American Economic Review 2019, 109(1): 153–170
https://doi.org/10.1257/aer.20130249

A Spatial Knowledge Economy†

By Donald R. Davis and Jonathan I. Dingel*

Leading empiricists and theorists of cities have recently argued that the generation and exchange of ideas must play a more central role in the analysis of cities. This paper develops the first system of cities model with costly idea exchange as the agglomeration force. The model replicates a broad set of established facts about the cross section of cities. It provides the first spatial equilibrium theory of why skill premia are higher in larger cities and how variation in these premia emerges from symmetric fundamentals. (*JEL* J24, J31, O31, R12, R23)

In modern economies driven by innovation and ideas, local economic outcomes increasingly depend on local idea generation. Empirically, the spatial distribution of human capital has consequences for productivity, prices, and inequality (Rauch 1993, Moretti 2004, Diamond 2016). Theoretically, however, the exchange of ideas in cities has often been treated as a special case of "black box" local external economies.[1]

This paper introduces a model in which costly exchange of ideas is the agglomeration force driving a variety of spatial phenomena. Heterogeneous individuals, drawn from a continuous distribution of ability, may produce tradables or non-tradables,

* Davis: Department of Economics, Columbia University, 420 W 118 Street, New York, NY 10027, and NBER (email: drdavis@columbia.edu); Dingel: Booth School of Business, University of Chicago, 5807 S Woodlawn Avenue, Chicago, IL 60637, and NBER (email: jdingel@chicagobooth.edu). This paper was accepted to the *AER* under the guidance of Debraj Ray, Coeditor. We thank four anonymous referees, Pol Antràs, Kerem Coşar, Arnaud Costinot, Gilles Duranton, Jessie Handbury, Walker Hanlon, Sam Kortum, Corinne Low, Ben Marx, Joan Monras, Suresh Naidu, Kentaro Nakajima, Stephen Redding, Holger Sieg, Daniel Sturm, Eric Verhoogen, Reed Walker, David Weinstein, and seminar participants at the CESifo conference on heterogeneous firms in international trade, Columbia applied micro and international trade colloquia, Conference on Urban and Regional Economics, Empirical Investigations in International Trade, NBER ITI meeting, NYU, Urban Economics Association annual meeting, Princeton IES Summer Workshop, Spatial Economic Research Centre annual conference, and University of Toronto for helpful comments on various drafts. We thank Yuxiao Huang, Paul Piveteau, and especially Meru Bhanot for research assistance. We are grateful to Enrico Moretti and Stuart Rosenthal for sharing their housing-price measures with us. Dingel gratefully acknowledges resources provided by the University of Chicago Research Computing Center and financial support from the Institute for Humane Studies, the Program for Economic Research at Columbia University, and the Kathryn and Grant Swick Faculty Research Fund at the University of Chicago Booth School of Business. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

† Go to https://doi.org/10.1257/aer.20130249 to visit the article page for additional materials and author disclosure statement(s).

[1] Abdel-Rahman and Anas (2004, p. 2300): "One way to interpret this black-box model [of Marshallian externalities] is that the productivity of each worker is enhanced by the innovative ideas freely contributed by the labor force working in close proximity." Fujita, Krugman, and Venables (1999, p. 4), and Fujita and Thisse (2002, p. 129) criticize the black-box version for being evanescent in empirical terms and close to assuming the conclusion in theoretical terms. Duranton and Puga (2004, p. 2065) describe "looking inside the black box… as one of the fundamental quests in urban economics."

and higher-ability individuals have comparative advantage in tradables. Tradables producers divide their time between producing and exchanging ideas with each other in order to raise their productivity. Cities with more numerous and higher-ability partners are better idea-exchange environments. Higher-ability individuals benefit more from these conversations, so they locate in larger cities, paying higher local prices to realize more valuable idea exchanges. In equilibrium, larger cities exhibit better idea-exchange opportunities because their tradables producers are more talented, greater in number, and devote more time to exchanging ideas. Less skilled individuals are employed in every city producing non-tradables, and larger cities have higher non-tradables prices to compensate them for their higher costs of living.

Our model replicates a broad set of empirical facts about the cross section of cities. First, while our model has symmetric fundamentals, idea-driven agglomeration generates cities of heterogeneous sizes.[2] Second, larger cities exhibit higher nominal wages, housing prices, and productivity in equilibrium (Glaeser 2008). Third, larger cities' higher wages are partly attributable to higher-ability individual sorting into those locations, but this sorting is incomplete and individuals of many skill types are present in every city (Combes, Duranton, and Gobillon 2008; Gibbons, Overman, and Pelkonen 2014; Carlsen, Rattsø, and Stokke 2016; de la Roca and Puga 2017).

This account of the spatial distribution of heterogeneous labor yields a novel prediction about spatial variation in skill premia. Since higher-ability tradables producers locate in larger cities and raise their productivity by exchanging ideas while non-tradables productivity does not vary across locations, the relative productivity of tradables producers is increasing in city size. This causes relative wages to increase with city size when the productivity gap is only partially offset by higher non-tradables prices. Empirically, Table 1 shows that the college wage premium rises significantly with metropolitan population. This measure of the skill premium ranges from about 47 percent in metros with 100,000 residents to about 71 percent in metros with 10 million residents. This relationship is robust to controlling for two other city characteristics that prior work has linked to cities' skill premia: the fraction of the population possessing a college degree and housing prices.[3] The positive correlation between cities' population sizes and skill premia is a robust, persistent, first-order feature of the data.[4]

Theoretically linking together cities, ideas, and skill premia is nontrivial. Unlike temporal differences in wage premia, spatial differences in wage premia are disciplined by a no-arbitrage condition. As Glaeser (2008, p. 85) notes, when people are mobile, differences in productivity "tend to show up exclusively in changes in quantities of skilled people, not in different returns to skilled people across space." The canonical spatial-equilibrium model, in which there are two homogeneous skill groups and preferences are homothetic, predicts that skill premia are spatially

[2] In classic models, heterogeneity across industries supports heterogeneous city sizes (Henderson 1974). In our model, heterogeneity across individuals supports this outcome.

[3] See Glaeser (2008); Glaeser, Resseger, and Tobio (2009); and Beaudry, Doms, and Lewis (2010) on college shares and Black, Kolesnikova, and Taylor (2009) on housing prices.

[4] Online Appendix C.2 reports regressions for 1990 and 2007 that also show a positive premia-size relationship. More broadly, Wheeler (2001); Glaeser, Resseger, and Tobio (2009); Behrens and Robert-Nicoud (2014); and Baum-Snow and Pavan (2013) relate other measures of wage inequality to city size.

Table 1—Skill Premia and Metropolitan Characteristics, 2000

Population (log)	0.031	0.029	0.034	0.026
	(0.0037)	(0.0053)	(0.0042)	(0.0049)
Rent (log)		0.019		0.097
		(0.033)		(0.035)
College ratio (log)			−0.036	−0.069
			(0.018)	(0.018)
R^2	0.151	0.153	0.171	0.199

Notes: Robust standard errors in parentheses. Each column reports an OLS regression with 325 observations. The dependent variable is a metropolitan area's difference in average log hourly wages between college and high school graduates. See online Appendix C for details.

invariant (Black, Kolesnikova, and Taylor 2009). Our departure from these standard assumptions yields a novel prediction that matches the data.

