WORKING PAPER SERIES

Search and Zipf A Model of Frictio-nal Spatial Equilibrium

Benoît Schmutz, Modibo Sidibé

N°01 / Jan 2021

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/

Search and Zipf

A Model of Frictional Spatial Equilibrium*

Benoît Schmutz† Modibo Sidibé ‡

December 31, 2020

Abstract

This paper proposes a theory of cities based on a general equilibrium search and matching model of individual careers with an explicit spatial component. Heterogeneous firms and workers continuously decide where to locate within a set of imperfectly connected local labor markets and engage in wage bargaining using both local and remote match opportunities as threat points. A parsimonious model allows us to introduce the structural origins of workers' sorting, firms' selection and matching-based agglomeration economies into a unified framework and discuss their relationship with the city size distribution. Simulations show that power laws in city size do not require increasing returns to scale in matching or production, but may simply result from the combination of imperfect labor mobility, positive assortative matching between labor and capital, and agglomeration economies in the matching between workers and firms. By-products of the model include sufficient statistics to identify sorting and agglomeration using city-level variation and a rationale for the geographic diversity of urban networks.

Keywords: city size; local labor market; frictions; on-the-job search; migration
JEL Classification: R1; J2; J3; J6

*We thank Gilles Duranton, Jan Eeckhout, Xavier Gabaix, Grégory Jolivet, Isabelle Méjean, Fabien Postel-Vinay, Frédéric Robert-Nicoud and Hélène Turon as well as many seminar and conference participants for helpful comments and questions. Financial support from Duke University and CREST is gratefully acknowledged. Schmutz also thanks support from a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (Labex ECODEC No. ANR-11-LABEX-0047)

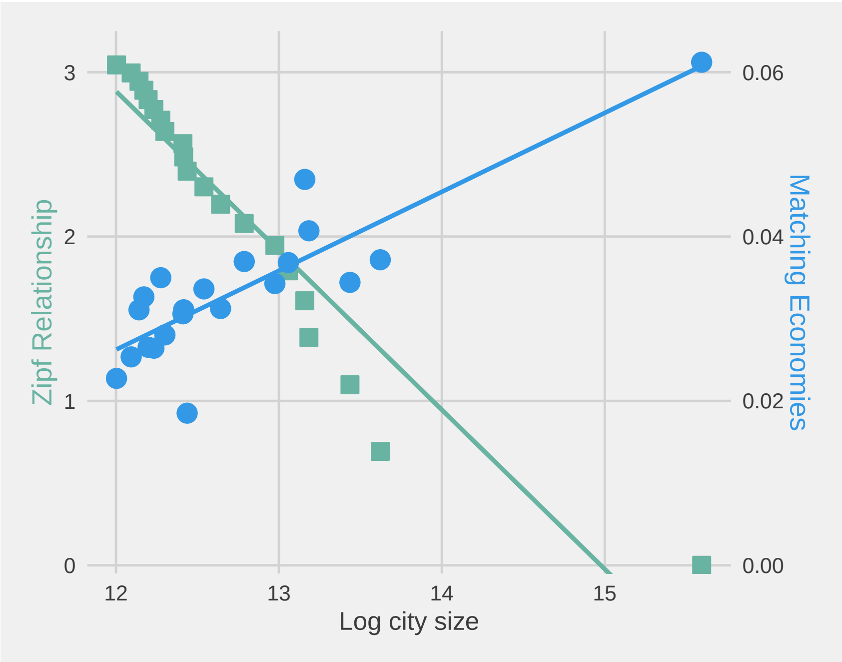
†Ecole Polytechnique and CREST; benoit.schmutz@polytechnique.edu
‡Duke University; modibo.sidibe@duke.edu

While space is not homogenous, it is futile to try to justify the marked unevenness of development solely on the basis of space being naturally heterogeneous (Duranton & Puga, 2004).

Introduction

The spatial allocation of production factors is remarkably skewed. The sorting of high ability workers, the selection of the most productive firms into the most productive cities, and the causal impact of city size on productivity are leading explanations of the positive relationship between city size and factor productivity. Yet, and despite the recent empirical interest in power laws, the reason why large cities exist in the first place is still controversial.

Figure 1: City size and matching economies in France



Notes: (i) Squares represent the log rank size distribution of the first 21 cities (metropolitan areas) in France based on the measure of local labor force aged between 15 and 64 and circles are point estimates of yearly local job arrival rates for employed workers residing in each of these first 21 cities; (ii) The solid lines represent linear regressions; the Zipf estimate is 97% ($R^2 = 91\%$) and the matching estimate is 0.96% ($R^2 = 74\%$).
Source: 2006 Census and 2002-2007 DADS panel.

In this paper, we argue that the limitations of the literature on city size distribution may be traced back to its paradoxical "aspatiality", as objects defined as "cities" are generally located over a continuum of locations. Assuming a discrete network of cities, we show that this unique source of heterogeneity can generate a Zipf distribution of cities in a model of frictional spatial equilibrium. In contrast to existing models, ours does not require any transfer of power laws from workers or firms to cities, series of long and persistent positive idiosyncratic shocks, nor aggregate production or matching functions with increasing returns (Champernowne, 1953; Simon, 1955; Duranton & Puga, 2004; Eeckhout, 2004; Behrens & Robert-Nicoud, 2015).

We embed an urban model with both production and consumption externalities in the spirit of Moretti (2011) into a search and matching model à-la Mortensen & Pissarides (1999). Combining insights from the trade literature (Melitz, 2003) with market-clearing from equilibrium search and matching models, we propose a setting where firm entry does not drive the profit of all vacancies to zero, allowing us to generate spatial heterogeneity in the distribution of firm productivities. Workers are continuously reallocating between cities through both off-the-job and on-the-job search. However, this reallocation is subject to spatial frictions that lower workers' ability to hear about or apply to remote job offers. After matching, workers and firms bargain over wages à-la Cahuc, Postel-Vinay & Robin (2006). The model delivers an equilibrium framework where heterogeneous workers and firms meet, bargain over wages, and sometimes separate, in local labor markets of endogenous size. Agglomeration economies emerge through the *matching* channel: as illustrated by Figure 1, large cities enable employed workers to keep searching for better jobs, so that observed matches may be closer to the Walrasian allocation.

The paper's main methodological contribution is to introduce urban channels into an equilibrium search and matching model.[1] The model operates through a surplus function, which summarizes the dynamic trade-off of heterogeneous workers and firms. This surplus function also includes feedback from the housing market since any spatial transition affects the local rents in both origin and destination cities. As such, it guides not only which matches are created but also a whole range of outcomes, such as the spatial distribution of population and human capital, spatial and skill wage premia, or low-productivity traps. For this reason, the model may prove useful to analyze the welfare consequences of very diverse place-based policies such as local minimum wages, optimal spatial taxation and industrial clusters (Cengiz, Dube, Lindner & Zipperer, 2019; Eeckhout & Guner, 2017; Kline & Moretti, 2014).

The key features of our model: discrete network of cities, heterogeneity in human capital and firm productivity, and on-the-job search, are not costless - analytical results on spatial equilibrium are out-of-reach. However, we can simulate the steady-state allocation of local outcomes under calibrated parameters. We use the model to derive simple statistics exploiting spatial heterogeneity in metrics such as the local level of skill, and the value of any given match to identify the presence of workers' sorting, firms' selection, and agglomeration economies in the data. This provides a transparent way of discussing the relative importance of sorting, selection, and agglomeration on city size distribution.

Our results suggest that the truncation of local productivity distribution tends to increase with city size. However, higher entry costs shield low-productivity firms against competition: when entry costs are large enough to reduce the equilibrium number of firms, all types of firms will exist in all cities. This mechanism may explain why firm selection, as measured by the level of left-truncation in local productivity distributions, may not be a large force behind agglomeration economies. Our results also show that a supermodular production function is required for high-skill workers to sort into larger cities. In this case, the level of assortativeness in the local labor market (as measured by the correlation between the worker's skill and the productivity of the firm) also increases with city

[1] The first papers seeking to incorporate matching frictions into the classical spatial equilibrium framework do not model the spatial foundations of local market tightness, do not model on-the-job search and search between local labor markets and do not allow for worker or firm heterogeneity (Kline & Moretti, 2013; Wrede, 2015).

size, but only when workers are able to search on-the-job. Finally, we show that matching-based agglomeration economies may arise in equilibrium, but only for matches involving productive firms, which workers have no incentive to quit.

We make use of our simulations to document the relationship between selection, sorting, agglomeration, and the city size distribution. We compare our benchmark simulation to cases where we shut down each channel separately. Absent positive assortative matching, cities all have roughly the same size. On the contrary, the absence of firm selection does not preclude the emergence of large cities, even if it seems to slightly limit it. In the absence of agglomeration economies, cities may still be quite heterogeneous in size, but the spatial equilibrium is centrifugal, and the size distribution is nowhere near exhibiting a linear log rank - log size relationship.

Finally, we describe the effect of spatial search frictions on the population's spatial distribution. Under moderate values of spatial frictions, a core-periphery equilibrium with a large central city and a positive correlation between city size and centrality can emerge. Under low-to-moderate spatial frictions, the distribution of city size may be quite close to the standard Zipf relationship. Conversely, a high (though not prohibitive) level of spatial frictions can generate regional clusters with a core-periphery structure. This last exercise shows that the model may be fruitfully used to rationalize observed differences in the shape of the urban network, from peripheral clusters in the US to core cities such as Paris or Madrid.

Relationship to the literature — This paper provides an additional step toward a comprehensive agglomeration theory, which could pave the way to more careful empirical work.[2] Existing empirical studies are set in partial equilibrium, either focusing on the worker's side to study sorting or on the firm's side to study selection. As these mechanisms interact with each other, a general approach is necessary. For example, wages, net of factor productivity, are often used to measure agglomeration economies, even though local wages depend on the relative scarcity of local production factors (Beaudry, Green & Sand, 2012).

Interestingly, several theories of agglomeration have been motivated by the desire to provide a micro-foundation to the existence of power laws for city size, which Gabaix (2009) coins as "one of the rare non-trivial quantitative laws in economics". Drawing from Champernowne's insights, Gabaix (1999) provides a mathematical micro-foundation to the observation of a power law, where it emerges as the steady-state of a proportional random growth process combined with a lower bound on city size. For the power law to be Zipf, one merely needs the lower bound to be small and the total population to be exogenously fixed. Similarly, we will use the steady-state and the fixed total population assumptions, and a lower bound on city size will result from the existence of construction costs. Rossi-Hansberg & Wright (2007), among others, have introduced more urban mechanisms by shifting the random component to the cross-section. Indeed, one possible way of generating a power law in city size is by transfer of another feature, such as TFP.[3] This is the case, for instance, of the model by Lee & Li (2013), where TFP is the product of several weakly correlated

[2] A unified state of agglomeration theory with heterogeneous agents absent search frictions can be found in Behrens & Robert-Nicoud (2015).

[3] In a search framework featuring endogenous firm size, Acemoglu & Hawkins (2014) also show that the firm size distribution may be Pareto if firm productivity is itself drawn from a Pareto.

random factors. However, the empirical counterpart to these factors is not easy to imagine.[4]

Given these earlier papers' focus on the interplay between space, city size, and agglomeration economies, the spatial allocation of heterogeneous factors is not addressed.[5] To the best of our knowledge, the most general theory with two-sided heterogeneity in an urban system to date is the model by Behrens, Duranton & Robert-Nicoud (2014). In a static framework, the authors can reproduce what they identify as the main empirical results regarding sorting, selection, agglomeration, agglomeration effects, higher returns to workers' skills in larger cities, and negligible remaining firm selection. Their model has many other appealing features: workers can choose to become either entrepreneurs or workers, providing a natural way to generate multi-dimensional heterogeneity. Agglomeration economies stem from a higher number of entrepreneurs, hence more diverse inputs in larger cities. Also, the number of cities emerges as an equilibrium outcome. However, we argue that additional insights may be gained from the addition of search frictions, for several reasons.

