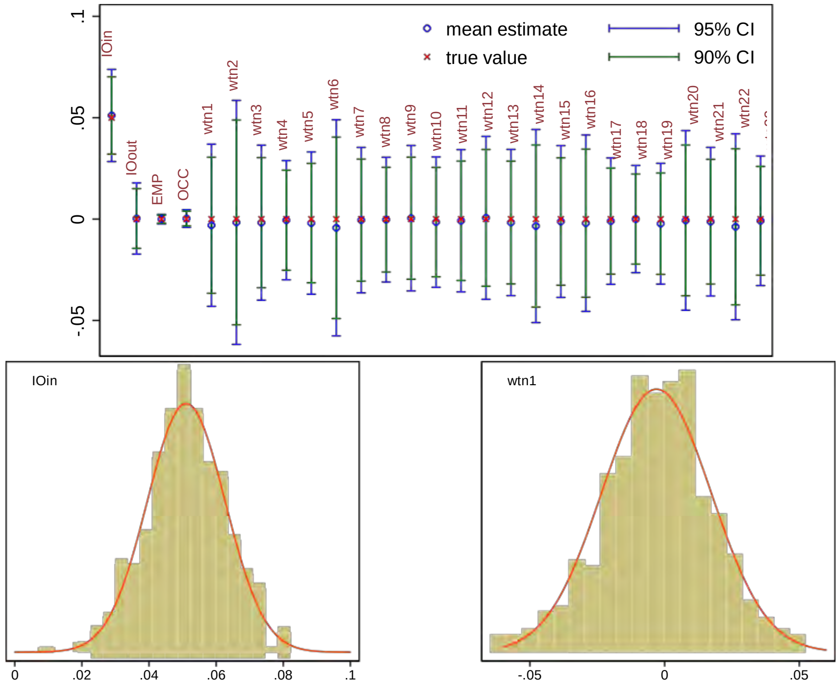


Figure 6: Simulated results with all parameters are set to zero vs. *IOin* estimate on real data

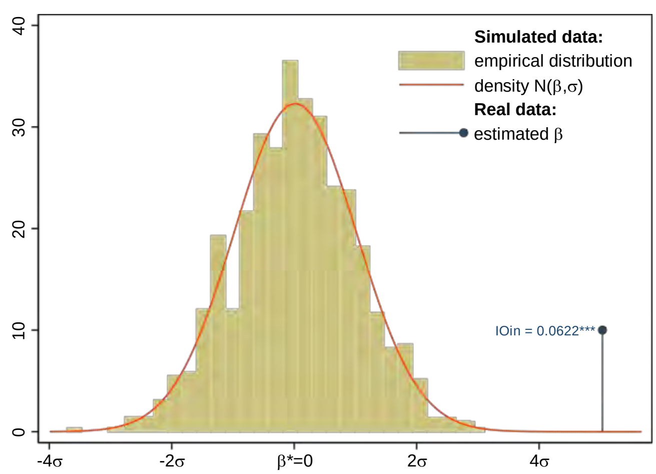


Table 11: Simulated results with all parameters are set to zero vs. parameter estimates from true data

	Simulated Data		True Data	
Variable	Mean	Std. Dev.	Coef.	p-value
EMP	.000	.001	.002	.198
IOin	.000	.012	.062	.000
IOout	.000	.011	-.018	.096
OCC	.000	.002	.007	.002
wtn1	-.002	.023	-.094	.000
wtn10	-.001	.018	-.124	.000
wtn11	-.001	.017	.041	.016
wtn12	.000	.024	-.026	.274
wtn13	-.002	.022	-.074	.001
wtn14	-.003	.027	-.102	.000
wtn15	-.002	.022	-.04	.068
wtn16	-.001	.024	-.069	.004
wtn17	-.001	.019	-.066	.001
wtn18	.000	.013	-.032	.016
wtn19	-.001	.017	-.028	.092
wtn2	-.001	.033	-.01	.752
wtn20	-.001	.023	-.005	.84
wtn21	-.001	.02	-.061	.002
wtn22	-.003	.028	-.142	.000
wtn23	-.001	.018	-.046	.011
wtn3	-.003	.021	-.032	.123
wtn4	-.001	.018	-.089	.000
wtn5	-.001	.021	-.062	.003
wtn6	-.003	.034	-.078	.023
wtn7	.000	.021	-.058	.007
wtn8	.000	.017	-.004	.818
wtn9	.000	.022	-.071	.001

For each of the key explanatory variables, the first two columns of this table present the mean and standard deviation of the distribution of coefficient estimates obtained from applying our estimation strategy to 1000 simulated datasets where the data have been generated with all spillover parameter values set to zero. Column 3 presents the coefficients estimated using the true data (as in Table 2, Column 6). Column 4 presents the p-value implied by comparing the coefficients estimated using the true data to the distribution of coefficient estimates obtained from the simulated data.