Our modeling of heterogeneous abilities, cities, and skill premia in a setting with spatially symmetric fundamentals distinguishes our theory from recent work that engages these topics by assuming either asymmetric fundamentals or talent-homogeneous cities. A number of recent contributions have sought to explain differences in outcomes for skilled and unskilled workers across cities by appealing to exogenous differences in fundamental characteristics of those cities.[5] A recent paper by Behrens, Duranton, and Robert-Nicoud (2014) assumes symmetric fundamentals and a continuum of abilities, as we do, but they focus on equilibria with heterogeneous cities in which each city is populated by individuals of only one ability. A theory of talent-homogeneous cities cannot explain spatial variation in skill premia.[6]

The role of cities in facilitating idea exchange has been noted by economists since at least Marshall (1890). Empirical studies suggest that larger cities reward cognitive and people skills rather than motor skills or physical strength (Bacolod, Blum, and Strange 2009; Michaels, Rauch, and Redding forthcoming). Physical proximity is associated with increased communication and intellectual interaction (Jaffe, Trajtenberg, and Henderson 1993; Gaspar and Glaeser 1998; Audretsch and Feldman 2004; Charlot and Duranton 2004; Arzaghi and Henderson 2008). Since much knowledge is tacit and requires face-to-face transmission, we treat cities as the loci of idea exchange.

Our model unites two strands of theoretical literature on the exchange of ideas. One focuses on individuals' spatial choices when knowledge spillovers are exogenous externalities (Henderson 1974, Black 1999, Lucas 2001). Another focuses on choices of learning activities within a single location of exogenous population (Jovanovic and Rob 1989; Helsley and Strange 2004; Berliant, Reed, and Wang 2006; Berliant and Fujita 2008). In our model, locational choices shape idea

[5] For example, Glaeser (2008) and Beaudry, Doms, and Lewis (2010) use skill-segmented housing markets and skill-biased housing supplies to explain spatial variation in skill premia. These neoclassical models do not relate skill premia to city sizes. Gyourko, Mayer, and Sinai (2013) model exogenous differences in housing supply elasticities.

[6] In their model, all within-city inequality is due to exogenous shocks that individuals experience after selecting their location. Educational attainment is neither randomly assigned nor location-specific.

exchanges because learning opportunities are heterogeneous and depend upon the time-allocation decisions of local participants.[7] Our characterization of idea exchanges is simple compared to the second strand of literature, but this allows us to tractably model endogenous exchanges of ideas in a system of cities.

We focus on the exchange of ideas between rather than within firms. Idea exchange within firms is surely important, but it does not motivate firms to locate in cities, since intrafirm idea exchange may occur in geographic isolation. Our model describes interfirm interactions because these are the idea exchanges that may underpin urban agglomeration.[8]

I. A Spatial Knowledge Economy

The economy consists of a continuum of individuals of mass L, whose heterogeneous abilities are indexed by z and distributed with density $\mu(z)$ on connected support on $\mathbb{R}_+$. There are a number of homogeneous sites that may be cities with endogenous population and ability composition.

A. *Preferences and Production*

Individuals consume three goods: tradables, non-tradable services, and (non-tradable) housing. Services and housing are strict necessities; after consuming $\overline{n}$ units of non-tradable services and one unit of housing, consumers spend all of their remaining income on tradables, which we use as the numéraire.[9] Therefore, the indirect utility function for a consumer with income y facing prices $p_{n,c}$ and $p_{h,c}$ in city c is

$$V(p_{n,c}, p_{h,c}, y) = y - p_{n,c}\overline{n} - p_{h,c}. \tag{1}$$

Individuals are perfectly mobile across cities and jobs, so their locational and occupational choices maximize $V(p_{n,c}, p_{h,c}, y)$.

An individual can produce tradables (t) or non-tradables (n). Non-tradables can be produced at a uniform level of productivity by all individuals, which we normalize to 1 by choice of units. Tradables, by contrast, make use of the underlying heterogeneity in ability. An individual's tradables output is $\tilde{z}(z, Z_c)$, which depends on both individual ability z and learning opportunities available through local interactions,

[7] Glaeser (1999) is an important precursor to our approach. His model specifies two locations, a city and a rural hinterland. In contrast to our approach, the fundamental difference between the two locations is exogenous, since learning is possible only in the city.

[8] Recent research suggests that physical proximity facilitates such activities. Allen, Raz, and Gloor (2010) examine interfirm communication amongst individual scientists at biotech firms in the Boston area and find that geographic proximity and firm size are both positively associated with interfirm communication on the extensive and intensive margins. Inoue, Nakajima, and Saito (2015) examine interfirm collaboration on Japanese patent applications and find that it is more geographically concentrated than intrafirm collaboration and this localization is stable over the last two decades.

[9] The merit of this stark specification is tractability. Section IIC shows that our assumption of perfectly inelastic demand for housing and services generates a compensation effect that in fact works against finding a positive premia-population relationship. Section IID shows that this inelastic specification is nonetheless consistent with realistic housing expenditure shares. Assuming unit demand for housing is common in urban theory (e.g., Moretti 2011; Behrens, Duranton, and Robert-Nicoud 2014).

Z_c. An individual working in sector σ earns income equal to the value of her output, which is

$$y = \begin{cases} p_{n,c} & \text{if } \sigma = n \\ \tilde{z}(z, Z_c) & \text{if } \sigma = t \end{cases}. \tag{2}$$

Tradables productivity depends both on an individual's ability and participation in idea exchanges. Tradables producers can raise their productivity by exchanging ideas with other tradables producers in their city.[10] Each person has one unit of time that they divide between interacting and producing. Exchanging ideas is an economic decision, because time spent interacting trades off with time spent producing output directly. The production function for tradables B depends on time spent exchanging ideas $(1-\beta)$, time spent producing (β), own ability (z), and local learning opportunities (Z_c). When time is allocated optimally, the output of a tradables producer of ability z is

$$\tilde{z}(z, Z_c) = \max_{\beta \in [0,1]} B(1-\beta, z, Z_c). \tag{3}$$

The value of local idea exchanges, Z_c, is determined by the time-allocation decisions of all the tradables producers living in city c. In particular, it is a function of both the time they devote to exchanges and their abilities.[11] We denote the time devoted to idea exchange by individuals of ability z in city c by $1-\beta_{z,c}$ and the population of individuals of ability z in city c by $L \cdot \mu(z,c)$, where $\mu(z,c)/\mu(z)$ is the share of z-ability individuals who live in c. The value of the local idea-exchange environment Z_c is a functional of the time-allocation decisions and the population composition,

$$Z_c = Z(\{1-\beta_{z,c}\}, \{L \cdot \mu(z,c)\}). \tag{4}$$

Key to the tractability of our model is that these local learning opportunities are summarized by a scalar. To further summarize behavior, denote the total time devoted to learning by tradables producers in city c by M_c, which is defined as

$$M_c = L \int_{z:\sigma(z)=t} (1-\beta_{z,c})\, \mu(z,c)\, dz.$$

This depends on the city's total population of tradables producers, their ability composition, and their time-allocation choices.