First, while sharing economies may be significant, recent research has emphasized matching economies' crucial role. For instance, Papageorgiou's (2018) quantitative model suggests that the wider span of occupations offered in larger cities, allowing for better sorting, may explain up to a third of the conditional urban wage gap. Positive assortative matching may have large macroeconomic consequences. In a recent AKM-type decomposition based on German data, Dauth, Findeisen, Moretti & Südekum (2018) estimate that the correlation between firm and worker fixed effects is six percentage points higher if one doubles city size. The authors argue that the increase in assortativeness in the German labor market between 1985 and 2014 is responsible for more than a 2% increase in total labor earnings over the period. By definition, these matching economies cannot be modeled within a competitive framework.

Another reason for the introduction of search frictions is that competitive settings predict perfect sorting by skill or degenerate local productivity distributions. In order to allow for firm selection, the choice made by Behrens et al. (2014) is to keep nondegenerate firm productivity distributions by imposing restrictive assumptions on workers' mobility: this is obviously true in the static version of the model -where workers are perfectly mobile when they make their location decision but perfectly immobile after being hit by a draw of luck but remains verified in a dynamic variant where workers can try their luck several times but at the cost of losing one period of urban income through a forced relocation in the hinterland. As explained by Eeckhout, Pinheiro & Schmidheiny (2014), a way to generate imperfect sorting would be to drop the discrete choice framework and allow housing consumption to increase with wages and hence with skills. An alternative way is to allow for equilibrium frictional migration flows between cities, as will be done in this paper.

A third reason for resorting to a frictional framework is the ability to accommodate career considerations: indeed, since migration is an investment, it requires not only a static trade-off between economic conditions but also a comparison between expected future economic conditions (Gallin, 2004). As argued by Kennan & Walker (2011), in the presence of mobility costs between cities, a dynamic framework is necessary. This is all the more important that workers' experience may accumulate from past stays in large cities. The crucial impact of learning-based agglomeration

[4]In a trade context, some papers have linked the distribution of local TFP with locational advantage (Behrens, Lamorgese, Ottaviano & Tabuchi, 2009).

[5]For instance, workers are homogeneous in Rossi-Hansberg & Wright (2007).

economies in workers' careers has recently been demonstrated by De La Roca & Puga (2017). While our framework does not account for the evolution of workers' skills throughout their life cycle, it still allows us to depict *spatial strategies*. The decision to accept a job in a given city depends not only on offered wages and local labor market primitives but also on current and future prospects in all other locations.

Finally, as previously explained, the existing literature on city size distribution usually assumes that the objects defined as "cities" are located over a continuum of locations.[6] While this modeling device may be appealing from a theoretical viewpoint, it complicates model-based empirical exercises.[7] Furthermore, the urban network of a country may be extremely persistent (Duranton, 2007), even when drastic long-run changes have made some of its nodes suboptimal (Michaels & Rauch, 2017). In this paper, we draw inspiration from the trade literature and take each city's position in the network as given. We show that this single element of heterogeneity may generate a large diversity of city profiles, even in the absence of any local idiosyncrasy affecting workers' utility or productivity.

The rest of the paper is organized as follows: in section 1, we define the urban setting and the working of the labor market; in section 2, we study the properties of the steady-state spatial equilibrium of the model. Our resolution strategy is detailed in section 3. Various simulation results are presented in section 4 and section 5 briefly concludes.

1 Framework

1.1 Cities and population

Space is divided into J non-overlapping local labor markets, or cities, indexed by $j \in \mathcal{J}$, where people both live and work. The location of each city within space is given by a fixed distance matrix $(d_{jl})_{(j,l)\in\mathcal{J}^2}$. The population consists of a continuum of workers of fixed measure normalized to 1. The population share of individuals living in city j, L_j is endogenous as well as the number of local firms N_j. Households consume 1 unit of housing. If ψ denotes relative land consumption by firms, city size s_j is given by $s_j = L_j + \psi N_j$. Rents are paid to absentee landlords.[8] We let the cost of supplying housing increase with city size, hence local rents R_j are competitively determined according to a market clearing equation $R_j = R(s_j)$, with $R'(\cdot) > 0$.[9]

Workers are heterogeneous in ability h, which is fixed over time and determined prior to labor market entry with uniform distribution $\ell(h) \hookrightarrow U[\underline{h}, \overline{h}]$.[10] The local skill distribution is endogenous

[6] Hsu (2012), who provides a micro-foundation of a power law based on central place theory when the distribution of scale economies is regularly varying, stands out as an exception.

[7] In particular, it seems to be key for maintaining analytical tractability. Behrens et al. (2014) discuss an extension of their model with a discrete set of cities and use simulations to show that their model's main properties hold in this alternative setting.

[8] We do not model real estate or firm ownership, which would mostly matter in a welfare analysis.

[9] This lack of segmentation is a first order approximation to the modeling of competition between the commercial and residential segments of the real estate market. Stronger limits to arbitrage in housing might yield higher volatility. However, the two sectors are indeed very correlated. For example, Gyourko (2009) shows that the simple correlation between appreciation rates on owner-occupied housing and commercial real estate is nearly 40% in the US over 1978-2008.

[10] Given our main objective here is to derive how production factors are allocated across cities, we do not need to assume a more complex underlying distribution of skills. Moreover, this ensures that the shape of equilibrium will not

and denoted $\ell_j(h)$. We follow the underlying model in Chapelle, Wasmer & Bono (2019) and feature an exogenous exit from urban space as follows: any urban worker may be hit with a Poisson shock Ξ and relocate out of the city. This acts as a death rate in perpetual youth models à-la Blanchard (1985). We normalize the value of exiting urban space/death to zero. Dead workers are replaced one-for-one by new agents of the same skill level, who start off as unemployed and are born proportionally to city size. Their local measure $\omega_j(h)$ then verifies:[11]

$$\omega_j(h) = \Xi \ell(h) L_j \tag{1}$$

There are no random shocks to utility and spatial equilibrium implies that migration between cities only occurs upon finding a job.

1.2 Labor Market

In this section, we describe the mechanism behind the spatial reallocation of production factors. We develop a continuous time model where differences between cities come from the constrained endogenous location decision of workers and firms. Matching between firms and workers is random. Workers and firms seek to maximize their lifetime income/profit. Both discount time at rate r. The key features of the model are as follows: i) workers are heterogeneous in ability or human capital h; ii) firms are heterogeneous in productivity p; iii) upon matching, they produce a generic good with a space-invariant production technology $\chi(h, p)$; iv) wage and employment dynamics are determined by city-specific bargaining and job-finding opportunities.

Workers — Workers may be either employed, or unemployed, and these states are denoted by $i = e, u$. Workers are risk neutral and endowed with a linear utility. Unemployed workers receive location-indifferent benefits $b(h)$, while employed workers enjoy a wage w. The stock of unemployed population in city j is denoted U_j, while productivity among the unemployed is determined by a distribution $u_j(h) \leq \ell_j(h)$.

Production — In each location, there is a continuum of latent firms. Firms are heterogenous in productivity p, with pdf $g(\cdot)$ and cdf $G(\cdot)$. Contrary to standard trade models, where firm production follows a Pareto distribution, we consider that p is uniformly distributed over $[\underline{p}, \overline{p}]$. Upon entry, the firm may match with a type-h worker. A type-p firm produces output defined by production technology $\chi(h, p)$, with positive first derivatives. Production and revenues are equivalent.

The equilibrium numbers of jobs and vacancies in city j are respectively given by N_j and V_j. Productivity among vacancies is distributed according to $v_j(p)$ and the joint distribution of matches is denoted $m_j(h, p)$. By definition, we have $L_j - U_j \equiv N_j - V_j$, where $V_j = \int v_j(p)dp$, $U_j = \int u_j(h)dh$ and $v_j(p) = n_j(p) - \int m_j(h, p)dh$, where $n_j(p)$ is the local distribution of firm productivity. The determination of these objects is deferred to the next section.[12]

be driven by a "transfer" mechanism such as the ones described in the Introduction.

[11]This effectively shuts down the impact of space on intergenerational linkages, despite empirical evidence to the contrary (Chetty, Hendren, Kline & Saez, 2014). In the model, the other polar case of perfect correlation would be given by: $\omega_j(h) = \Xi \ell_j(h)$.

[12]Hereafter, $\int$ denotes integration over the full span of h or p and $\int_{A(x)} f(x)dx$ denotes $\int 1_{x \in A(x)} f(x)dx$.

Search and Matching — Workers engage in both off-the-job and on-the-job search. The probability of receiving a job offer depends upon the spatial link between the locations of the worker and the firm. Similarly, the probability that a vacancy is filled depends on its location. We define the efficiency-adjusted search intensity in city j as:

$$E_j = \sum_{k \in \mathcal{J}} \varrho_{kj} \left[U_k + \zeta (L_k - U_k) \right] \tag{2}$$

where ζ is the relative search efficiency of employed workers and ϱ_{jl} measures the connectedness of jobseekers in city j to jobs in city l (described below).

The matching technology $\mathcal{M}(\cdot)$ is the same across cities. The number of matches in city j is given by $\mathcal{M}(E_j, V_j)$. Following Lise, Meghir & Robin (2016), we define $\theta_j = \mathcal{M}(E_j, V_j)/(E_j V_j)$ and $\theta_{lj} = \varrho_{lj}\theta_j$. The random aspect of job search is reflected in the fact that the instantaneous rate at which an unemployed (resp., employed) worker in city j meets a type-p vacancy in city l is $\theta_{jl} v_l(p)$ (resp., $\zeta\theta_{jl} v_l(p)$). Similarly, the instantaneous rate for any vacancy in city j to make a contact with a type-h unemployed worker (resp., a type-h worker employed in a type-p firm) in city l is $\theta_{lj} u_l(h)$ (resp., $\zeta\theta_{lj} m_l(h, p)$).

Matches are destroyed at exogenous Poisson rate δ, reflecting location indifferent shocks. These exogenous separations do not lead to the destruction of the job. They are best thought of as reflecting unobserved taste shocks from either the employer or the worker.