A.3.2 KP test appendix

The standard errors in all of our main regressions are clustered along multiple dimensions. When using 2sls regressions, it is useful to be able to calculate the Kleibergen & Paap (2006) test statistics for under- and weak-identification using the appropriately clustered covariance matrix. The KP statistics can easily be computed using existing Stata routines, but only for up to two non-nested dimensions of clustering (Kleibergen (2010)). None of these routines can handle a higher number of clusters so we developed our own package, which we will make available to the benefit of other researchers.

Our strategy builds on Thompson (2011) and Cameron *et al.* (2011) to compute a multi-clustered covariance of the orthogonality condition for any number of clusters. We then use a modified version of the Stata program *ranktest* to compute the appropriate KP statistics based on this covariance matrix. It can be verified that our program exactly reproduces the rk statistic (under-identification) and Wald statistic computed by ranktest in the case of two clusters. The weak-identification test statistic is then computed by transforming the Wald statistic into an F statistic. Notice that the value of our F statistic does not exactly match the one computed by *ivreg2* due to the very small differences in the small sample adjustment.

A.4 Results appendix

A.4.1 Robustness of results to dropping cities or industries

Figure 7 presents histograms of t-statistics for each cross-industry term obtained from running regressions equivalent to Column 6 of Table 2, where in each regression a different city is dropped from the dataset. This allows us to assess the extent to which our results are robust to changes in the set of cities included in the analysis. These results indicate that our estimates are not sensitive to dropping individual cities from the analysis database.

Figure 7: Robustness to dropping one city at a time – distribution of t-statistics