We make three assumptions about the production of tradables and exchange of ideas. First, tradables output is increasing in both ability z and the idea-exchange environment Z_c. In the absence of idea-exchange opportunities $(Z_c = 0)$ or time

[10] Our static model focuses on the location of idea exchange and abstracts from dynamic accumulation of knowledge. Lucas and Moll (2014) and Perla and Tonetti (2014) study knowledge accumulation while abstracting from the spatial dimension. We hope that future work might unify these topics.

[11] These elements distinguish our agglomeration mechanism from a "black box" function that depends on a location's population size. Heterogeneous individuals' choices of locations and time allocations are economic decisions that both have opportunity costs and determine local opportunities for idea exchange.

devoted to them ($\beta = 1$), tradables output is the product of time spent producing and ability, βz.

ASSUMPTION 1: *The production function for tradables* $B(1-\beta, z, Z_c)$ *is continuous, strictly concave in* $1-\beta$, *strictly increasing in* z, *and increasing in* Z_c. $B(1-\beta, z, 0) = \beta z$ *and* $B(0, z, Z_c) = z \; \forall z$.

Second, individual ability and local learning opportunities are complements.[12] When exchanging ideas, the output gain from greater ability is increasing in Z_c.

ASSUMPTION 2: *Tradables output* $\tilde{z}(z, Z_c)$ *is supermodular and is strictly supermodular on* $\otimes \equiv \{(z, Z) : \tilde{z}(z, Z) > z\}$.

Third, the idea-exchange environment Z_c is better when those devoting time to idea exchange are of higher ability and when all tradables producers devote more time to idea exchange.

ASSUMPTION 3: *The idea-exchange functional* $Z(\{1-\beta_{z,c}\}, \{L \cdot \mu(z,c)\})$ *is continuous, equal to* 0 *if* $M_c = 0$, *and bounded above by* $\sup\{z : 1-\beta_{z,c} > 0, \mu(z,c) > 0\}$. *If* $M_c > M_{c'}$ *and* $\{(1-\beta_{z,c})\mu(z,c)\}$ *stochastically dominates* $\{(1-\beta_{z,c'})\mu(z,c')\}$, *then* $Z(\{1-\beta_{z,c}\}, \{L \cdot \mu(z,c)\}) > Z(\{1-\beta_{z,c'}\}, \{L \cdot \mu(z,c')\})$.

Assumption 3 implies that knowledge has both horizontal and vertical dimensions. There is horizontal differentiation in the sense that individuals can learn something from anyone and are therefore better off when all tradables producers devote more time to exchange. Vertical differentiation means that they learn more from more able counterparts.

For some of our analysis, we focus on particular functional forms for $B(\cdot)$ and $Z(\cdot)$:

$$B(1-\beta, z, Z_c) = \beta z(1 + (1-\beta) A Z_c z), \tag{5}$$

$$Z(\{1-\beta_{z,c}\}, \{L \cdot \mu(z,c)\}) = \big(1 - \exp(-\nu M_c)\big)\,\overline{z}_c, \tag{6}$$

$$\overline{z}_c = \begin{cases} \displaystyle\int_{z:\sigma(z)=t} \frac{z(1-\beta_{z,c})\,\mu(z,c)}{\int_{z:\sigma(z)=t}(1-\beta_{z,c})\,\mu(z,c)\,dz}\,dz & \text{if } M_c > 0 \\ 0 & \text{otherwise} \end{cases}.$$

Online Appendix A.6 shows that these functions satisfy Assumptions 1–3.

In this special case, productivity gains are the product of random matches between individuals devoting time to idea exchange. The term A indexes the scope for gains

[12] Assumption 2 is stated in terms of $\tilde{z}(z, Z_c)$, which incorporates the optimal choice of $1-\beta$. In online Appendix A.5, we identify sufficient conditions for $B(1-\beta, z, Z_c)$ that imply Assumption 2. This Assumption 2′ is sufficient but not necessary. For example, equation (5) satisfies Assumption 2 but not Assumption 2′.

from such interactions. With random matching, the expected value of devoting a moment of time to idea exchange in a city is the probability of encountering another individual during that moment times the expected ability of the individual encountered.[13] Since idea exchanges are instantaneous and individuals devote an interval of time to idea exchange, every individual devoting time to exchanging ideas realizes the expected gains from these exchanges, Z_c.

The probability of encountering someone during each moment of time spent seeking idea exchanges is $1 - \exp(-\nu M_c)$. As in Diamond (1982), exchange is easier when there are more potential exchange partners. Thus, one potential benefit of larger cities is that idea exchanges may occur with greater frequency there (Glaeser 1999).[14] The population of individuals available for such encounters is determined endogenously by tradables producers' time-allocation choices.

The average ability of the individuals encountered in these matches is $\bar{z}_c$. This is a weighted average of the abilities of local tradables producers, in which the weights are the time each type of individual devotes to interactions. Conditional on meeting another learner and one's own ability, conversations with more able individuals are more valuable.

The agglomeration mechanism described by Assumptions 1–3 trades off with a simple congestion force. Each individual in a city of population L_c pays a net urban cost (in units of the numéraire) of

$$p_{h,c} = \theta L_c^{\gamma}, \tag{7}$$

with $\theta, \gamma > 0$. We will refer to $p_{h,c}$ as the price of housing in city c, though this object incorporates both land rents and commuting costs when given standard microfoundations.[15]

[13] Random matching is not particularly realistic, but it is tractable. Related work in growth theory, Lucas and Moll (2014) and Perla and Tonetti (2014), also assumes random meetings. In those models, lower-ability agents spend time in order to observe a random competitor and improve their productivity through imitation of higher-ability agents. In our model, both participants in a meeting opt to spend time exchanging ideas. The benefit of the meeting can be interpreted as combining both what one can learn from the other and one's ability to learn from any encounter. Thus, even when participants in a meeting have different skill levels, both benefit. By contrast, those growth-theory papers feature meetings that benefit only the lower-ability participant, who imitates the higher-ability participant. Such perfectly asymmetric accounts of idea exchange are poorly suited as a foundation for spatial models. The highest-ability agent would never pay to be in a city where meetings occurs because idea exchanges with lower-ability agents yield zero benefit. By iteration, every agent would want to move to locations where higher-ability agents reside, while the latter would leave locations populated by lower-ability agents. Our approach in which idea exchange is mutually beneficial provides a framework consistent with spatial equilibrium.

[14] This city-level scale effect embodies the horizontal dimension of knowledge in Assumption 3 and implies an upper bound on the number of heterogeneous cities. See online Appendices A.2 and A.6.1. Most empirical evidence on matching processes describes job search, which is distinct from idea exchange in numerous dimensions. Early job-search studies, while noisy, were often interpreted to suggest constant returns (Petrongolo and Pissarides 2001). More recent studies have found results more favorable to increasing returns to scale (Petrongolo and Pissarides 2006, Di Addario 2011, Bleakley and Lin 2012).