Wage determination — The value of a match between a type-p firm in city j and a type-h worker paid at wage w and the value of a vacancy for a type-p firm in city j are given respectively by $\Pi_j^f(h, w, p)$ and $\Pi_j^v(p)$. Similarly, the value functions for unemployed and employed workers are respectively denoted $\mathcal{V}_j^u(h)$ and $\mathcal{V}_j^e(h, w, p)$. We assume that utility is transferable between workers and firms. If we denote by $\widetilde{\mathcal{S}}_{jl}(h, p)$ the surplus associated with a match between a type-h unemployed worker in city j and a type-p vacancy in city l and by $\mathbb{P}_l(h, p) \equiv \Pi_l^f(h, \cdot, p) + \mathcal{V}_l^e(h, \cdot, p)$ the product of the match, which therefore does not depend on the level of wage, this gives:

$$\widetilde{\mathcal{S}}_{jl}(h, p) = \mathbb{P}_l(h, p) - \mathcal{V}_j^u(h) - \Pi_l^v(p) \tag{3}$$

Surplus is shared according to the bargaining framework proposed in Lise et al. (2016), adapted to a spatially segmented labor market. Jobseekers are characterized by a parameter of bargaining power β independent of labor market status or location. When they are unemployed and are contacted by a firm, workers either get hired or stay unemployed, through a *reservation wage strategy*. When they are employed and are contacted by another firm, workers may either dismiss this new opportunity, stay in the same firm and negotiate a wage increase (the *wage-increase strategy*), or move to the challenging firm (the *firm-switching strategy*). Labor market transitions may take place within or between cities. Different cities feature different option values, driven by the local structural parameters and firm productivity distributions. Contrary to what happens in the single market case, both the wage-increase and the firm-switching strategies may be triggered either by a less productive or by a more productive firm than the incumbent firm. We describe below the reservation-wage and firm-switching strategies.[13]

Reservation-wage strategy – Let $\psi_{jl}(h,p)$ be the wage offered by a type-p firm in l to a type-h worker unemployed in j. A match is feasible only when the associated surplus is positive. Given the reservation sets $\mathcal{H}^0_{jl}(p) = \left\{x : \widetilde{\mathcal{S}}_{jl}(x,p) \geq 0\right\}$ and $\mathcal{P}^0_{jl}(h) = \left\{y : \widetilde{\mathcal{S}}_{jl}(h,y) \geq 0\right\}$, the rent sharing process between a type-h unemployed worker living in city j and a type-p firm located in city l verifies:

$$\forall x \in \mathcal{H}^0_{jl}(p), \quad \Pi^f_l(x, \psi_{jl}(x,p), p) - \Pi^v_l(p) = (1-\beta)\widetilde{\mathcal{S}}_{jl}(x,p) \tag{4}$$

$$\forall y \in \mathcal{P}^0_{jl}(h), \quad \mathcal{V}^e_l(h, \psi_{jl}(h,y), y) - \mathcal{V}^u_j(h) = \beta\widetilde{\mathcal{S}}_{jl}(h,y) \tag{5}$$

Firm-switching strategy – Let $\psi^c_{jl}(h,p,p')$ be the wage offered by a type-p' firm in l to a type-h worker initially employed by a type-p firm in j, where the challenging firm generates a higher surplus. The corresponding strategy set $\mathcal{P}^c_{jl}(h,p) = \left\{y : \widetilde{\mathcal{S}}_{jj}(h,p) \leq \widetilde{\mathcal{S}}_{jl}(h,y)\right\}$ does not depend on the current level of wage w because workers, during the bargaining process, use the total surplus value of their current match as threat point. The firm-side equivalent of this set is $\mathcal{P}^f_{lj}(h,p) = \left\{y : \widetilde{\mathcal{S}}_{ll}(h,y) \leq \widetilde{\mathcal{S}}_{lj}(h,p)\right\}$, consists of all firms in city l employing a type-h worker who can be poached by a type-p firm in city j. The firm-switching strategy of a type-h worker employed in a type-p firm in city j contacted by a type-p' firm in city l and the worker-poaching strategy of a type-p firm in city j poaching a type-h worker employed in a type-p' firm in city l are determined by a sequential auction process that verifies:

$$\forall p' \in \mathcal{P}^c_{jl}(h,p), \quad \mathcal{V}^e_l(h, \psi^c_{jl}(h,p,p'), p') - \mathcal{V}^u_j(h) = \widetilde{\mathcal{S}}_{jj}(h,p) + \beta[\widetilde{\mathcal{S}}_{jl}(h,p') - \widetilde{\mathcal{S}}_{jj}(h,p)] \tag{6}$$

$$\forall p' \in \mathcal{P}^f_{lj}(h,p), \quad \Pi^f_j(h, \psi^c_{lj}(h,p',p), p) - \Pi^v_j(p) = (1-\beta)[\widetilde{\mathcal{S}}_{lj}(h,p) - \widetilde{\mathcal{S}}_{ll}(h,p')] \tag{7}$$

Even though workers only switch firms when this generates a higher match surplus, note that this strategy may sometimes translate into a wage cut, if the challenging firm is located in a very dynamic city, featuring many future opportunities.

1.3 Spatial constraints

There are three possible ways of introducing spatial constraints in this model: through trade costs, mobility costs and spatial search efficiency. Trade costs impact firms' revenues, mobility costs impact the value of the match surplus and spatial search efficiency impacts the meeting rate between jobseekers and vacancies.[14]

Trade costs — In a model with trade costs, production and revenues are no longer equivalent. For example, local revenues $\chi_j(h,p)$ could be given by equation 8:

$$\chi_j(h,p) = T_j\chi(h,p), \text{ with } T_j \equiv \frac{\sum_{k\in\mathcal{J}} \tau_{jk}L_k}{\frac{1}{J}\sum_{(k,l)\in\mathcal{J}^2} \tau_{lk}L_k} \tag{8}$$

[13]By definition, the wage-increase strategy does not impact the local distribution of skills and productivities. Since the primary focus of this model is not to model wage dynamics, we defer the presentation of this strategy to Appendix A.

[14]For a discussion on the equivalence between trade costs and commuting costs, see Allen & Arkolakis (2019).

where $\tau_{jl} \equiv \tau(d_{jl})$, with $\tau(0) = 1$, $\tau'(\cdot) \leq 0$ and $\tau(\infty) \geq 0$. In this setting, τ_{jl} measures the net profit between the mark-up associated with selling in a different city and potential iceberg cost and the TFP factor T_j is a measure of city relative centrality. If $\tau(d_{jl}) = \tau(0) = 1$, all cities have the same centrality and $T_j = 1$. Conversely, if $\forall l \neq j, \tau(d_{jl}) = 0$, then $T_j = \frac{L_j}{L}$ and the home-market effect increases productivity in larger cities.[15] However, this way of introducing spatial constraints is arguably more relevant if one seeks to explain international trade patterns than domestic trade so we focus on the case where $\tau_{jl} = 1$.

Mobility costs — Mobility cost can be incorporated into the model as follows: assume that moving a worker between two cities costs $\phi_{jl} \equiv \phi(d_{jl}) > 0$ and that this cost is sunk; then, a spatial match induces lower surplus, and hence lower wages and profits. Let $\mathcal{S}_{jl}(h,p) = \widetilde{\mathcal{S}}_{jl}(h,p) - \phi_{jl}$. This adjusted surplus is the one that will be taken into account in firms' and workers' decisions depicted in strategy sets $\mathcal{H}^0$, $\mathcal{P}^0$, $\mathcal{P}^c$ and $\mathcal{P}^f$. All else equal, a worker, if given a choice, would prefer working for a local firm, and a firm would prefer hiring a local worker. While such a decision never has to be made in a continuous time framework, it may be the case that some jobs are not available to remote workers, while they are available to a local worker of the same skill level.[16]

Spatial frictions — Finally, space may directly impact the effectiveness of the search and matching process. The spatial instability of the matching function has long been documented (Burda & Profit, 1996) and spatial search costs are distinct from mobility costs (Kennan & Walker, 2011; Schmutz & Sidibé, 2019). In the model, spatial search frictions operate through a spatial decay function $\varrho(\cdot)$ such that $\varrho_{jl} \equiv \varrho(d_{jl})$, with $\varrho(0) = 1$, $\varrho'(\cdot) \leq 0$ and $\lim_{+\infty} \varrho(\cdot) \geq 0$. This function measures a compound of informational frictions (workers are less likely to hear about remote vacancies) and spatial impediments to the application process (workers who apply remotely are less likely to receive call-backs).

1.4 Recursive formulation

Equations 9 and 10 respectively describe the program of type-h unemployed workers living in city j and of type-p vacancies located in city j:

$$(r + \Xi)\mathcal{V}_j^u(h) = b(h) - R_j + \beta \sum_{k \in \mathcal{J}} \theta_{jk} \int_{\mathcal{P}_{jk}^0(h)} \mathcal{S}_{jk}(h,y) v_k(y) dy \tag{9}$$

$$\begin{aligned} r\Pi_j^v(p) &= -\psi R_j + (1-\beta) \sum_{k \in \mathcal{J}} \theta_{kj} \int_{\mathcal{H}_{kj}^0(p)} \mathcal{S}_{kj}(x,p) u_k(x) dx \\ &+ (1-\beta) \sum_{k \in \mathcal{J}} \zeta \theta_{kj} \int_{\mathcal{H}_{kj}^0(p)} \int_{\mathcal{P}_{kj}^f(x,p)} \left[\mathcal{S}_{kj}(x,p) - \mathcal{S}_{kk}(x,y)\right] m_k(x,y) dy dx \end{aligned} \tag{10}$$

There are two departures from the traditional aspatial search framework: first, both equations feature local rent level in the instant utility. Unemployed workers have a lower instant disposable wage in larger cities, as verified in the urban literature. However, they are still willing to live there

[15] Such a pricing mechanism may be micro-founded under a Helpman-Krugman model as in Behrens et al. (2009).
[16] Note that the bargaining process about the sharing of the mobility cost is not theoretically well-grounded, although this issue is somewhat mitigated if one assumes $\beta = 0.5$.

because their financial expectations are high enough.[17] The cost of opening a vacancy is also larger in big cities, because of commercial real estate. However, some firms are able to make up for this loss thanks to shorter expected vacancy duration or high-quality prospective hires.

The second departure resides in the fact that option values are driven by the position of the city in the urban network. For example, if spatial frictions permit, more central cities may feature more matches and tend to be larger in equilibrium. However, the relationship between centrality and size is not trivial, because large cities may cannibalize their smaller neighbors. For example, given the trade-off between real estate costs and matching externalities, one might think that a good location for some unemployed workers may be a small city, near a large one, from which many job offers emanate. However, this will depend on the relative scarcity of their skill type in and around the large city. If many residents of the large city have the same level of skill, the competition for jobs will be fierce, especially since workers from the small city are both less efficient at searching and make up for less valuable hires, because of mobility costs. The same type of reasoning applies to vacancy posting.

The value functions of employed workers $\mathcal{V}^e_j(h,w,p)$ and filled vacancies $\Pi^f_j(h,w,p)$ are described by equations 19 and 20 in Appendix A. Combining those two equations, equation 11 describes a recursive formulation for match product:

$$\begin{aligned} r\mathbb{P}_j(h,p) &= \chi(h,p) - (1+\psi)R_j - \delta_j(h,p)\mathcal{S}_{jj}(h,p) - \Xi\mathcal{V}^u_j(h) \\ &+ \sum_{k\in\mathcal{J}} \zeta\theta_{jk} \int_{\mathcal{P}^c_{jk}(h,p)} \left[(1-\beta)\mathcal{S}_{jj}(h,p) + \beta\mathcal{S}_{jk}(h,y)\right] v_k(y)dy \end{aligned} \tag{11}$$

where $\delta_j(h,p) = \Xi + \delta + \sum_{k\in\mathcal{J}} \zeta\theta_{jk} \int_{\mathcal{P}^c_{jk}(h,p)} v_k(y)dy$ is the match destruction rate, which features all sources of job destruction (idiosyncratic match shock, death, and job to job mobility).

2 Equilibrium

2.1 Steady state

In the steady-state equlibrium, the distribution of type-h unemployed workers and the distribution of type-(h,p) matches are stationary in each local labor market $j \in \mathcal{J}$. These conditions are respectively given by equations 12 and 13.