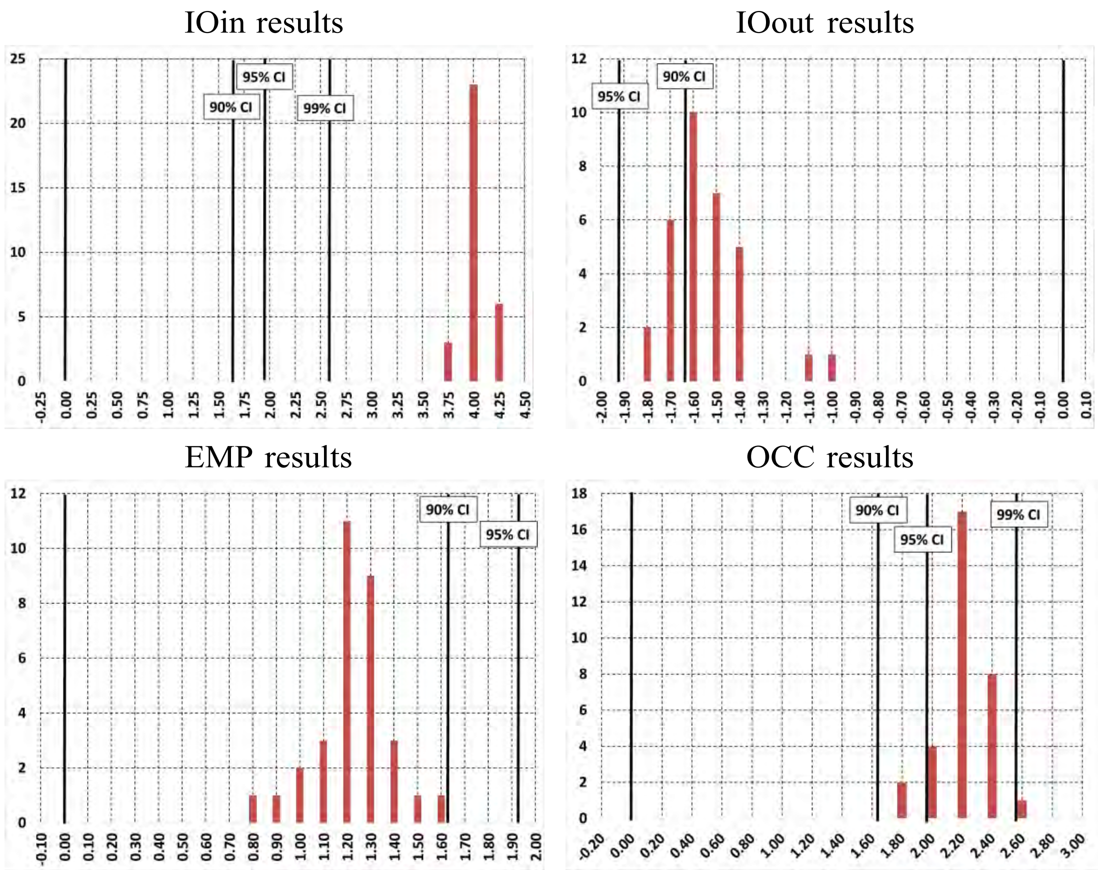
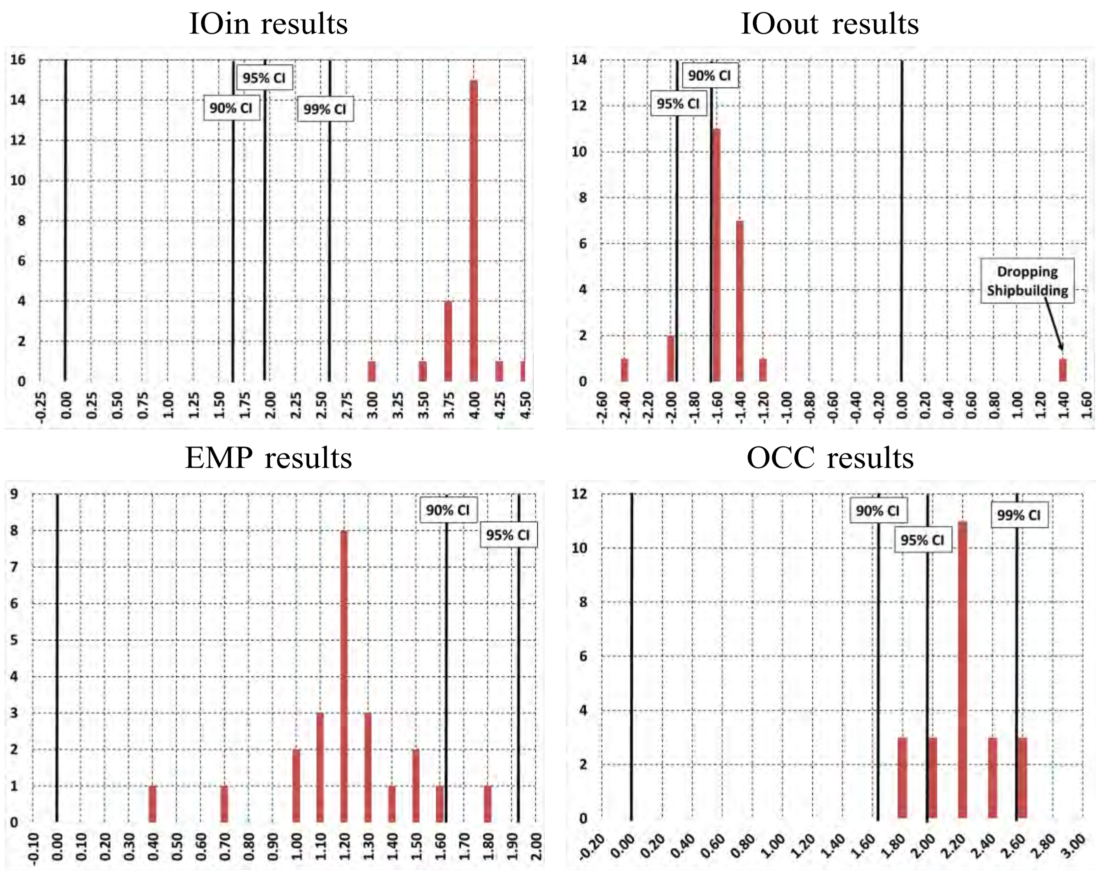


Figure 8 presents histograms of t-statistics for each cross-industry term obtained from running regressions equivalent to Column 6 of Table 2, where in each regression a different industry is dropped from the dataset. This allows us to assess the extent to which our results are robust to changes in the set of industries included in the analysis. We can see that in general our estimated coefficients are not sensitive to dropping individual industries. However, this does not apply when looking at the IO out coefficient. The top-right graph shows that when we drop shipbuilding from the data, the IO out coefficient changes substantially. In particular, the estimated coefficient changes from negative and occasionally statistically significant to positive and not statistically significant. This suggests that the negative coefficient estimated on the IO out coefficient is driven entirely by the Shipbuilding industry. This is an unusual industry because presumably it can only operate in coastal cities or those with access to a major navigable river. Thus, the IO out results obtained when dropping this industry seem more reasonable. These results suggest that in general the impact of local customers is weakly positive.