[15] Behrens, Duranton, and Robert-Nicoud (2014) provide microeconomic foundations for this functional form, which they derive from a standard model of the internal structure of a monocentric city in which commuting costs increase with population size as governed by the technological parameters θ and γ. See online Appendix A.1 for details.

B. *Equilibrium*

Individuals choose their locations, occupations, and time allocations optimally. Since individuals are perfectly mobile, two individuals with the same ability z will obtain the same utility in equilibrium wherever they are located.

An equilibrium for a population L with ability distribution $\mu(z)$ in a set of locations $\{c\}$ is a set of prices $\{p_{h,c}, p_{n,c}\}$ and populations $\mu(z,c)$ such that workers maximize (1) by their choices of c, σ, and β and markets clear.[16] Markets clear when $\beta_{z,c} = \beta(z, Z_c)$, and equations (4), (7), and the following conditions hold:

$$\mu(z) = \sum_c \mu(z,c) \quad \forall z, \tag{8}$$

$$L_c = L \int \mu(z,c)\, dz \quad \forall c, \tag{9}$$

$$\bar{n} L_c = L \int_{z:\sigma(z)=n} \mu(z,c)\, dz \quad \forall c. \tag{10}$$

The equilibrium value of local idea exchanges $Z_c = Z(\{1-\beta_{z,c}\}, \{L \cdot \mu(z,c)\})$ in equation (4) is a fixed point, since individuals' choices of location and $\beta_{z,c}$ depend on local learning opportunities Z_c. Equation (7) defines the market-clearing housing price in each city. Equations (8) and (9) are adding-up constraints for worker types and city populations. Equation (10) equalizes demand and supply of non-tradable services within each location. The tradables market clears by Walras' Law.

II. The Cross Section of Cities in Equilibrium

We now characterize equilibrium outcomes. By comparative advantage, higher-ability individuals produce tradables and lower-ability individuals produce non-tradables. Since individual ability and idea-exchange opportunities are complements, there is spatial sorting of tradables producers by ability. When cities are heterogeneous, tradables producers are sequentially segmented so that the lowest-ability tradables producers locate in the smallest city and the highest-ability tradables producers locate in the largest city.[17] For the marginal tradables producer, a larger city's idea-exchange benefits are exactly offset by its higher cost of living. Together, sorting and idea exchange cause larger cities to exhibit higher skill premia.

A. *Equilibrium Occupations and Prices*

Occupational choices are governed by comparative advantage. High-ability individuals produce tradables since labor heterogeneity matters in that sector.

[16] In this exposition, we define equilibrium where each member of the set $\{c\}$ is populated, $L_c > 0$. In online Appendix A.2, we discuss the endogenous number of cities that make up this set, since not all potential city locations must be populated.

[17] Spatial models with a continuum of heterogeneous individuals sorting across a finite number of locations date to at least Westhoff (1977), who studied conditions for the existence of an equilibrium in a model of local public finance.

LEMMA 1 (Comparative Advantage): *Suppose that Assumption 1 holds. There is an ability level* z_m *such that individuals of greater ability produce tradables and individuals of lesser ability produce non-tradables*:

$$\sigma(z) = \begin{cases} t & \text{if } z > z_m \\ n & \text{if } z < z_m \end{cases}.$$

The proofs of Lemma 1 and subsequent results appear in online Appendix A.6. By Lemma 1 and equations (8) through (10), the ability level of the individual indifferent between producing tradables and non-tradables, z_m, is given by

$$\int_0^{z_m} \mu(z)\, dz = \overline{n}. \tag{11}$$

Since individual ability and local learning opportunities are complements, there is spatial sorting of tradables producers engaged in idea exchange. Higher-ability tradables producers locate in cities with better idea-exchange environments.

LEMMA 2 (Spatial Sorting of Tradables Producers Engaged in Idea Exchange): *Suppose that Assumptions 1 and 2 hold. For* $z > z' > z_m$, *if* $\mu(z,c) > 0$, $\mu(z',c') > 0$, $\beta(z,Z_c) < 1$, *and* $\beta(z',Z_{c'}) < 1$, *then* $Z_c \geq Z_{c'}$.

As a result, individuals of ability z_m producing tradables will be located in the city with the lowest value of Z_c. Label cities in order of the value of their idea exchanges, $Z_c \geq Z_{c-1}$, so that $Z_1 = \min_c\{Z_c\}$. Indifference between producing tradables and non-tradables implies that $p_{n,1}$ satisfies

$$p_{n,1} = \tilde{z}(z_m, Z_1). \tag{12}$$

There is a population of non-tradables producers located in each city. In spatial equilibrium, each of these individuals obtains the same utility, so equation (1) implies that spatial differences in non-tradables prices exactly compensate for spatial differences in housing prices:

$$(1-\overline{n})\,p_{n,c} - p_{h,c} = (1-\overline{n})\,p_{n,c'} - p_{h,c'} \quad \forall c, c'. \tag{13}$$

All equilibria exhibit this pattern of occupations and prices. We now distinguish between equilibria based on whether cities vary in size.

B. *Equilibrium Systems of Cities*

There are two classes of equilibria for this economy: equilibria in which all cities have the same population size and equilibria with heterogeneous cities. The latter are the empirically relevant class. We analyze the properties of equilibria with heterogeneous cities after describing why systems of equal-sized cities are only stable equilibria if the marginal gains from idea exchange are too small relative to marginal congestion costs to break the symmetric arrangement. When idea exchange is sufficiently rewarding, a system of heterogeneous cities is an equilibrium configuration.[18]

Systems of Equal-Sized Cities.—Given symmetric fundamentals, systems of equal-sized cities are possible equilibria. By equations (7) and (13), equal-sized cities have equal local prices. To be in equilibrium, they must also have equal idea-exchange benefits for the marginal tradables producer.

When idea exchange occurs nowhere, $Z_c = 0\ \forall c$, these benefits are equal because every tradables producer devotes zero time to idea exchange. This is individually rational when others do the same. While not the focus of our paper, the no-idea-exchange equilibrium illustrates an important aspect of the economic mechanism: ideas are not manna from heaven but the outcome of a costly allocation of time by those acquiring knowledge. Though not the empirically relevant case, this possibility highlights the relevant economic trade-off.

A system of equal-sized cities in which idea exchange occurs can only be a stable equilibrium when the marginal benefits of a better idea-exchange environment are small relative to marginal congestion costs. Denote the city with the best idea-exchange environment by C. Given identical prices, living in C is optimal for all individuals of ability such that $\beta(z, Z_C) < 1$. If C is the only city with idea exchanges, this is an equilibrium only if both Z_C and the ability distribution are so low that all those possibly gaining from idea exchange fit in this single city and the marginal tradables producer's idea-exchange benefit is zero. If another city also has idea exchanges, the value of its idea-exchange environment must be the same. An equilibrium with two cities with equal idea-exchange environments is locally stable only if the marginal gains from idea exchange are small relative to marginal congestion costs. Otherwise, the movement of some high-ability tradables producers from one city to the other would improve the latter's idea-exchange environment, thereby drawing in more tradables producers and breaking the symmetric arrangement.[19] In particular, if any improvement in the idea-exchange environment is sufficiently valuable to the highest-ability tradables producers, then an equilibrium of equal-sized cities cannot be stable.[20]

Thus, a system of equal-sized cities is a stable equilibrium in the trivial case that no one exchanges ideas or if marginal congestion costs are large relative to the benefits of a better idea-exchange environment. We henceforth focus on the empirically relevant equilibrium configuration, a system of heterogeneous cities.