Unemployed workers may either find a job or die. Inflows of unemployed workers are made out of locally employed workers who experience the exogenous destruction of their previous job (no voluntary unemployment and no mobility of unemployed workers absent of job-related reasons) and of young workers:

$$u_j(h) = \frac{\delta \int m_j(h,y)dy + \omega_j(h)}{\Xi + \sum_{k\in\mathcal{J}} \theta_{jk} \int_{\mathcal{P}^0_{jk}(h)} v_k(y)dy} \tag{12}$$

Outflows of type-(h,p) matches are either caused by exogenous destruction, worker's death or worker's departure, while inflows are matches involving type-p vacancies and type-h jobseekers,

[17] The lack of participation decision means that $\min_{j,h} \mathcal{V}^u_j(h) \geq 0$, thanks to high enough unemployment benefits $b(h)$.

unemployed or not, who came to the city to work in a firm at least as productive as p but had to settle for a type-p firm.

$$m_j(h,p) = \frac{\sum_{k\in\mathcal{J}}\left(\theta_{kj}1_{h\in\mathcal{H}^0_{kj}(p)}u_k(h) + \zeta\theta_{kj}\int_{\mathcal{P}^f_{kj}(h,p)} m_k(h,y)dy\right)v_j(p)}{\delta_j(h,p)} \tag{13}$$

Each city is then fully characterized by the combination of equations 12 and 13.

Brain drain and brain gain — An interesting feature of the model is that contrary to the single-market case, population $\ell_j(h)$ is endogenous and satisfies the demographic constraint:

$$\omega_j(h) - \Xi\ell_j(h) = \mathcal{O}_j(h) - \mathcal{I}_j(h) \tag{14}$$

where $\omega_j(h) - \Xi\ell_j(h)$ features natural growth of type-h workers in city j and $\mathcal{O}_j(h)$ (resp., $\mathcal{I}_j(h)$) is the number of type-h workers leaving (resp, migrating into) city j.[18] The role of each city in the urban system may vary. In particular, some cities may be receiving more "foreign" workers of a given skill level. Cities are more skill-intensive when they are more likely to attract or retain more high-skilled workers.

2.2 Firm entry

The entry decision of firms leads to profit exhaustion for the least productive entrant. This zero-profit cutoff $\underline{p}_j$ is defined by $\Pi^v_j(\underline{p}_j) = 0$. It will be unique if the option value of a vacancy increases with its productivity and there will be (spatial) selection if $\partial\underline{p}_j/\partial L_j > 0$.

Firm entry does not drive the profit of all vacancies to zero. To achieve this feature, which is necessary to introduce firm selection as a local minimum productivity threshold below which firms are unable to operate, we consider a two-step entry process in the spirit of Melitz (2003). Before entry, firms are uncertain about their level of productivity. Entry requires an entry cost c and free entry in each city implies:

$$c = \int 1_{p>\underline{p}_j}\Pi^v_j(p)g(p)dp \tag{15}$$

This process determines the local distribution of firms, given by $n_j(p) = N_j g(p)/(1 - G(\underline{p}_j))$ if $p > \underline{p}_j$ and $n_j(p) = 0$ otherwise. Combining free entry with equation 10 allows us to recover an expression for θ_j:

$$\theta_j = \frac{rc + \psi R_j\left(1 - G(\underline{p}_j)\right)}{\int_{\underline{p}_j}^{\overline{p}} A_j(p)g(p)dp} \tag{16}$$

with $A_j(p) = (1-\beta)\sum_{k\in\mathcal{J}}\varrho_{kj}\int_{\mathcal{H}^0_{kj}(p)}\{\mathcal{S}_{kj}(x,p)u_k(x) + \zeta\int_{\mathcal{P}^f_{kj}(x,p)}[\mathcal{S}_{kj}(x,p) - \mathcal{S}_{kk}(x,y)]m_k(x,y)dy\}dx$. Note that contrary to the skill distribution that must verify $\ell(h) = \sum_{k\in\mathcal{J}}\ell_k(h)$, the aggregate produc-

[18] With $\mathcal{J}_j = \mathcal{J} - \{j\}$, we have: $\mathcal{O}_j(h) = u_j(h)\sum_{k\in\mathcal{J}_j}\theta_{jk}\int_{\mathcal{P}^0_{jk}(h)}v_k(y)dy + \int m_j(h,p)\{\sum_{k\in\mathcal{J}_j}\zeta\theta_{jk}\int_{\mathcal{P}^c_{jk}(h,p)}v_k(y)dy\}dp$ and: $\mathcal{I}_j(h) = \sum_{k\in\mathcal{J}_j}\{u_k(h)\theta_{kj}\int_{\mathcal{P}^0_{kj}(h)}v_j(y)dy + \int m_k(h,p)[\zeta\theta_{kj}\int_{\mathcal{P}^c_{kj}(h,p)}v_j(y)dy]dp\}$.

tivity distribution $n(p) = \sum_{k \in \mathcal{J}} n_k(p)/N_k$ may be different from the latent productivity distribution $g(p)$.

2.3 Definition of the equilibrium

Given a system $\mathcal{J}$ of cities j defined by relative location d_{jl}, a fixed population of workers with ability h, a continuum of latent firms with productivity p, production, matching and land supply technologies, an exogenous job destruction rate δ, a worker death rate Ξ, a frictional spatial equilibrium is defined by $\{\ell_j(h), u_j(h), v_j(p), m_j(h,p)\}_{j \in \mathcal{J}}$ and a match product sharing rule given by $\{S_{jl}(h,p)\}_{(j,l) \in \mathcal{J}^2}$ such that:

1. There is a strictly positive measure of workers of any skill level in each city.
2. All workers and firms maximize expected payoffs, taking the strategies of all other agents as given.
3. Firms offer a set of contracts contingent on location and skills, which defines the local skill distribution in the population $\ell_j(h)$.
4. Given local population, firms make their entry decision which defines the local distribution $v_j(p)$ of the productivity of vacancies. This distribution verifies the free-entry condition defined in equation 16.
5. Firms and workers meet, bargain over wages according to the process defined by equations 4 to 7 and become matched according to the joint distribution $m_j(h,p)$. This distribution verifies the steady state constraint defined in equation 13.
6. Unmatched workers are defined by the local skill distribution $u_j(h)$. This distribution verifies the steady state constraint defined in equation 12.

This equilibrium describes the spatial distribution of skill and productivity as well as labor flows between firms and cities. It is compatible with various wage dynamics within a given (h,p) match – see Appendix A for details. While these dynamics could matter to achieve better identification of the model, they are not crucial to city size distribution.

Existence and uniqueness — Proving that the steady-state exists in an equilibrium search and matching model is a difficult undertaking. For a class of non-transferable utility models, Burdett & Coles (1997) exploit the known structure of the matching to prove existence. Shimer & Smith (2000) prove existence when the production function is either super- or submodular. The idea is to map matching sets into value functions, which are then mapped into the unmatched densities. Finally, the composite of value function and matching sets map into new values. While our algorithm to construct the steady-state builds on this intuition, proving the existence of a steady-state equilibrium is out of reach because of on-the-job search and the induced competition between employed and unemployed workers in the matching technology. However, many papers have extended this argument in the setting with on the job search (Lise et al., 2016; Hagedorn, Law & Manovskii, 2017; Bagger & Lentz, 2018). For exogenous values of the parameters, there are many instances where an equilibrium with positive population measures does not exist. For example, in a city where the level of rent is very high while matching technology yields very few matches, no firms would operate and all workers would be permanently unemployed, with a negative instant utility. As a consequence, we restrict ourselves to the subset of parameters where an equilibrium exists.

Proving uniqueness is even harder and given the spatial nature of the search problem, multiple equilibria may be especially acute. However, many ingredients of the model alleviate concerns regarding multiple equilibria. First, search is random and, as detailed in Section 3, the matching function is CRS. Second, the rent function, which acts a shifter in the value of residing/operating in a location, clearly distinguishes between the role of firms and workers. Third, the value of a vacancy is monotonic in firm productivity, guaranteeing the uniqueness of the firm productivity cutoff. Fourth, utilities are linear and transferable between workers and firms, and there is no heterogeneity in preferences. As a consequence, the trade-off between the consumption value of a location (R_j) and local labor market attractiveness (θ_j) creates a set of indifference conditions that rule out off-the job migrations and strategic unemployment migrations.[19] Finally, we note in the simulations that if an equilibrium exists, it is unique.

3 Quantitative Analysis

We calibrate the model in steady state on a network of 21 cities with a distance matrix based on a uniform grid with small normal deviations (see Figure 7).[20] In principle, we can simulate the equilibrium on any network.

3.1 Functional forms

We first detail our choice of functional forms for production, housing supply and matching technologies as well as the impact of distance on spatial constraints. Those are summarized in Table 1.

Table 1: Functional forms

Model	Parameterization
$\chi(h,p)$	$[h^{(\sigma-1)/\sigma} + p^{(\sigma-1)/\sigma}]^{\sigma/(\sigma-1)}$
R_j	$R_0 + R_1(L_j + \psi N_j)^{R_2}$
θ_j	$m_0 E_j^{m_1-1} V_j^{-m_1}$
ϱ_{jl}	$\exp(-\varrho d_{jl})$
ϕ_{jl}	$1_{j\neq l} \times \phi$

First, given our focus is on the complementarity between skill and productivity, we assume a standard CES production function with constant returns to scale. Second, as justified and used by Behrens et al. (2014), housing supply is a concave technology, where $R_2 < 1$ is the population (of both workers and firms) elasticity to urban cost, which reflects both supply constraints and

[19]This means that historical episodes like the Klondike Gold rush or Alaska Oil Boomers are out of equilibrium. As shown in Figure 2 in Section 3.3, the value of unemployment is constant across space.

[20]The urban network is trivial if $J = 1$ and symmetric if $J = 2$. The number $J = 21$ is arbitrary but allows for a nice circular representation. Small normal deviations avoid having generically identical locations for several cities.

commuting-related congestion and is assumed invariant across space.[21] Yet, in contrast with the latter, we also feature a fixed cost $R_0 > 0$, meant to capture constraints on the construction sector. The third question has to do with the functional form of the matching function. According to most existing studies, the aggregate matching function features CRS and is generally compatible with Cobb-Douglas properties (Petrongolo & Pissarides, 2001).[22] Therefore, we keep a Cobb-Douglas matching function.

As for spatial constraints, these are summarized by parameters ϕ_{jl} and ϱ_{jl}. We assume a fixed mobility cost ϕ. Following Fujita & Thisse's (2002) use of an exponential to model knowledge spillovers, spatial frictions in the matching process are captured by a negative exponential of distance. The resulting heterogeneity in the level of connectedness by location is in line with the variability in spatial friction estimates of Schmutz & Sidibé (2019). Under this parametrization, the set of parameters Θ may be decomposed as follows: $\{r, \Xi\}$ relate to demography and time; $\{\psi, R_0, R_1, R_2\}$ characterize the real estate market; $\{b, \beta, c, \delta, \zeta, \sigma, m_0, m_1\}$ are structural determinants of the labor market; and $\{\varrho, \phi\}$, measure spatial constraints.

3.2 Calibration

The parameter values used in the simulations are given in Table 2. Panel A describes the parameters that generate sorting, selection and agglomeration and parameters driving spatial frictions. These parameters will vary in section 4. Panel B describes a set of fixed parameters.

A unit time interval in the model is set to a year. The discount rate is set to $r = 0.01$, which roughly corresponds to a 99% discount rate at annual frequency. We set the bargaining parameter $\beta = 0.5$, which dates back from the Hosios condition, and implies equal surplus sharing between the worker and the firm. The consumption value of unemployment $b(h)$ is given by $\chi(h, b)$, with $b = \underline{p}/2$. We set the job separation rate to $\delta = 0.03$, the death rate $\Xi = 0.02$, which together with a matching elasticity of $m_1 = 0.5$ and a scale of the matching function $m_0 = 0.5$ yield an aggregate unemployment rate of 4% for our benchmark value of spatial constraints ($\varrho = 0.2$).