Overall, the results in Figure 8 indicate that our estimates are much more sensitive to dropping industries than they are to dropping cities. This suggests that heterogeneity across industries is more important than heterogeneity across cities.

Figure 8: Robustness to dropping one industry at a time – distribution of t-statistics



A.4.2 Heterogeneous effects

In this section we look at heterogeneity in the pattern of cross-industry and within-industry effects across different industries. We begin by considering heterogeneous cross-industry effects. Specifically, we run two alternative versions of Equation 18,

$$\Delta \ln(L_{ict+1}) = \tilde{\gamma}_{ii} \ln(L_{ict}) + \beta^{i} \sum_{k \neq i} CONNECT_{ki} \ln(L_{kct}) + \theta_{ct} + \chi_{it} + e_{ict} \tag{26}$$

$$\Delta \ln(L_{i \neq k\, ct+1}) = \tilde{\gamma}_{ii} \ln(L_{ict}) + \beta^{k}\, CONNECT_{ki} \ln(L_{kct}) + \theta_{ct} + \chi_{it} + e_{ict} \tag{27}$$

where $CONNECT_{ki}$ is one of our four measures of cross-industry connections. Equation 26 allows us to estimate industry-specific coefficients β^i describing how much each industry i benefits from cross-industry connections. This specification can be estimated using the same approach as was used for our baseline regressions. Using Equation 27, we estimate industry-specific coefficients β^k that reflect the extent to which industry k generates spillovers that benefit other industries. Estimating this value requires a different approach to avoid conflating the within and between impact of industry k when estimating β^k. Specifically, we run separate regressions corresponding to Equation 27 for each industry k. In each of these regressions, only employment in industry k (interacted with a cross-industry connection measure) is included as an explanatory variable and observations from industry k are not included in the dependent variable.

Once the industry-specific β^i and β^k terms are estimated, we compare them to available measures of industry characteristics: firm size in each industry, the share of output exported, the share of output sold to households, the industry labor cost share, and the industry intermediate cost share. In each case we run a simple univariate regression where the dependent variable is the estimated industry-specific cross-industry spillover coefficient and the independent variable is one of the industry characteristics. Univariate regressions are used because we are working with a relatively small number of observations. These results can provide suggestive evidence about the characteristics of industries that produce or benefit from different types of cross-industry spillovers, but because of the small sample size we will not be able to draw any strong conclusions.

Table 12 describes the characteristics of industries that *benefit from* cross-industry connections. In rows 1-2, we see evidence that small firm size in an industry is associated with more cross-industry spillover benefits, but this pattern is not statistically significant at standard confidence levels. The only strong result coming out of this table is that industries that benefit from connections to other local industries with similar labor pools tend to have a larger labor cost share relative to overall industry sales, as well as a smaller intermediate cost share. This seems like a very reasonable result which provides some additional confidence that the estimates we have obtained are reasonable.

Table 12: Features of industries that benefit from each type of cross-industry spillover

Coefficients from univariate regressions				
	DV: Estimated industry-specific β^i coefficient			
Spillovers channel:	IO-in	IO-out	EMP	OCC
Average firm size	-0.210	-0.902	-0.0353	-0.209
	(0.319)	(0.559)	(0.0350)	(1.122)
Median worker's firm size	-0.0179	-0.108	-0.00289	-0.0441
	(0.0377)	(0.0651)	(0.00417)	(0.131)
Share of industry output exported abroad	0.0185	-0.122	-0.0163	-0.313
	(0.0982)	(0.178)	(0.0114)	(0.333)
Share of industry output sold to households	0.0300	0.121	0.00746	0.170
	(0.0443)	(0.0864)	(0.00519)	(0.150)
Labor cost/output ratio	-0.137	-0.185	-0.00769	0.426**
	(0.147)	(0.280)	(0.0100)	(0.191)
Intermediate cost/output ratio	0.0196	0.0819	-0.000385	-0.373***
	(0.109)	(0.194)	(0.00737)	(0.125)

Estimated coefficients from univariate regressions. Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$. The dependent variable in each regression is the estimated β^i coefficient from Eq. 26. Firm size data comes from the 1851 Census of Population. The share of industry output exported or sold to households is from the 1907 input-output table. The labor cost share is constructed from industry wage bills from the 1907 Census of Manufactures. The intermediate cost share is based on the 1907 input-output table. We do not report robust standard errors because these generate smaller confidence intervals, probably due to small-sample bias. We have also explored regressions in which we weight results by the inverse of the standard error of each estimated within-industry coefficient in order to account for the precision of those estimates and these deliver similar results.