Systems of Heterogeneous Cities.—Equilibria with heterogeneous cities exhibit cross-city patterns that can be established independent of the number of cities that arise.[21] Proposition 1 characterizes the characteristics of heterogeneous cities in equilibrium.

[18] We provide sufficient conditions for the existence of a stable equilibrium with two heterogeneous cities in online Appendix A.3. While we do not provide conditions for the existence of equilibria with more than two cities, we have numerically computed equilibria with many more cities. Section IID reports an example with 275 cities and online Appendix B summarizes many more numerical results.

[19] See online Appendix A.4 for our definition of local stability and the relevant argument.

[20] See Proposition 3 in online Appendix A.4 for the formal result.

[21] Since these patterns characterize all equilibria with heterogeneous cities, we do not address issues of uniqueness or determine the equilibrium number of cities. See online Appendix A.2 for further discussion.

PROPOSITION 1 (Heterogeneous Cities' Characteristics): *Suppose that Assumptions 1 and 2 hold. In any equilibrium, a larger city has higher housing prices, higher non-tradables prices, a better idea-exchange environment, and higher-ability tradables producers. If* $L_c > L_{c'}$ *in equilibrium, then* $p_{h,c} > p_{h,c'}$, $p_{n,c} > p_{n,c'}$, $Z_c > Z_{c'}$, *and* $z > z' > z_m \Rightarrow \mu(z,c)\,\mu(z',c') \geq \mu(z,c')\,\mu(z',c) = 0$.

The mechanics of Proposition 1 are straightforward. Larger cities have higher housing prices due to congestion, so non-tradables producers require higher wages in these locations. Larger cities attract tradables producers because the benefits of more valuable idea exchanges offset their higher housing and non-tradables prices. More able tradables producers benefit more from participating in better idea exchanges, so there is spatial sorting of tradables producers.[22] This spatial sorting supports equilibrium differences in idea-exchange environments because these high-ability individuals are better idea-exchange partners, conditional on population size and time allocations.

Equilibria with heterogeneous cities match the fundamental facts that cities differ in size and these size differences are accompanied by differences in wages, housing prices, and productivity (Glaeser 2008). Empirically, larger cities exhibit higher nominal wages in industries that produce tradable goods, which means that productivity is higher in these locations (Moretti 2011). Our model of why larger cities generate more productivity-increasing idea exchanges is a microfounded explanation of these phenomena. Having matched these well-established facts, we now describe the novel implication that skill premia will be higher in larger cities.

C. *Skill Premia with Heterogeneous Cities*

We define a city's skill premium as the average income of its tradable producers relative to the wage of its non-tradable producers. Our model typically predicts that skill premia are higher in more populous cities. After discussing the mechanisms contributing to spatial variation in skill premia, we formally state this prediction for two cities in Proposition 2. Numerical analysis, detailed in online Appendix B, suggests that this prediction generalizes from two cities to a large number of heterogeneous cities.

The nominal wages of both non-tradables and tradables producers are higher in larger cities. For non-tradables producers, higher nominal wages in larger cities are compensation for higher housing prices that keeps their utility constant across cities, per equation (13).

Differences in tradables producers' wages across cities can be expressed as the sum of three components: composition, learning, and compensation effects. First, due to spatial sorting, tradables producers in larger cities have higher innate abilities that generate higher incomes in any location. Second, since one's own ability complements others' abilities in idea exchanges, these tradables producers realize

[22] Our prediction of sorting among the more able and indifference among the less able is consistent with the limited evidence available. Using Norwegian administrative data, Carlsen, Rattsø, and Stokke (2016) estimate unobserved ability with individual fixed effects in wage regressions and find sorting among college-educated workers but none among those with primary and secondary education.

larger income gains in larger cities' better idea-exchange environments. Third, producers who are indifferent between two cities realize learning gains in the larger city that exactly compensate for its higher non-tradables and housing prices. For convenience, let z_b denote the ability of this boundary tradables producer who is indifferent, define inframarginal learning $\Delta(z, c, c')$ as the idea-exchange gains accruing to a producer of ability z from locating in environment $Z_c > Z_{c'}$ compared to those gains for ability z_b, and define the density of tradables producers' abilities in city c by $\tilde{\mu}(z, c)$.[23] When a tradables producer of ability z_b is indifferent between cities c and c', the difference in the cities' average tradables wages can be expressed as

$$\overline{w}_c - \overline{w}_{c'} \equiv \frac{\int_{z:\sigma(z)=t} \tilde{z}(z, Z_c)\mu(z, c)\,dz}{\int_{z:\sigma(z)=t} \mu(z, c)\,dz} - \frac{\int_{z:\sigma(z)=t} \tilde{z}(z, Z_{c'})\,\mu(z, c')\,dz}{\int_{z:\sigma(z)=t} \mu(z, c')\,dz}$$

$$= \underbrace{\int_{z_m}^{\infty} [\tilde{\mu}(z, c) - \tilde{\mu}(z, c')]\tilde{z}(z, Z_{c'})\,dz}_{\text{composition}} + \underbrace{\int_{z_m}^{\infty} \tilde{\mu}(z, c)\,\Delta(z, c, c')\,dz}_{\text{inframarginal learning}} + \underbrace{p_{n,c} - p_{n,c'}}_{\text{compensation}}.$$

Cross-city variation in skill premia can also be expressed in terms of these three components. We define a city's observed skill premium as its average tradables wage divided by its (common) non-tradables wage, $\overline{w}_c/p_{n,c}$. When a tradables producer of ability z_b is indifferent between cities c and c', this skill premium is higher in c if and only if

$$(14) \qquad \underbrace{\int_{z_m}^{\infty} [\tilde{\mu}(z, c) - \tilde{\mu}(z, c')]\tilde{z}(z, Z_{c'})\,dz}_{\text{composition}} + \underbrace{\int_{z_m}^{\infty} \tilde{\mu}(z, c)\,\Delta(z, c, c')\,dz}_{\text{inframarginal learning}} \geq \underbrace{(p_{n,c} - p_{n,c'})\left(\frac{\overline{w}_{c'}}{p_{n,c'}} - 1\right)}_{\text{relative compensation}}.$$

The composition and inframarginal learning effects yield higher nominal incomes for tradables producers in larger cities. These raise tradables producers' wages relative to non-tradables producers' wages in larger cities and therefore generate a positive premium-population relationship. The compensation effect that reflects differences in local prices makes the nominal wages of both tradables and non-tradables producers in larger cities higher by the same amount. Since higher-ability individuals earn higher incomes, this compensation is a larger proportion of the non-tradables producers' incomes and therefore pushes toward a negative premium-population relationship. When the composition and learning effects dominate this implication of the compensation effect, the skill premium is higher in the larger city.