R_0 is set to a low strictly positive value.[23] R_1 is calibrated to 0.5, which yields a rent level equal to between 35% and 60% of the lowest unemployment income.[24] R_2 is set to 0.3 to reflect recent empirical evidence on the elasticity of urban costs to city size (Combes, Duranton & Gobillon, 2019). Finally, we set ψ to 0.2 to match the ratio of commercial square footage to total real estate square footage in the US.[25]

On the remaining 4 parameters - ζ, σ, ϱ and c, we illustrate the properties of the model for a grid of values. We simulate the model for values of ϱ between 0 and 20 and for values of σ between -1 and 1, allowing for quasi-linearity in an extra simulation where $\sigma = 100$. Estimates and calibrated

[21] As shown by Saiz (2010), housing supply is less elastic in cities with higher geographic amenities, both directly, via reductions in the amount of land availability (the obvious example is coastal cities), and indirectly, via higher incentives for anti-growth regulations. We abstract from these important features.

[22] In a very disaggregated approach (but only based on job search by unemployed workers), Manning & Petrongolo (2017) also fail to reject the CRS assumption.

[23] This feature is meant to represent technological constraints in the construction sector. It could act as a lower bound on city size, in the spirit of Gabaix (1999). In practice, its role is however limited.

[24] Housing accounts for between 15% and 30% of household income in OECD countries.

[25] Total home square footage in the US is around 260 billion square feet, to be compared to 70 billion for commercial real estate.

Table 2: Parameter values

A - Key parameters

Mechanism	Parameter	Description	Benchmark	Range
Spatial Constraints	ϱ	Spatial search efficiency	0.2	$[0, 20]$
Sorting	σ	Substitution	0.2	$[-1, 1] \cup \{100\}$
Selection	c	Cost of opening a vacancy	100	$[0, 200]$
Agglomeration	ζ	Efficiency of on-the-job search	0.2	$[0, 0.4]$

B - Fixed parameters

Type	Parameter	Description	Benchmark
Elasticities	m_1	Matching	0.5
	R_2	Housing supply	0.3
Rates	r	Discount	0.01
	δ	Job destruction	0.03
	Ξ	Death	0.02
	m_0	Matching: scale	0.5
Sharing	β	Bargaining power	0.5
	ψ	Cost of commercial	0.2
Costs	b	Productivity of unemployment	1
	R_0	Fixed housing cost	0.01
	R_1	Housing cost: scale	0.5
	ϕ	Mobility cost	0.01
Grid	$[\underline{h}, \overline{h}]$	Skills	$[3, 20]$
	$[\underline{p}, \overline{p}]$	Productivity	$[2, 20]$

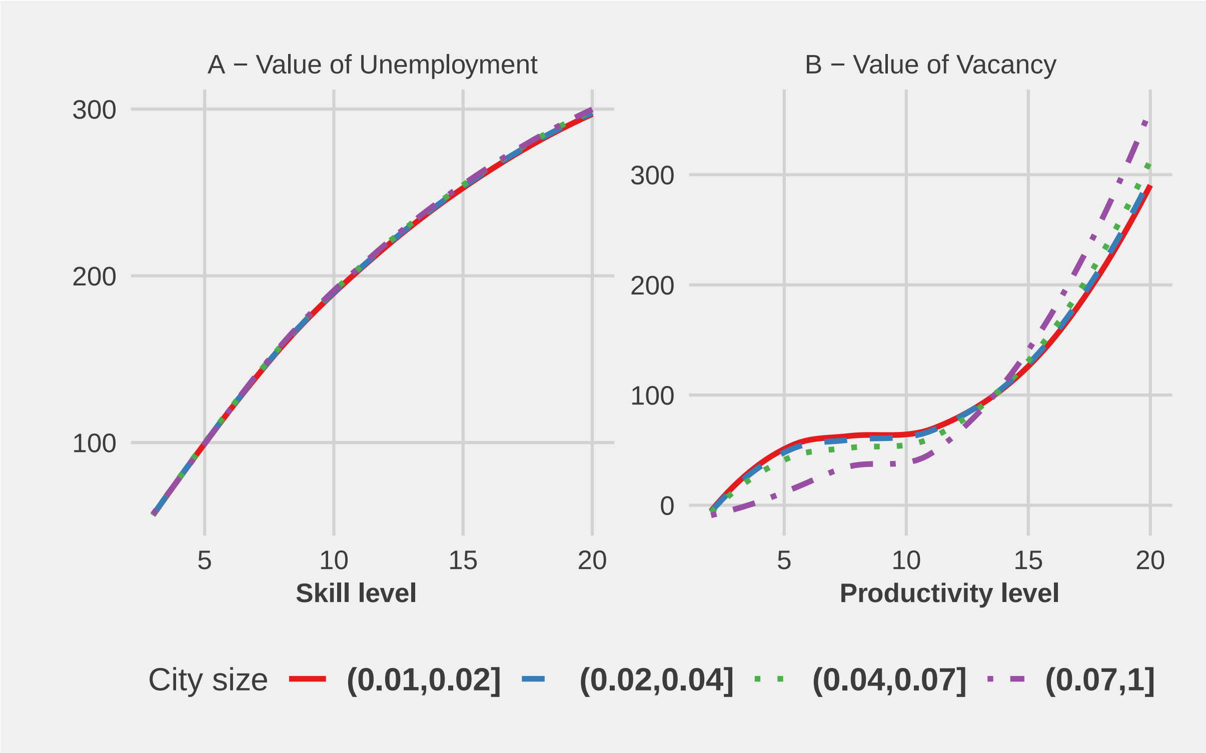
values of the relative search intensity of employed workers ζ range between 0.1 and 0.4 and we allow this parameter to vary between 0 and 0.4. Finally, we adopt a much wider range for c, which turns out to be crucial in generating reasonable values of the vacancy rate.

3.3 Resolution

For a given value of Θ, the model can be solved using a recursive algorithm - see Appendix B for details.[26] As argued in Section 2.3, spatial equilibrium implies equal unemployment values across space. Graph A in Figure 2 confirms that the equilibrium values of $\mathcal{V}^u_j(h)$ for all cities $j \in \mathcal{J}$ are virtually indistinguishable from each other, even though the algorithm constructs each one separately. On the contrary, Graph B displays substantial spatial heterogeneity in the value of a vacancy $\Pi^v_j(p)$: low-productivity vacancies have higher values in smaller cities thanks to lower operating costs and the opposite stands true for high-productivity vacancies, which are able to reap benefits from better labor market conditions in larger cities.

[26]Simulation codes are written in R and make heavy use of the Rcpp package. They are available upon request.

Figure 2: Spatial heterogeneity: the difference between unemployment and vacancy



Notes: Simulation results from the benchmark case; Graph A: Value of unemployment $\mathcal{V}_j^u(h)$ for all cities as a function of individuals' skill level; Variations between cities are very small and hardly appear on the graph; Graph B: Value of a vacancy $\Pi_j^v(p)$ for all cities (grouped by size) as a function of firms' productivity level.

4 Results

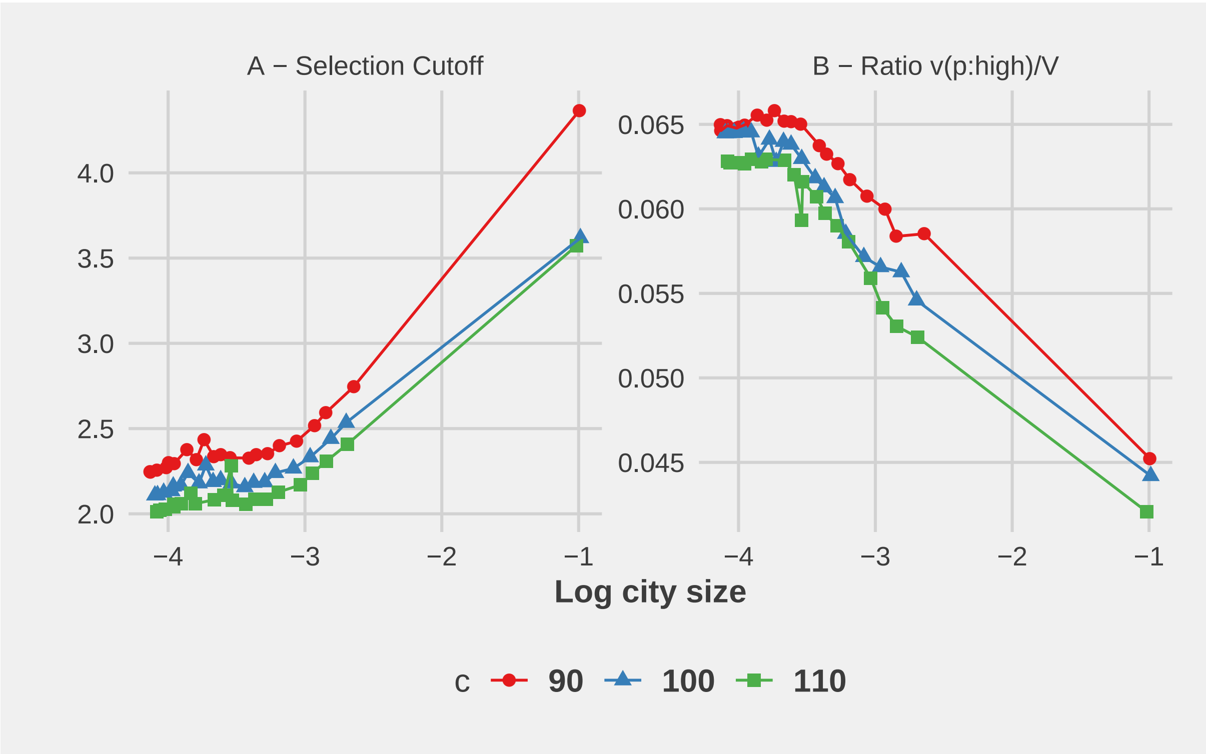
4.1 The mechanisms of selection, sorting and agglomeration

We first illustrate the workings of the model regarding firm selection, worker sorting, and agglomeration economies. We provide sufficient statistics to characterize each mechanism, we document the impact of technological parameters on the working of the model and we describe how this impact may or may not change with city size. To this end, we look at marginal changes in the parameters of interest to compare equilibria that are similar in other aspects.

Selection — The mechanism for selection is based on a trade-off between city-specific expected profit on the one hand, and productivity-invariant operating costs (entry and real estate) on the other hand. Given that the expected profit associated with posting a vacancy increases with productivity, the existence of operating cost has the potential to generate firm selection. However, firm selection is not uniform across cities. In particular, the impact of firm productivity on the profit that they can expect from the opening of a vacancy will be mostly beneficial to the most productive firms in the larger cities, because they are the ones which can actually tackle the competition in the hiring process, especially when it comes to the poaching of employed workers.

Our measure of firm selection, $\underline{p}_j$, which is defined as the lowest productivity level compatible with operating in city j, is represented by Graph A in Figure 3. The cost of opening a vacancy c has a direct effect on the intensity of firm selection. We illustrate the effect of marginal changes of c around the benchmark value of 100. Our results suggest an increasing relationship between the selection cutoff $\underline{p}_j$ and city size. Although smaller cities may exhibit some degree of selection, there is no clear relationship, which implies that smaller cities offer matching opportunities that are high enough to offset the higher rents.