Table 13 describes the characteristics of industries that *produce* cross-industry connections. These results also suggest that industries with smaller firm sizes produce more beneficial cross-industry spillovers, but again, these results are not statistically significant. As before, the only clear relationship that we observe is that industries with a greater labor cost share (and smaller intermediate cost share) relative to overall sales produce more cross-industry benefits to occupationally similar industries.

Table 13: Features of industries that produce each type of cross-industry spillover

Coefficients from univariate regressions				
	DV: Estimated industry-specific β^k coefficient			
Spillovers channel:	IO-in	IO-out	EMP	OCC
Average firm size	-1.496	-3.899	0.0487	-1.022
	(1.239)	(6.452)	(0.174)	(2.133)
Median worker's firm size	-0.163	-0.626	0.00115	-0.149
	(0.147)	(0.752)	(0.0206)	(0.250)
Share of industry output exported abroad	0.0742	-0.797	0.00417	-0.341
	(0.407)	(1.994)	(0.0539)	(0.648)
Share of industry output sold to households	0.149	0.0470	-0.0169	0.418
	(0.201)	(0.905)	(0.0241)	(0.280)
Labor cost/output ratio	-0.324	0.651	-0.0251	0.983*
	(0.625)	(3.212)	(0.0440)	(0.524)
Intermediate cost/output ratio	-0.197	-0.0637	0.0142	-0.870**
	(0.420)	(2.261)	(0.0311)	(0.338)

Estimated coefficients from univariate regressions. The dependent variable in each regression is the estimated β^k coefficient from Eq. 27. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Firm size data comes from the 1851 Census of Population. The share of industry output exported or sold to households is from the 1907 input-output table. The labor cost share is constructed from industry wage bills from the 1907 Census of Manufactures. The intermediate cost share is based on the 1907 input-output table. We do not report robust standard errors because these generate smaller confidence intervals, probably due to small-sample bias. We have also explored regressions in which we weight results by the inverse of the standard error of each estimated within-industry coefficient in order to account for the precision of those estimates and these deliver similar results.

Next, we undertake a similar exercise with our estimated within-industry coefficients. In Table 14 we consider some of the industry characteristics that may be related to the range of different within-industry spillover estimates we observe. Columns 1-2 focus on the role of firm size using two different measures. We observe a positive relationship between firm size in an industry and the strength of within-industry spillovers, but this results is not statistically significant due to the small number of available observations. There is also weak evidence that more labor intensive industries benefit more from within-industry spillovers.

Table 14: Features of industries that benefit from within-industry spillovers

DV: Estimated industry-specific within-industry spillover coefficients						
Average firm size	0.289					
	(0.196)					
Median worker's firm size		0.0253				
		(0.0236)				
Exports share of industry output			0.0428			
			(0.0708)			
Households share of industry output				-0.0384		
				(0.0314)		
Labor cost/output ratio					0.136	
					(0.0983)	
Intermediate cost/output ratio						-0.0115
						(0.0755)
Observations	20	20	23	23	16	16
R-squared	0.107	0.060	0.017	0.066	0.121	0.002

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The number of observations varies because the explanatory variables are drawn from different sources and are not available for all industries. The within coefficients come from the specification used in Column 6 of Table 2. Firm size data comes from the 1851 Census of Population. The export's and household's share of industry output come from the input-output table. Total labor cost and total output values come from the 1907 Census of Production. Intermediate cost is constructed based on data from the 1907 Input-Output matrix. We do not report robust standard errors because these generate smaller confidence intervals, probably due to small-sample bias. We have also explored regressions in which we weight results by the inverse of the standard error of each estimated within-industry coefficient in order to account for the precision of those estimates and these deliver similar results.

A.4.3 Robustness: Alternative functional forms

In this table we replace the logarithms on the right-hand side of the estimating equation with plausible alternative functional forms based on either the second root or fifth root. These results show that adjusting the functional form in this way has little impact on the estimated results.