The sizes of these three effects depend on the distribution of abilities, $\mu(z)$, the strength of the complementarity between z and Z_c in $\tilde{z}(z, Z_c)$, and equilibrium differences in cities' sizes. The composition and inframarginal learning effects necessarily

[23] That is, $\Delta(z, c, c') \equiv [\tilde{z}(z, Z_c) - \tilde{z}(z, Z_{c'})] - [\tilde{z}(z_b, Z_c) - \tilde{z}(z_b, Z_{c'})]$ and $\tilde{\mu}(z, c) \equiv \frac{\mu(z, c)}{\int_{z':\sigma(z')=t} \mu(z', c)\,dz'}$.

depend on heterogeneity in tradables producers' abilities.[24] Inframarginal learning also depends on the degree to which higher-ability individuals experience larger gains from locating in a better idea-exchange environment. The relative compensation effect is the product of size-related differences in costs of living $\left(p_{n,c} - p_{n,c'}\right)$ and the level of the skill premium $\left(\overline{w}_{c'}/p_{n,c'}\right)$.[25]

Proposition 2 states three different sets of sufficient conditions under which, when the smallest city has population L_1 and the second-smallest city has population $L_2 > L_1$, the skill premium is higher in the more populous city. These sufficient conditions depend jointly on assumptions about $\mu(z)$, $\tilde{z}(z, Z_c)$, and equilibrium city sizes. For the production function, the relevant property concerns the ability elasticity of tradable output, $\partial \ln \tilde{z}(z, Z_c)/\partial \ln z$.

CONDITION 1: *The ability elasticity of tradable output is non-decreasing in z and Z_c.*

Condition 1 is satisfied by the production function in equation (5), as well as other production functions, as described in online Appendix A.5.

PROPOSITION 2 (Skill Premia): *Suppose that Assumptions 1 and 2 hold. In an equilibrium in which the smallest city has population L_1 and the second-smallest city has population $L_2 > L_1$*

(*i*) *If the ability distribution is decreasing, $\mu'(z) \leq 0$, $\tilde{z}(z, Z_c)$ is log-convex in z, and $\tilde{z}(z, Z_c)$ is log-supermodular, then $\overline{w}_2/p_{n,2} > \overline{w}_1/p_{n,1}$;*

(*ii*) *If the ability distribution is Pareto, $\mu(z) \propto z^{-k-1}$ for $z \geq z_{\min}$ and $k > 0$, and the production function satisfies Condition 1, then $\overline{w}_2/p_{n,2} > \overline{w}_1/p_{n,1}$;*

(*iii*) *If the ability distribution is uniform, $z \sim U\left(z_{\min}, z_{\max}\right)$, the production function satisfies Condition 1, and $\frac{L_2 - L_1}{L_1^2} > \frac{1}{L}\frac{(1-\overline{n})(z_{\max} - z_{\min})}{z_{\min} + \overline{n}(z_{\max} - z_{\min})}$, then $\overline{w}_2/p_{n,2} > \overline{w}_1/p_{n,1}$.*

These three cases trade off stronger assumptions about the production function with weaker assumptions about the ability distribution. In case (i), the log-supermodularity of $\tilde{z}$ implies large inframarginal learning, and the log convexity of $\tilde{z}$ implies a large composition effect when higher-ability individuals are not relatively abundant. Together, these are sufficient to dominate the relative compensation effect, such that the skill premium is higher in the larger city. In case (ii), the Pareto ability distribution and the ability elasticity of tradable output jointly

[24] In a model with only two skill types (i.e., homogeneous tradables producers and homogeneous non-tradables producers), the skill premium is lower in the larger city. Homogeneity makes the composition and inframarginal learning components zero, leaving only the compensation term. This two-type case is the basis for the prediction by Black, Kolesnikova, and Taylor (2009) that skill premia will be lower in cities with higher housing prices. Empirically, larger cities have both higher housing prices and skill premia.

[25] Since $p_{n,1} = \tilde{z}(z_m, Z_1)$, the level of the skill premium in city 1 depends on the heterogeneity in tradables producers' abilities, $\mu(z, 1)$, and the complementarity between z and Z_1 in $\tilde{z}(z, Z_1)$.

generate composition and inframarginal-learning effects sufficient to dominate the relative compensation effect. In case (iii), the uniform ability distribution generates weaker compensation and inframarginal-learning effects for a given ability elasticity of tradable output. The sufficient condition in case (iii) is written in terms of endogenous equilibrium outcomes. It establishes a value of L_1 small enough relative to the heterogeneity in tradables producers' abilities such that the relative compensation effect is less than these effects.[26] Note that this condition is far from necessary, as demonstrated in online Appendix B.1. The relative compensation effect approaches 0 as $L_1 \to L_2$ because $p_{n,2} - p_{n,1} \to 0$.

Since the sufficient condition in case (iii) depends on endogenous city sizes, we study the two-city, uniform-ability case further by numerically characterizing equilibria for the special case of equations (5) and (6) for a wide range of parameter values in online Appendix B. We do find examples of parameter combinations that yield equilibria in which a larger city has a lower skill premium, but they are rare (less than 0.3 percent of the parameter combinations for which equilibria exist). These examples generate very large relative compensation effects (the right-hand side of inequality (14)) by generating values of $\overline{w}_1/p_{n,1}$ an order of magnitude larger than those in the data. They also require values of γ, the congestion cost elasticity, that are implausibly large relative to empirical estimates, though non-increasing skill premia are atypical in equilibrium even for extreme parameter values.[27]

To extend the prediction of Proposition 2 to more than two cities, in online Appendix B we numerically solve the special case of equations (5) and (6) for a wide number of heterogeneous cities and wide range of parameter values for the uniform- and Pareto-ability cases. In the uniform-ability case, we again find that skill premia are monotonically increasing in population size in almost all equilibria. The exceptions to this pattern occur when $\overline{w}_1/p_{n,1}$ and γ are implausibly large relative to empirical values, consistent with the two-city case. In the Pareto-ability case, we find that all equilibria examined exhibit monotonically increasing skill premia.[28] Thus, our two-city result appears to generalize to many cities.

To summarize, our model typically predicts that skill premia are higher in more populous cities, in line with the empirical pattern documented in Table 1. Proposition 2 analytically characterizes the pattern of premia for the two smallest cities in an equilibrium, and numerical computations reported in online Appendix B show that this pattern of premia generalizes across all cities in equilibrium for a wide range of parameter values. This novel prediction distinguishes our model from the canonical spatial-equilibrium model, which predicts spatially invariant skill premia.

[26] This sufficient condition can also be written as $(L_2 - L_1)/L_1^2 > (z_{\max} - z_m)/(L \times z_m)$.

[27] For example, we find non-increasing skill premia in parameter combinations in which $\gamma = 5$, but less than 0.5 percent of parameter combinations with $\gamma = 5$ for which equilibria exist exhibit non-increasing skill premia.