Figure 3: Selection and city size



Notes: productivity cutoff $\underline{p}_j$ (Graph A) and share of high-productivity vacancies $v_j(\overline{p})/V_j$ (Graph B) as a function of city size ($\log L_j$) for three values of vacancy creation cost c; all other parameters are set to benchmark values.

On the contrary, selection is higher in the largest cities: firms with low productivity cannot afford higher rents. Note that in these calibrations, the level of firm selection remains relatively low, even in the largest city. This observation echoes Behrens et al.'s (2014) results of a small role for firm selection in spatial equilibrium. Interestingly, we also find that increasing the entry cost *decreases* selection. While this relationship may seem counter-intuitive at first, it highlights a limitation of selection measures based on productivity cutoffs. The vacancy cost affects the entry decision both on the intensive and extensive margins. As the entry cost increases, the number of firms decreases in the economy. With fewer firms in the economy, hiring competition attenuates (and the commercial rent level is lower), allowing low-productivity firms to survive, even in large cities.

The spatial dimension of hiring competition is illustrated in Graph B in figure 3, which shows the probability of randomly meeting a vacant firm with the highest productivity level: in larger cities, more productive firms are less likely to stay unmatched. However, Graph B also shows that higher entry costs benefit more to high-productivity firms, which are exposed to lower congestion.

Altogether, these mechanisms illustrate the difficulty of identifying selection using observational data and may explain why studies find little role for firm selection in the productivity advantages of large cities (Combes, Duranton, Gobillon, Puga & Roux, 2012).

Sorting — We now turn to our strategy to measure sorting. Contrary to firms, which do not experience demographic shocks and whose population emerges as an equilibrium outcome, all worker types may be observed in all cities. However, the distribution of skills, and in particular $h_j = \int h(\ell_j(h)/L_j)dh$ the local average skill level may vary along the city size distribution: for example, if the most productive firms, which are able to attract the most productive workers, are disproportionately located in larger cities.

Empirically, it is difficult to rule out the possibility that skill may be positively correlated with a stronger preference for large cities, which is not modeled here. In the model, sorting is not primarily a spatial phenomenon, but a technological one, which depends on the level of factor complementarity in the production function. Workers' spatial preferences are uncorrelated with their skill level and all workers face the same local rent irrespective of their skill. Therefore, if $\partial h_j/\partial L_j > 0$, this is indicative of skill-based spatial sorting. Following the literature that builds on Abowd, Kramarz & Margolis (1999), we measure the degree of assortativeness in the data using the correlation between workers' skill and firms' productivity, and we complement this measure with the simple average human capital.

The parameter that guides the degree of assortativeness in the economy is σ. In our benchmark, σ is set to 0.2. Figure 4 represents the correlations for each city in the spirit of Dauth et al. (2018) and the average level of human capital, for three values of σ implying positive assortative matching (PAM) between h and p and one value implying negative assortative matching (NAM).

It illustrates a clear contrast between NAM and PAM. Under NAM, all cities have roughly the same size and, if anything, both measures of sorting are flat with city size. Under PAM, cities are very heterogeneous in size and both measures of sorting are sharply increasing with city size. Consequently, we view the relationship between city size and these metrics as a testable prediction of positive assortative matching. Given existing empirical evidence on both outcomes (Combes, Duranton & Gobillon, 2008; Dauth et al., 2018), city-size wage premium is evidence of positive assortativeness in the economy.

In our benchmark simulation, the correlation between human capital and firm productivity is around 0.1 in the smallest cities compared to approximately 0.5 in the largest city. The average human capital level reaches 14 in the largest city, compared to 9 in the smallest city. These differences are not implausible as the fraction of college-educated in cities at the top of the education distribution is four times larger than the fraction of college-educated in cities at the bottom of the education distribution (Moretti, 2012).

Under PAM, the level of production complementarities has limited effect on the local average skill level but a strong effect on the correlation between human capital and productivity, which increases by almost 10 p.p. from $\sigma = 0.2$ to $\sigma = 0.1$, regardless of location. Altogether, these results suggest that the strength of positive sorting may be identified using either time or space variation in observational data.

Figure 4: Sorting and city size

Notes: Mean skill level $\int h(\ell_j(h)/L_j)dh$ (Graph A) and within-match correlation between skill and productivity (Graph B) by city size ($\log L_j$), for four values of the elasticity of substitution σ; all other parameters are set to benchmark values.

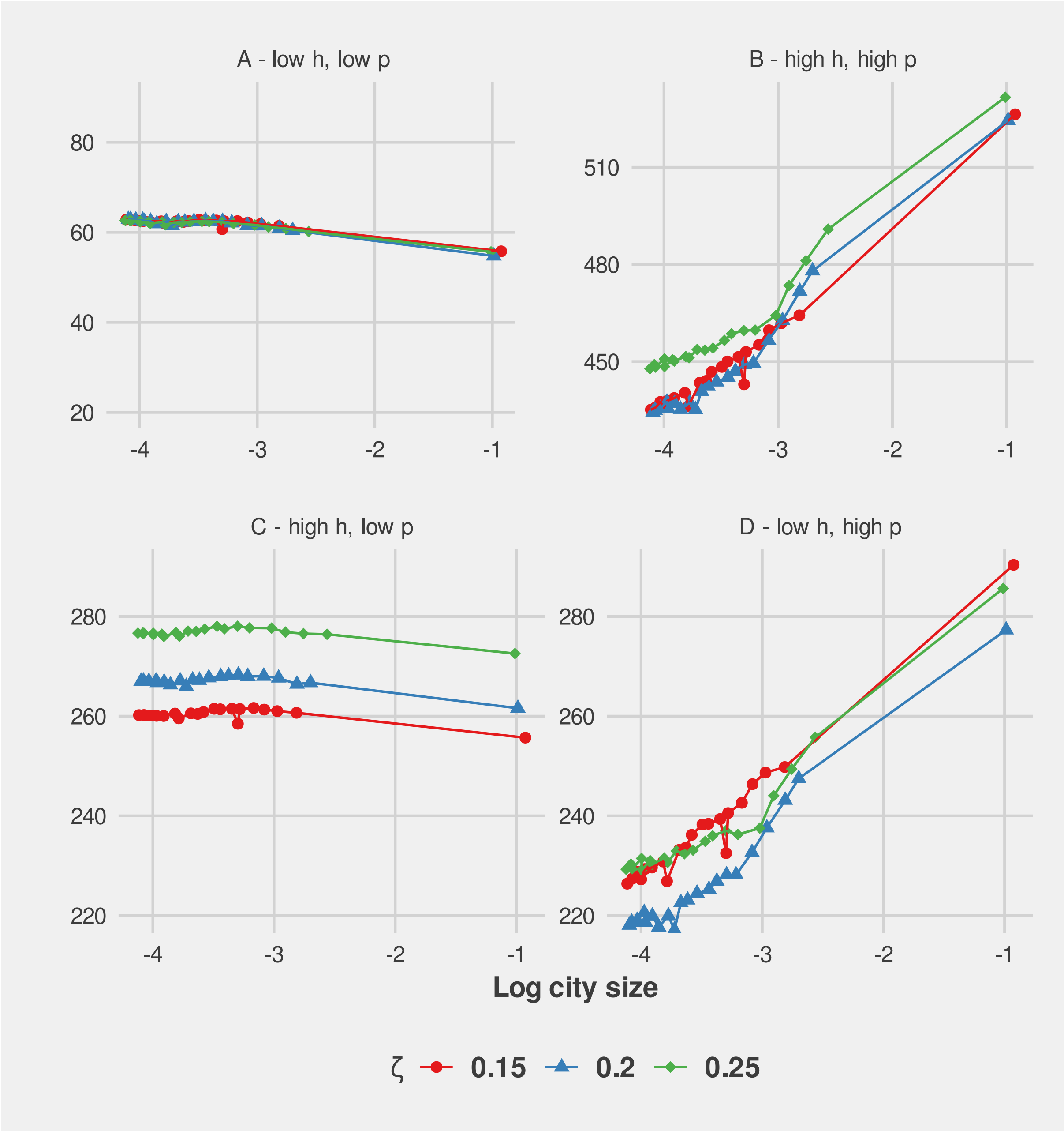
Agglomeration — Finally, we turn to agglomeration economies. They are microfunded by the existence of matching economies in a context of on-the-job search. Therefore, the technological driver of agglomeration economies in our setting is the job search efficiency of employed workers (ζ). If the probability of filling a vacancy originates from a matching technology that generates higher rates for larger cities, it is straightforward to show that employed workers will have higher financial expectations in larger cities.

However, a higher matching probability also implies a higher poaching probability, so that a firm may suffer from competition. The strength of agglomeration economies depends on the relative magnitude of these two forces, which determines whether matches in larger cities are closer to being optimal. Formally, this means that agglomeration economies are defined at the match level. There will be agglomeration economies for a type-(h, p) match if $\partial \mathbb{P}_j(h, p)/\partial L_j > 0$.

Figure 5 displays four different values of the match product, for low- and high-skill workers and for low- and high-productivity firms. For matches involving a high-productivity firm (Graphs B and D), there is a positive relationship between the match product and city size. Conversely, for matches involving a low-productivity firm (Graphs A and C), the correlation with city size is negative. Quantitatively, marginal changes in the value of ζ have little impact on the value of the match product.

Altogether, these results illustrate why an aggregate measure of agglomeration economies is not easy to interpret. According to our simulations, positive gradients are steeper than negative ones. Combined with the fact that larger cities have more productive firms, this feature should generate

Figure 5: Match product and city size



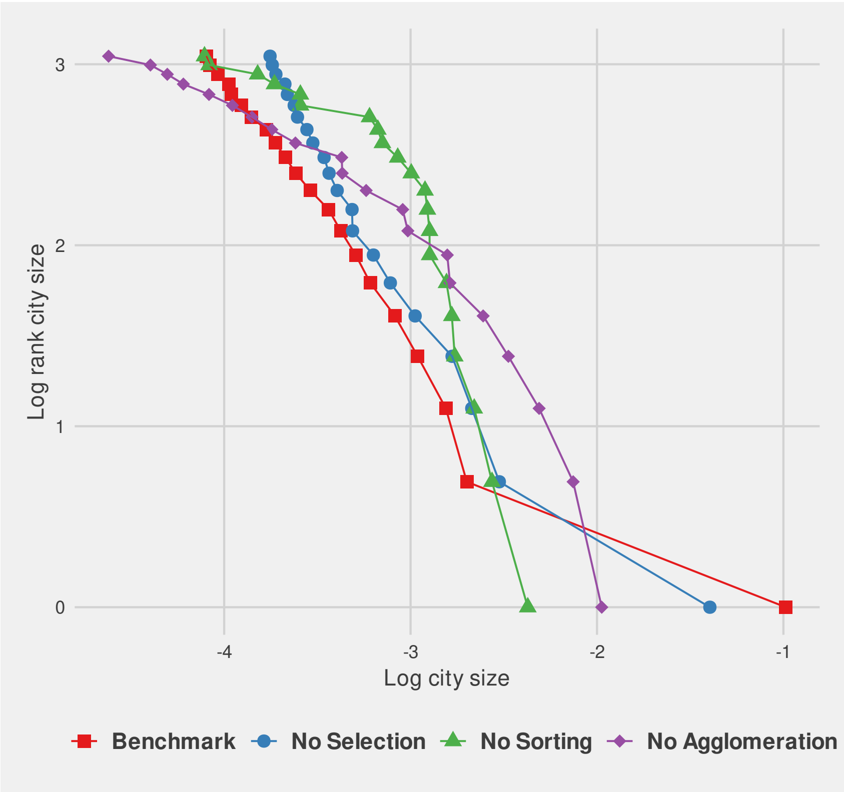
Notes: Values of match product $\mathbb{P}_j(3.2, 2.2)$ (Graph A), $\mathbb{P}_j(16.6, 16.4)$ (Graph B), $\mathbb{P}_j(16.6, 2.2)$ (Graph C) and $\mathbb{P}_j(3.2, 16.4)$ (Graph D), by city size ($\log L_j$) and for three values of the efficiency of job search by employed workers ζ; in practice, the values of h and p are chosen as either the second or the penultimate grid point; all other parameters are set to benchmark values.

buoyant agglomeration economies in the aggregate. However, we cannot rule out that it may not always be the case, for example, if the underlying productivity distribution is Pareto.