Table 15: Regression results with alternative functional forms

FF:	Square root			Fifth root		
	(1)	(2)	(3)	(4)	(5)	(6)
IOin	0.0017***	0.0016***	0.0016***	0.0779***	0.0651***	0.0659***
	(0.0005)	(0.0005)	(0.0005)	(0.0205)	(0.0187)	(0.0190)
IOout	-0.0003	-0.0004	-0.0004	-0.0098	-0.0149	-0.0173
	(0.0004)	(0.0004)	(0.0004)	(0.0126)	(0.0125)	(0.0127)
EMP	-0.0000	0.0000	0.0000	-0.0001	0.0022	0.0019
	(0.0001)	(0.0000)	(0.0001)	(0.0021)	(0.0015)	(0.0017)
OCC	0.0003***	0.0003*	0.0003**	0.0108***	0.0085**	0.0088**
	(0.0001)	(0.0001)	(0.0001)	(0.0036)	(0.0040)	(0.0040)
Observations	4,263	3,554	3,554	4,263	3,554	3,554
Estimation	ols	2sls	2sls	ols	2sls	2sls
Instrumented	none	wtn	wtn-btn	none	wtn	wtn-btn
KP under		20.74	20.92		26.73	28.71
KP weak		103.9	62.81		79.44	49.83

Multi-level clustered standard errors by city-industry, city-year, and industry-year in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. A full set of industry-specific *within* terms, industry-year and city-year effects are included in all regressions but not displayed. Regressions in Columns 2 and 4 instrument the *within* terms with lagged values. Regressions in Columns 3 and 5 instrument both the *within* and *between* terms with lagged values. Acronyms: wtn = *within*, btn = *between*. "KP under" denotes the test statistic for the Lagrange Multiplier underidentification test based on Kleibergen & Paap (2006). "KP weak" denotes the test statistic for a weak instruments test based on the Kleibergen-Paap Wald statistic.

A.4.4 Robustness: Alternative connections matrices

Next, we revisit the analysis using some alternative measures of inter-industry connections. In particular, we use an alternative matrix of input-output connections constructed by Horrell *et al.* (1994) for Britain in 1841. Generating results with this alternative matrix, which comes from before the study period, can help address concerns that the results we find are dependent on the specific set of matrices we consider or are due to a process of endogenous inter-industry connection formation. The cost of using this matrix is that we are forced to work with a smaller set of 12 more aggregated industry categories.[44]

Because we are now working with a smaller number of industry categories, we focus

[44]The industry categories are: "Mining & quarrying," "Food, drink & tobacco", "Metals & Machinery," "Oils, chemicals & drugs," "Textiles, clothing & leather goods," "Earthenware & bricks," "Other manufactured goods," "Construction," "Gas & water," "Transportation," "Distribution," and "All other services."

on regressions that incorporate one spillover channel at a time. Table 16 describes the results. As in the main results, we observe positive effects occurring through the IOin channel and these results are generally statistically significant. There is also evidence that industries may have benefited from the presence of local buyers, but this result is clearly sensitive to the underlying set of industries used, so it should be interpreted with some caution. There is also some evidence of benefits through the presence of occupationally similar local industries.

Table 16: Alternative matrix regressions with one channel at a time

	(1)	(2)	(3)	(4)	(5)	(6)
IOin1841	0.0490***	0.0346**	0.0421**			
	(0.0134)	(0.0152)	(0.0164)			
IOout1841				0.0383***	0.0555***	0.0570***
				(0.0141)	(0.0151)	(0.0152)
Observations	2,232	1,860	1,860	2,232	1,860	1,860
Estimation	ols	2sls	2sls	ols	2sls	2sls
Instrumented	none	wtn	wtn-btn	none	wtn	wtn-btn
KP under id.		326.29	263.61		351.47	297.76
KP weak id.		292.93	179.49		366.68	239.76
	(7)	(8)	(9)	(10)	(11)	(12)
EMP	0.0028	0.0050***	0.0051**			
	(0.0019)	(0.0019)	(0.0020)			
OCC				0.0058*	0.0049	0.0048
				(0.0035)	(0.0041)	(0.0041)
Observations	2,232	1,860	1,860	2,232	1,860	1,860
Estimation	ols	2sls	2sls	ols	2sls	2sls
Instrumented	none	wtn	wtn-btn	none	wtn	wtn-btn
KP under id.		493.03	445.87		301.94	308.52
KP weak id.		494.71	391.98		349.01	262.24

Multi-level clustered standard errors by city-industry, city-year, and industry-year in parenthesis. *** p<0.01, ** p<0.05, * p<0.1. A full set of *within* regressors, city-time and industry-time effects are included in all regressions but not displayed. 2SLS regressions use lagged instruments. Note that the number of observations falls for the instrumented regressions because the instruments require a lagged employment term. Thus, data from 1851 are not available for these regressions. Acronyms: wtn = *within*, btn = *between*. "KP under id." denotes the test statistic for the Lagrange Multiplier underidentification test based on Kleibergen & Paap (2006). "KP weak id." denotes the test statistic for a weak instruments test based on the Kleibergen-Paap Wald statistic.