[28] The numerical findings for hundreds of thousands of parameter values are strongly suggestive. Unfortunately, our analytical proof of the two-city result for a Pareto ability distribution does not extend naturally to an arbitrary number of cities.

D. *An Illustrative Example with 275 Cities*

To illustrate our model's capacity to match empirical patterns linking cities' sizes, wages, and prices, we report an example of an equilibrium for the special case of equations (5) and (6) that is consistent with 3 empirical moments of interest. First, Zipf's law says that a city's size rank is inversely proportional to its size (Gabaix 1999). Second, there is the positive correlation between skill premia and population size documented in Table 1. Third, Davis and Ortalo-Magné (2011) document that housing expenditure shares vary little across cities.[29]

Our illustrative example is a uniform-ability equilibrium with 275 heterogeneous cities, akin to the number of US metropolitan areas.[30] The exogenous parameter values are $A = 3, \overline{n} = 0.4, \theta = 1, \gamma = 0.1, L = 2{,}062.5, \nu = 50, z_{\min} = 1, z_{\max} = 2$. The large values of ν and L make the equilibrium matching rate $m(M_c)$ exceed 0.99 in every city, so cross-city variation in idea-exchange environments is much more due to variation in average ability $\overline{z}_c$ than scale effects. The average time devoted to idea exchange ranges from 0.41 to 0.46 and is monotonically increasing in city size. Regressing log population rank on log size yields a coefficient of -1.025, near the typical empirical estimate of this power-law exponent. Regressing log skill premium on log size yields a coefficient of 0.092, which is greater than those reported in Table 1 but plausible. Housing expenditure shares vary from 0.32 to 0.36 and have a population elasticity reasonably close to 0, -0.023. Thus, this illustrative example exhibits properties consistent with empirical patterns.[31]

While our model does not yield closed-form comparative statics, this illustrative example exhibits local comparative statics consistent with economic shifts in recent decades. Work in labor economics has emphasized skill-biased technical change, which we interpret as an increase in A, as one reason for growth in the (economy-wide) skill premium (Acemoglu and Autor 2011). Around the illustrative equilibrium, a 10 percent increase in A leaves the power-law exponent virtually unchanged and increases both the economy-wide average skill premium and the population elasticity of skill premia by about 7 to 8 percent. Table C.2 shows that the population elasticity of skill premia did increase from 1990 to 2007. Thus, our model's mechanics are qualitatively consistent with and introduce a spatial dimension to the leading explanation for changes in the skill premium in recent decades.

III. Conclusion

The presence of skyscrapers is a defining characteristic of cities' central business districts. These attest to an intense desire to concentrate large numbers of people in

[29] In light of their finding, Davis and Ortalo-Magné (2011) use Cobb-Douglas preferences. Our results show that such preferences are not necessary to obtain housing expenditure shares that are approximately spatially invariant, as previously established by Behrens, Duranton, and Robert-Nicoud (2014).

[30] The geographic delineations used in Census 2000 publications define 280 (consolidated) metropolitan statistical areas, including 4 in Puerto Rico.

[31] While this example matches empirical regularities well, our model does not generically generate these patterns. Other parameterizations can yield equilibria in which the power-law exponent is quite far from -1. In our numerical explorations, small changes to the parameter values cause small changes in the equilibrium distribution of outcomes and all eight parameters are quantitatively important to the joint determination of outcomes such as the power-law exponent.

a tiny geography. This extreme concentration is neither to exchange goods nor to facilitate hiring. One benefit of this concentration is that it facilitates idea exchange. While idea exchange within firms is surely of great importance, an individual firm need not pay the costs to be in a central business district for this benefit. Idea exchange outside the boundaries of the firm provides a foundation for agglomeration.

It is precisely this costly, voluntary interaction that we seek to capture in our model of idea exchange. In our theory, individuals allocate their time according to the expected gains from exchanging ideas in their city. The gains from idea exchange are greater in places where conversation partners are more numerous, devote more time to idea exchange, and are of higher ability. The highest-ability tradables producers reap the most from such learning opportunities. This simple setup, designed to overcome the "black box" critique that has inhibited research in this crucial area, nonetheless yields a rich set of spatial patterns. Larger cities are places with more idea exchanges between higher-ability participants, and they in turn exhibit higher wages, productivity, housing prices, and skill premia—all prominent features in the data.

This account suggests important implications for various aspects of urban policy that affect the city as a locus of idea exchange. Transportation policy determines the frequency with which meetings may feasibly occur. Zoning shapes not only the population density of potential participants but the venues in which idea exchanges may arise. And our model provides an account in which larger cities' higher nominal wage inequality does not imply that their lower-income residents have lower welfare than their counterparts in other locations.

Our static model characterizes the cross section of cities resulting from the complementarity between individual ability and idea-exchange opportunities. We thus provide a microfounded account of the spatial distribution of economic activity in a world in which cities are defined by the skills and ideas of those who choose to live in them. Future theoretical work might also capture the dynamics of knowledge accumulation and innovation, in light of the empirical evidence in Wang (2016); de la Roca and Puga (2017); and Carlsen, Rattsø, and Stokke (2016) that larger cities' benefits to high-ability individuals accrue over time.

REFERENCES

Abdel-Rahman, Hesham M., and Alex Anas. 2004. "Theories of Systems of Cities." In *Handbook of Regional and Urban Economics*, Vol. 4, edited by J. Vernon Henderson and Jacques-François Thisse, 2293–339. Amsterdam: Elsevier.

Acemoglu, Daron, and David Autor. 2011. "Skills, Tasks and Technologies: Implications for Employment and Earnings." In *Handbook of Labor Economics*, Vol. 4B, edited by Orley Ashenfelter and David Card, 1043–171. Amsterdam: Elsevier.

Allen, Thomas J., Ornit Raz, and Peter Gloor. 2010. "Does Geographic Clustering Still Benefit High Tech New Ventures? The Case of the Cambridge/Boston Biotech Cluster." In *Innovation in Complex Social Systems*, edited by Petra Ahrweiler, 73–87. London: Routledge.

Arzaghi, Mohammad, and J. Vernon Henderson. 2008. "Networking Off Madison Avenue." *Review of Economic Studies* 75 (4): 1011–38.

Audretsch, David B., and Maryann P. Feldman. 2004. "Knowledge Spillovers and the Geography of Innovation." In *Handbook of Regional and Urban Economics*, Vol. 4, edited by J. Vernon Henderson and Jacques-François Thisse, 2713–39. Amsterdam: Elsevier.

Bacolod, Marigee, Bernardo S. Blum, and William C. Strange. 2009. "Skills in the City." *Journal of Urban Economics* 65 (2): 136–53.

Baum-Snow, Nathaniel, and Ronni Pavan. 2013. "Inequality and City Size." *Review of Economics and Statistics* 95 (5): 1535–48.

Beaudry, Paul, Mark Doms, and Ethan Lewis. 2010. "Should the Personal Computer Be Considered a Technological Revolution? Evidence from U.S. Metropolitan Areas." *Journal of Political Economy* 118 (5): 988–1036.