4.2 City size and agglomeration economies

Large cities are costly in many dimensions but prevalent across the world. Standard attempts at providing a microfoundation for the existence of power laws in the distribution of city size are based

Figure 6: Zipf plots: Selection, sorting and agglomeration



Notes: Log rank size - Log size relationship for our benchmark simulation, a simulation without firm selection ($c = 150$), a simulation without worker sorting ($\sigma = 100$) and a simulation without agglomeration economies ($\zeta = 0.01$); all other parameters are set to benchmark values.

on a theoretical mix, including agglomeration economies, sorting (positive assortativeness between workers and firms) and selection of more productive firms into larger cities. In this section, we open the black box that justifies the very existence of a Zipf distribution of cities. In our equilibrium model, we can isolate the effect of these channels' effect in generating a Zipf distribution of cities. More specifically, given our parameter values, we can simulate counterfactual equilibria where we shut down each mechanism separately.

Figure 6 shows the results from such exercise for our benchmark simulation and for three simulations where we either discard firm selection (by setting $c = 150$), worker sorting (by setting $\sigma = 100$) and agglomeration economies (by setting $\zeta = 0.001$). Under the benchmark calibration, the relationship is close to linear, with a unit slope. If we shut down firm selection, this linear relationship is still verified, even if the distribution displays larger small cities and smaller large cities. This suggests that firm selection does not play a crucial role in shaping the urban landscape. On the contrary, if we shut down the sorting channel by assuming a very high elasticity of substitution, size heterogeneity disappears. Finally, in the absence of agglomeration economies, cities may be quite heterogeneous in size, but the spatial equilibrium is centrifugal (see Figure C.1 in Appendix C) and the size distribution is nowhere near exhibiting a linear log-rank - log size relationship.

Nevertheless, the fact that the absence of agglomeration economies does not preclude the existence of larger cities may seem paradoxical and calls into question whether our model allows us to separately identify sorting and agglomeration. By definition, the two are intertwined: high-skill workers will only cluster in large, expensive cities if those offer some specific prospects. If the gains from agglomeration are similar to those from sorting, a statistic such as city-level average human capital will fail to differentiate sorting from agglomeration.

We investigate this issue by comparing sorting patterns in economies with and without on-the-job search. As shown in Figure C.2 in Appendix C, if we shut down on-the-job search, the patterns of spatial sorting become quite different. Graph A shows that the economy still displays a positive gradient of average human capital with respect to city size - even though the gradient is more than twice lower than in the case featuring on-the-job search. This remaining gradient, which represents the gain from sorting, makes it difficult to separately identify the two mechanisms using the average measure of skills at the city level. Yet, as illustrated in Graph B, in the absence of on-the-job search, the pattern of increasing assortativeness of matches with respect to city size is no longer verified: the correlation stays roughly constant and equal to that of the smallest cities in the benchmark case. This result confirms the potential use of city average correlations between firm and worker fixed effects (Dauth et al., 2018) as a sufficient statistic for the existence of agglomeration economies, for a given level of sorting.

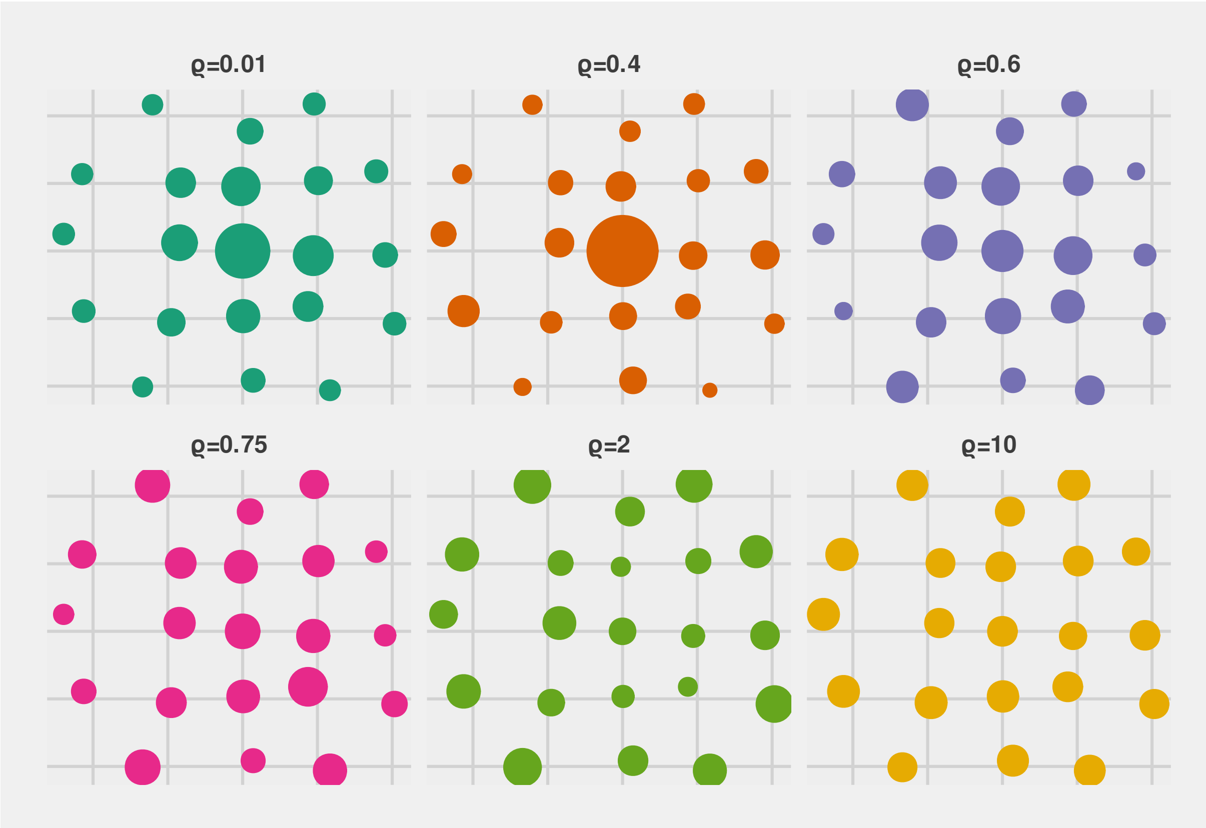
4.3 Core cities and regional clusters

Finally, we turn to the role of spatial frictions in shaping the urban network, not only in terms of size distribution but also in terms of spatial distribution of city size. Spatial frictions are summarized by the parameter ϱ, the only parameter to be directed by spatial heterogeneity in the model. When $\varrho = 0$, the labor market is perfectly integrated, while large positive values are associated with city autarky, with equilibrium city size distribution mirroring the initial condition. Under intermediate values of spatial frictions, workers and firms may allocate unevenly. Figure 7 shows maps of the distribution of the population for different values of ϱ. As soon as frictions are introduced, a core-periphery structure emerges: for $\varrho = 0.01$, the most central city reaches 22% of the population and its three closest neighbors respectively host 10%, 9% and 8% of the population, while cities located on the outer ring are halved. Minimal differences in the value of the different locations for firms pile up and generate a heterogeneous equilibrium at the steady-state.

As we keep increasing the level of spatial frictions, centrality still matters and a macropolis emerges, with a maximum population of over 40% reached around $\varrho = 0.4$, while some of the most remote cities become extremely small. However, for this level of frictions, centrifugal forces emerge as some cities interact only with their close neighbors. These forces grow in importance with even higher frictions, and when $\varrho = 0.6$, the distribution is much more even, with a central city hosting 10% of the population. When $\varrho = 0.75$, the network no longer exhibits a positive correlation between city size and centrality. Under even higher values of ϱ, the economy is no longer interconnected, and regional clusters can emerge ($\varrho = 2$). Finally, as shown for $\varrho = 10$, when frictions become so high that jobseekers cannot even hear about or apply to jobs located in their immediate surroundings, each city is bound to autarky.

These variations in city size are summarized in Figure C.3 in Appendix C, which displays the traditional Zipf relationship for different values of ϱ. Even if the associated spatial configurations

Figure 7: Spatial Configurations and spatial frictions



Notes: Spatial distribution of L_j for different values of ϱ; the spatial unit is set to 1; all other parameters are set to benchmark values.

may differ in some respect, the Zipf coefficient may be roughly the same for different levels of spatial frictions. Moderate spatial frictions ($\varrho \in \{0.05, 0.2, 0.5\}$) yield a coefficient remarkably close to unity, whereas higher, but not prohibitive, spatial frictions ($\varrho \in \{1, 5\}$) are associated with a coefficient almost twice as high. Altogether, these variations also illustrate how the model, absent local amenity or productivity idiosyncrasies or historical determination, can rationalize very different profiles of urban networks, from peripheral clusters in the US to core cities like Paris or Madrid.

5 Conclusion

In this paper, we construct a dynamic model of individual careers with an explicit spatial component. Location impacts workers' careers through localized job-finding probabilities, wage offer distributions and bargaining opportunities. The combination of the imperfect mobility of production factors, combined with positive assortative matching in production and better matching in more central cities generate a Zipf distribution of city sizes. The model is highly tractable numerically and delivers several sufficient statistics to test for the presence of agglomeration economies, sorting and selection. Future works could use a similar framework to disentangle between the three main forces behind the urban premium, and quantify the welfare effect of policies aimed at equalizing labor market outcomes across space.

References

Abowd, J. M., Kramarz, F., & Margolis, D. N. (1999). High wage workers and high wage firms. *Econometrica*, *67*(2), 251–334.

Acemoglu, D. & Hawkins, W. B. (2014). Search with multi-worker firms. *Theoretical Economics*, *9*(3).

Allen, T. & Arkolakis, C. (2019). The Welfare Effects of Transportation Infrastructure Improvements. NBER Working Papers 25487, National Bureau of Economic Research, Inc.

Bagger, J. & Lentz, R. (2018). An Empirical Model of Wage Dispersion with Sorting. *The Review of Economic Studies*, *86*(1), 153–190.

Beaudry, P., Green, D. A., & Sand, B. M. (2012). Does industrial composition matter for wages? a test of search and bargaining theory. *Econometrica*, *80*(3), 1063–1104.

Behrens, K., Duranton, G., & Robert-Nicoud, F. (2014). Productive Cities: Sorting, Selection, and Agglomeration. *Journal of Political Economy*, *122*(3), 507 – 553.

Behrens, K., Lamorgese, A. R., Ottaviano, G. I., & Tabuchi, T. (2009). Beyond the home market effect: Market size and specialization in a multi-country world. *Journal of International Economics*, *79*(2), 259 – 265.

Behrens, K. & Robert-Nicoud, F. (2015). *Agglomeration Theory with Heterogeneous Agents*, volume 5 of *Handbook of Regional and Urban Economics*, chapter 0, (pp. 171–245). Elsevier.