A.4.5 Robustness: Cross-city effects

There is substantial variation in the proximity of cities in our database to other nearby cities. Some cities, particularly those in Lancashire, West Yorkshire, and the North Midlands, are located in close proximity to a number of other nearby cities. Others, such as Norwich, Hull, and Portsmouth are relatively more isolated. In this section, we extend our analysis to consider the possibility that city-industry growth may also be affected by forces due to other nearby cities.

We consider two potential channels for cross-city effects. First, industries may benefit from proximity to consumers in nearby cities. This *market potential* effect has been suggested by Hanson (2005), who finds that regional demand linkages play an important role in generating spatial agglomeration using modern U.S. data. Second, industries may benefit from spillovers from other industries in nearby towns, through any of the channels that we have identified. We analyze these effects using the more detailed industry categories from Section 5.

We begin our analysis by collecting data on the distance (as the crow flies) between each of the cities in our database, which we call $distance_{ij}$. Using these, we construct a measure for the remoteness of one city from another $d_{ij} = exp(-distance_{ij})$.[45] Our measures of market potential for each city is then,

$$MP_{ct} = \ln \left(\sum_{j \neq c} POP_{jt} * d_{cj} \right).$$

where POP_{jt} is the population of city j. This differs slightly from Hanson's approach, which uses income in a city instead of population, due to the fact that income at the city level is not available for the period we study.

We also want to measure the potential for cross-industry spillovers occurring across cities. We measure proximity to an industry i in other cities as the distance-weighted sum of log employment in that industry across all other cities. Our full regression specification, including both cross-city market potential and spillover effects, is then,

$$\Delta \ln(L_{ict+1}) = \tilde{\gamma}_{ii} \ln(L_{ict})$$

[45] This distance weighting measure is motivated by Hanson (2005). We have also explored using $d_{ij} = 1/distance_{ij}$ as the distance weighting measure and this delivers similar results.

$$
\begin{aligned}
&+ \beta_1 \sum_{k \neq i} IOin_{ki} \ln(L_{kct}) + \beta_2 \sum_{k \neq i} IOout_{ki} \ln(L_{kct}) \\
&+ \beta_3 \sum_{k \neq i} EMP_{ki} \ln(L_{kct}) + \beta_4 \sum_{k \neq i} OCC_{ki} \ln(L_{kct}) \\
&+ \beta_5 \left(\sum_{k \neq i} IOin_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt}) \right) + \beta_6 \left(\sum_{k \neq i} IOout_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt}) \right) \\
&+ \beta_7 \left(\sum_{k \neq i} EMP_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt}) \right) + \beta_8 \left(\sum_{k \neq i} OCC_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt}) \right) \\
&+ \beta_9 MP_{ct} + \log(WORKpop_{ct}) + \theta_c + \chi_{it} + \epsilon_{ict}.
\end{aligned}
$$

One difference between this and our baseline specification is that we now include city fixed effects (θ_c) in place of city-year effects because city-year effects would be perfectly correlated with the market potential measure. To help deal with city-size effects, we also include the log of $WORKpop_{ct}$, the working population of city c in period t. To simplify the exposition and in analogy with the previous section, we will refer to the cross-city term $\sum_{k \neq i} IOin_{ki} \sum_{j \neq c} d_{jc} * \ln(L_{kjt})$ as $IOin * d$, and similarly for the other cross-city terms $IOout * d$, $EMP * d$, and $OCC * d$.

The results generated using this specification are shown in Table 17. The first thing to take away from this table is that our baseline results are essentially unchanged when we include the additional cross-city terms. The city employment term in the fifth column reflects the negative growth impact of city size. The coefficients on the market potential measure is always positive but not statistically significant.

Table 17: Regression results with cross-city variables

	(1)	(2)	(3)
IOin	0.0571***	0.0604***	0.0586***
	(0.0144)	(0.0154)	(0.0166)
IOout	-0.0248**	-0.0252**	-0.0257**
	(0.0109)	(0.0108)	(0.0111)
EMP	-0.0027	-0.0029	-0.0029
	(0.0018)	(0.0018)	(0.0018)
OCC	0.0064*	0.0062*	0.0061*
	(0.0033)	(0.0034)	(0.0034)
City employment	-0.3377***	-0.3295***	-0.3321***
	(0.0762)	(0.0753)	(0.0767)
Market Potential	0.1592		0.1176
	(0.1611)		(0.2631)
IOin*dist		0.0012	0.0004
		(0.0017)	(0.0024)
IOout*dist		-0.0008	-0.0007
		(0.0010)	(0.0010)
EMP*dist		0.0002	0.0001
		(0.0001)	(0.0001)
OCC*dist		-0.0001	-0.0001
		(0.0002)	(0.0002)
Observations	3,549	3,549	3,549
KP under	19.34	20.79	19.38
KP weak	2.07	2.3	2.07