Behrens, Kristian, Gilles Duranton, and Frederic Robert-Nicoud. 2014. "Productive Cities: Sorting, Selection, and Agglomeration." *Journal of Political Economy* 122 (3): 507–53.

Behrens, Kristian, and Frederic Robert-Nicoud. 2014. "Survival of the Fittest in Cities: Urbanisation and Inequality." *Economic Journal* 124 (581): 1371–400.

Berliant, Marcus, and Masahisa Fujita. 2008. "Knowledge Creation as a Square Dance on the Hilbert Cube." *International Economic Review* 49 (4): 1251–95.

Berliant, Marcus, Robert R. Reed, and Ping Wang. 2006. "Knowledge Exchange, Matching, and Agglomeration." *Journal of Urban Economics* 60 (1): 69–95.

Black, Dan, Natalia Kolesnikova, and Lowell Taylor. 2009. "Earnings Functions When Wages and Prices Vary by Location." *Journal of Labor Economics* 27 (1): 21–47.

Black, Duncan. 1999. "Local Knowledge Spillovers and Inequality." Unpublished.

Bleakley, Hoyt, and Jeffrey Lin. 2012. "Thick-Market Effects and Churning in the Labor Market: Evidence from US Cities." *Journal of Urban Economics* 72 (2–3): 87–103.

Carlsen, Fredrik, Jørn Rattsø, and Hildegunn E. Stokke. 2016. "Education, Experience, and Urban Wage Premium." *Regional Science and Urban Economics* 60: 39–49.

Charlot, Sylvie, and Gilles Duranton. 2004. "Communication Externalities in Cities." *Journal of Urban Economics* 56 (3): 581–613.

Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon. 2008. "Spatial Wage Disparities: Sorting Matters!" *Journal of Urban Economics* 63 (2): 723–42.

Davis, Donald R., and Jonathan I. Dingel. 2019. "A Spatial Knowledge Economy: Dataset." *American Economic Review.* https://doi.org/10.1257/aer.20130249.

Davis, Morris A., and Francois Ortalo-Magné. 2011. "Household Expenditures, Wages, Rents." *Review of Economic Dynamics* 14 (2): 248–61.

De La Roca, Jorge, and Diego Puga. 2017. "Learning by Working in Big Cities." *Review of Economic Studies* 84 (1): 106–42.

Di Addario, Sabrina. 2011. "Job Search in Thick Markets." *Journal of Urban Economics* 69 (3): 303–18.

Diamond, Peter A. 1982. "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy* 90 (5): 881–94.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000." *American Economic Review* 106 (3): 479–524.

Duranton, Gilles, and Diego Puga. 2004. "Micro-Foundations of Urban Agglomeration Economies." In *Handbook of Regional and Urban Economics*, Vol. 4, edited by J. Vernon Henderson and Jacques-François Thisse, 2063–117. Amsterdam: Elsevier.

Fujita, Masahisa, Paul Krugman, and Anthony J. Venables. 1999. *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge, MA: MIT Press.

Fujita, Masahisa, and Jacques-François Thisse. 2002. *Economics of Agglomeration: Cities, Industrial Location, and Regional Growth*. Cambridge, UK: Cambridge University Press.

Gabaix, Xavier. 1999. "Zipf's Law for Cities: An Explanation." *Quarterly Journal of Economics* 114 (3): 739–67.

Gaspar, Jess, and Edward L. Glaeser. 1998. "Information Technology and the Future of Cities." *Journal of Urban Economics* 43 (1): 136–56.

Gibbons, Stephen, Henry G. Overman, and Panu Pelkonen. 2014. "Area Disparities in Britain: Understanding the Contribution of People Vs. Place through Variance Decompositions." *Oxford Bulletin of Economics and Statistics* 76 (5): 745–63.

Glaeser, Edward L. 1999. "Learning in Cities." *Journal of Urban Economics* 46 (2): 254–77.

Glaeser, Edward L. 2008. *Cities, Agglomeration, and Spatial Equilibrium.* Cambridge, UK: Oxford University Press.

Glaeser, Edward L., Matt Resseger, and Kristina Tobio. 2009. "Inequality in Cities." *Journal of Regional Science* 49 (4): 617–46.

Gyourko, Joseph, Christopher Mayer, and Todd Sinai. 2013. "Superstar Cities." *American Economic Journal: Economic Policy* 5 (4): 167–99.

Helsley, Robert W., and William C. Strange. 2004. "Knowledge Barter in Cities." *Journal of Urban Economics* 56 (2): 327–45.

Henderson, J. V. 1974. "The Sizes and Types of Cities." *American Economic Review* 64 (4): 640–56.

Inoue, Hiroyasu, Kentaro Nakajima, and Yukiko Umeno Saito. 2015. "Localization of Collaborations in Knowledge Creation." Unpublished.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (3): 577–98.

Jovanovic, Boyan, and Rafael Rob. 1989. "The Growth and Diffusion of Knowledge." *Review of Economic Studies* 56 (4): 569–82.

Lucas, Robert E., Jr. 2001. "Externalities and Cities." *Review of Economic Dynamics* 4 (2): 245–74.

Lucas, Robert E., Jr., and Benjamin Moll. 2014. "Knowledge Growth and the Allocation of Time." *Journal of Political Economy* 122 (1): 1–51.

Marshall, Alfred. 1890. *Principles of Economics.* London: MacMillan.

Michaels, Guy, Ferdinand Rauch, and Stephen J Redding. Forthcoming. "Task Specialization in U.S. Cities from 1880 to 2000." *Journal of the European Economic Association*.

Moretti, Enrico. 2004. "Human Capital Externalities in Cities." In *Handbook of Regional and Urban Economics*, Vol. 4, edited by J. Vernon Henderson and Jacques-François Thisse, 2243–91. Amsterdam: Elsevier.

Moretti, Enrico. 2011. "Local Labor Markets." In *Handbook of Labor Economics*, Vol. 4B, edited by David Card and Orley Ashenfelter, 1237–313. Amsterdam: Elsevier.

Perla, Jesse, and Christopher Tonetti. 2014. "Equilibrium Imitation and Growth." *Journal of Political Economy* 122 (1): 52–76.

Petrongolo, Barbara, and Christopher Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature* 39 (2): 390–431.

Petrongolo, Barbara, and Christopher Pissarides. 2006. "Scale Effects in Markets with Search." *Economic Journal* 116 (508): 21–44.

Rauch, James E. 1993. "Productivity Gains from Geographic Concentration of Human Capital: Evidence from the Cities." *Journal of Urban Economics* 34 (3): 380–400.

Wang, Zhi. 2016. "Wage Growth, Ability Sorting, and Location Choice at Labor-Force Entry: New Evidence from U.S. Census Data." *Journal of Urban Economics* 96: 112–20.

Westhoff, Frank. 1977. "Existence of Equilibria in Economies with a Local Public Good." *Journal of Economic Theory* 14 (1): 84–112.

Wheeler, Christopher H. 2001. "Search, Sorting, and Urban Agglomeration." *Journal of Labor Economics* 19 (4): 879–99.