Multi-level clustered standard errors by city-industry, city-year, and industry-year in parenthesis. Significance levels: *** p<0.01, ** p<0.05, * p<0.1. A full set of *within* regressors, city-time and industry-time effects are included in all regressions but not displayed. All regressions instrument the *within* and *between* regressors with lagged instruments. Acronyms: wtn = *within*, btn = *between*. "KP under" denotes the test statistic for the Lagrange Multiplier underidentification test based on Kleibergen & Paap (2006). "KP weak" denotes the test statistic for a weak instruments test based on the Kleibergen-Paap Wald statistic.

The results do not provide statistically significant evidence that cross-city spillovers matter through any of the channels that we measure. However, these results are imprecisely measured. The coefficients estimated on the $IOin * dist$ term suggest that a one standard deviation increase in the presence of suppliers in other nearby cities could increase city-industry growth by 6.1-18.3%. The coefficients on the EMP term are consistent with effects of a similar magnitude. Thus, we should not rule out important cross-city effects based on these results. However, it is clear that omitted cross-city effects are not driving our findings regarding the importance of within-city cross-industry agglomeration forces.

A.4.6 Additional results for the city-size effects

We may be concerned that the results described in Table 4 are driven in part by the inclusion of industry-time effects in the regressions used to obtain the θ_{ct} terms. One way to assess this is to estimate alternative city-time effects from,

$$6 \ln(L_{ict+1}) = \theta^{FE}_{ct} + \chi_{it} + e_{ict}, \tag{28}$$

and then estimate,

$$\theta^{FE}_{ct} = c_0 + c_1 \ln(WORKpop_{ct}) + f_{ct}. \tag{29}$$

Because the only difference between the specification in Equation 18 and that in Equation 28 is the inclusion of the within and cross-industry agglomeration terms, we can be sure that any differences between the estimated θ_{ct} terms and the θ^{FE}_{ct} terms are due to these agglomeration forces.

The results in Table 18 mirror those shown in Table 4 except that the relationship between city size and the estimated θ_{ct} coefficients are now compared against the relationship between city size and the estimated $\theta^{FE}{}_{ct}$ coefficients from Eq. 29. In essence, this comparison is ensuring that the convergence results we obtain are not driven by the inclusion of industry-year effects in the regressions. We can see that the results in Table 18 are very similar to the results in Table 4, which suggests that the inclusion of industry-year effects is not playing an important role in generating our results.

Table 18: Measuring the aggregate strength of the agglomeration forces against an estimated baseline

Results based on unweighted regressions

	Results based on θ^{CROSS}		Results based on θ^{FE}		Difference: aggregate strength of agglomeration force (implied divergence
	Estimated city-size coefficient	Implied divergence Beta	Estimated city-size coefficient	Implied divergence Beta	
1861-1871	-0.133	1.43%	-0.055	0.56%	0.87%
1871-1881	-0.121	1.29%	-0.044	0.45%	0.84%
1881-1891	-0.089	0.93%	-0.012	0.12%	0.81%
1891-1901	-0.087	0.91%	-0.010	0.10%	0.81%
1901-1911	-0.093	0.97%	-0.013	0.13%	0.84%

Results based on regressions weighted by city-industry size in 1851

	Results based on θ^{CROSS}		Results based on θ^{FE}		Difference: aggregate strength of agglomeration force (implied divergence
	Estimated city-size coefficient	Implied divergence Beta	Estimated city-size coefficient	Implied divergence Beta	
1861-1871	-0.068	0.71%	-0.049	0.50%	0.21%
1871-1881	-0.059	0.60%	-0.038	0.39%	0.21%
1881-1891	-0.039	0.40%	-0.019	0.19%	0.21%
1891-1901	-0.034	0.34%	-0.014	0.14%	0.20%
1901-1911	-0.024	0.24%	-0.004	0.04%	0.20%

Column 1 presents the a_1 coefficients from estimating Equation 19 for each decade (cross-sectional regressions). Column 2 presents the decadal divergence rates implied by these coefficients. Column 3 presents the c_1 coefficients from estimating Equation 29 and Column 4 presents the decadal divergence rates implied by these coefficients. Column 5 gives the aggregate strength of the divergence force represented by the agglomeration economies, which is equal to the difference between the decadal convergence coefficients. The results in the top panel are based on city-time effects estimated from unweighted regressions while the results in the bottom panel are based on city-time effects estimated using weighted regressions based on city-industry employment in 1851.