Blanchard, O. J. (1985). Debt, Deficits, and Finite Horizons. *Journal of Political Economy*, *93*(2), 223–247.

Burda, M. C. & Profit, S. (1996). Matching across space: evidence on mobility in the czech republic. *Labour Economics*, *3*(3), 255–278.

Burdett, K. & Coles, M. G. (1997). Marriage and class. *The Quarterly Journal of Economics*, *112*(1), 141–168.

Cahuc, P., Postel-Vinay, F., & Robin, J.-M. (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, *74*(2), 323–364.

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The Effect of Minimum Wages on Low-Wage Jobs*. *The Quarterly Journal of Economics*, *134*(3), 1405–1454.

Champernowne, D. G. (1953). A Model of Income Distribution. *The Economic Journal*, *63*(250), 318–351.

Chapelle, G., Wasmer, E., & Bono, P.-H. (2019). Spatial misallocation and rent controls. *AEA Papers and Proceedings*, *109*, 389–92.

Chetty, R., Hendren, N., Kline, P., & Saez, E. (2014). Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, *129*(4), 1553–1623.

Combes, P., Duranton, G., Gobillon, L., Puga, D., & Roux, S. (2012). The productivity advantages of large cities: Distinguishing agglomeration from firm selection. *Econometrica*, *80*(6), 2543–2594.

Combes, P.-P., Duranton, G., & Gobillon, L. (2008). Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, *63*(2), 723–742.

Combes, P.-P., Duranton, G., & Gobillon, L. (2019). The Costs of Agglomeration: House and Land Prices in French Cities. *The Review of Economic Studies*, *86*(4), 1556–1589.

Dauth, W., Findeisen, S., Moretti, E., & Südekum, J. (2018). Matching In Cities. CEPR Discussion Papers 13347, C.E.P.R. Discussion Papers.

De La Roca, J. & Puga, D. (2017). Learning by Working in Big Cities. *The Review of Economic Studies*, *84*(1), 106–142.

Duranton, G. (2007). Urban Evolutions: The Fast, the Slow, and the Still. *American Economic Review*, *97*(1), 197–221.

Duranton, G. & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J. F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, volume 4 of *Handbook of Regional and Urban Economics* chapter 48, (pp. 2063–2117). Elsevier.

Eeckhout, J. (2004). Gibrat's law for (all) cities. *American Economic Review*, *94*(5), 1429–1451.

Eeckhout, J. & Guner, N. (2017). Optimal Spatial Taxation: Are Big Cities too Small? Working Papers 1705, CEMFI.

Eeckhout, J., Pinheiro, R., & Schmidheiny, K. (2014). Spatial sorting. *Journal of Political Economy*, *122*(3), 554 – 620.

Fujita, M. & Thisse, J.-F. (2002). *Economics of agglomeration: Cities, industrial location, and regional growth*. Cambridge University Press.

Gabaix, X. (1999). Zipf's law for cities: An explanation. *The Quarterly Journal of Economics*, *114*(3), 739–767.

Gabaix, X. (2009). Power Laws in Economics and Finance. *Annual Review of Economics*, *1*(1), 255–294.

Gallin, J. H. (2004). Net migration and state labor market dynamics. *Journal of Labor Economics*, *22*(1), 1–22.

Gyourko, J. (2009). Understanding commercial real estate: How different from housing is it? *The Journal of Portfolio Management*, *35*(5), 23–37.

Hagedorn, M., Law, T. H., & Manovskii, I. (2017). Identifying equilibrium models of labor market sorting. *Econometrica*, *85*(1), 29–65.

Hsu, W.-T. (2012). Central place theory and city size distribution. *The Economic Journal*, *122*(563), 903–932.

Kennan, J. & Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, *79*(1), 211–251.

Kline, P. & Moretti, E. (2013). Place based policies with unemployment. *American Economic Review*, *103*(3), 238–43.

Kline, P. & Moretti, E. (2014). People, places, and public policy: Some simple welfare economics of local economic development programs. *Annual Review of Economics*, *6*(1), 629–662.

Lee, S. & Li, Q. (2013). Uneven landscapes and city size distributions. *Journal of Urban Economics*, *78*(C), 19–29.

Lise, J., Meghir, C., & Robin, J.-M. (2016). Matching, Sorting and Wages. *Review of Economic Dynamics*, *19*, 63–87.

Manning, A. & Petrongolo, B. (2017). How local are labor markets? evidence from a spatial job search model. *American Economic Review*, *107*(10), 2877–2907.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, *71*(6), 1695–1725.

Michaels, G. & Rauch, F. (2017). Resetting the urban network: 117–2012. *The Economic Journal*, *128*(608), 378–412.

Moretti, E. (2011). Local labor markets. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics*, volume 4, Part B chapter Chapter 14, (pp. 1237 – 1313). Elsevier.

Moretti, E. (2012). *The new geography of jobs*. Houghton Mifflin Harcourt.

Mortensen, D. T. & Pissarides, C. A. (1999). New developments in models of search in the labor market. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics*, volume 3 of *Handbook of Labor Economics* chapter 39, (pp. 2567–2627). Elsevier.

Papageorgiou, T. (2018). Worker Sorting and Agglomeration Economies. 2011 Meeting Papers 660, Society for Economic Dynamics (updated).

Petrongolo, B. & Pissarides, C. A. (2001). Looking into the black box: A survey of the matching function. *Journal of Economic Literature*, *39*(2), 390–431.

Rossi-Hansberg, E. & Wright, M. (2007). Urban structure and growth. *Review of Economic Studies*, *74*(2), 597–624.

Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics*, *125*(3), 1253–1296.

Schmutz, B. & Sidibé, M. (2019). Frictional Labour Mobility. *The Review of Economic Studies*, *86*(4), 1779–1826.

Shimer, R. & Smith, L. (2000). Assortative matching and search. *Econometrica*, *68*(2), 343–370.

Simon, H. A. (1955). On a Class of Skew Distribution Functions. *Biometrika*, *42*(3-4), 425–440.

Wrede, M. (2015). Wages, rents, unemployment, and the quality of life: a consistent theory-based measure. *Journal of Regional Science*, *4*(55), 609–625.

A Additional expressions

The wage-increase strategy — Let $\psi^i_{jl}(h,p,p')$ the wage offered by a type-p firm in j to a type-h worker already employed in this firm but who has been contacted by a type-p' firm in l but will choose to stay in her current firm. The strategy is summarized by the set: $\mathcal{P}^i_{jl}(h,w,p) = \{y : (1-\beta)\mathcal{S}_{jj}(h,p) + \beta\mathcal{S}_{jl}(h,y) \leq \mathcal{V}^e_j(h,w,p) - V^u_j(h) \leq \mathcal{S}_{jl}(h,y)\}$ and the wage increase process of an employed worker in a type-p firm in city j at wage w contacted by a type-p' firm in city l is determined by a take-it-or-leave-it strategy by the incumbent firm, that verifies:

$$\text{for} \quad p' \in \mathcal{P}^i_{jl}(h,w,p),$$
$$\mathcal{V}^e_j(h,\psi^i_{jl}(h,p,p'),p) - \mathcal{V}^u_j(h) = \widetilde{\mathcal{S}}_{jl}(h,p') \tag{17}$$
$$\Pi^f_j(h,\psi^i_{jl}(h,p,p'),p) - \Pi^v_j(p) = \widetilde{\mathcal{S}}_{jj}(h,p) - \mathcal{S}_{jl}(h,p') \tag{18}$$

Value functions for matched agents — Equation 19 describes the program of type-h workers living in city j and employed in a type-p firm for a wage w:

$$\begin{aligned}
[r+\Xi]\mathcal{V}^e_j(h,w,p) &= w - R_j \\
&- \left(\delta + \sum_{k\in\mathcal{J}} \zeta\theta_{jk}\int_{\mathcal{P}^i_{jk}(h,w,p)} v_k(y)dy + \sum_{k\in\mathcal{J}} \zeta\theta_{jk}\int_{\mathcal{P}^c_{jk}(h,p)} v_k(y)dy\right)[\mathcal{V}^e_j(h,w,p) - \mathcal{V}^u_j(h)] \\
&+ \sum_{k\in\mathcal{J}} \zeta\theta_{jk}\int_{\mathcal{P}^i_{jk}(h,w,p)} \mathcal{S}_{jk}(h,y)v_k(y)dy \\
&+ \sum_{k\in\mathcal{J}} \zeta\theta_{jk}\int_{\mathcal{P}^c_{jk}(h,p)} \left[\beta\mathcal{S}_{jk}(h,y) + (1-\beta)\mathcal{S}_{jj}(h,p)\right] v_k(y)dy
\end{aligned} \tag{19}$$

Equation 20 describes the program of a type-p firm located in city j paying a wage w to a type-h worker:

$$\begin{aligned}
r\Pi^f_j(h,w,p) &= \chi(h,p) - w - \psi R_j \\
&- \left(\delta + \Xi + \sum_{k\in\mathcal{J}} \zeta\theta_{jk}\int_{\mathcal{P}^i_{jk}(h,w,p)} v_k(y)dy + \sum_{k\in\mathcal{J}} \zeta\theta_{jk}\int_{\mathcal{P}^c_{jk}(h,p)} v_k(y)dy\right)[\Pi^f_j(h,w,p) - \Pi^v_j(p)] \\
&+ \sum_{k\in\mathcal{J}} \zeta\theta_{jk}\int_{\mathcal{P}^i_{jk}(h,w,p)} \left[\mathcal{S}_{jj}(h,p) - \mathcal{S}_{jk}(h,y)\right] v_k(y)dy
\end{aligned} \tag{20}$$

B Resolution

Algorithm — In order to solve the model, we use the following sequential algorithm:

0. Assume $\mathcal{S}_{jl}(h,p)$, $u_j(h)$, $m_j(h,p)$ and θ_j known.
1. Use accounting relationships to construct $\ell_j(h) = u_j(h) + \int m_j(h,p)dp$, $U_j = \int u_j(h)dh$ and $L_j = \int \ell_j(h)dh$.
2. Construct $E_j = \sum_{k\in\mathcal{J}} \varrho_{kj}\left[U_k + \zeta(L_k - U_k)\right]$ and use $\mathcal{M}(\cdot)$ to recover $V_j = (m_0 E_j^{m_1-1}/\theta_j)^{1/m_1}$.
3. Use accounting relationship to construct $N_j = V_j + \int\int m_j(h,p)dhdp$ and inverse housing supply function to recover $R_j = R_0 + R_1(L_j + \psi N_j)^{R_2}$.
4. Construct $\Pi^v_j(p)$ from equation 10.
5. Find $\underline{p}_j \leftarrow \Pi^v_j(\underline{p}_j) = 0$ from the zero-profit cutoff condition.
6. Use accounting relationship to recover $v_j(p) = 1_{p\geq\underline{p}_j}[N_j/(\overline{p} - \underline{p}_j) - \int m_j(x,p)dx]$.
7. Update $u_j(h)$ from steady-state constraint 12.
8. Update $m_j(h,p)$ from steady-state constraint 13.
9. Construct $\mathcal{V}^u_j(h)$ from equation 9 and $\mathbb{P}_l(h,p)$ from equation 11.
10. Update $\mathcal{S}_{jl}(h,p) = \mathbb{P}_l(h,p) - \mathcal{V}^u_j(h) - \Pi^v_l(p) - \phi_{jl}$.
11. Update θ_j from equation 16.
12. Repeat steps 1 to 11 until convergence.

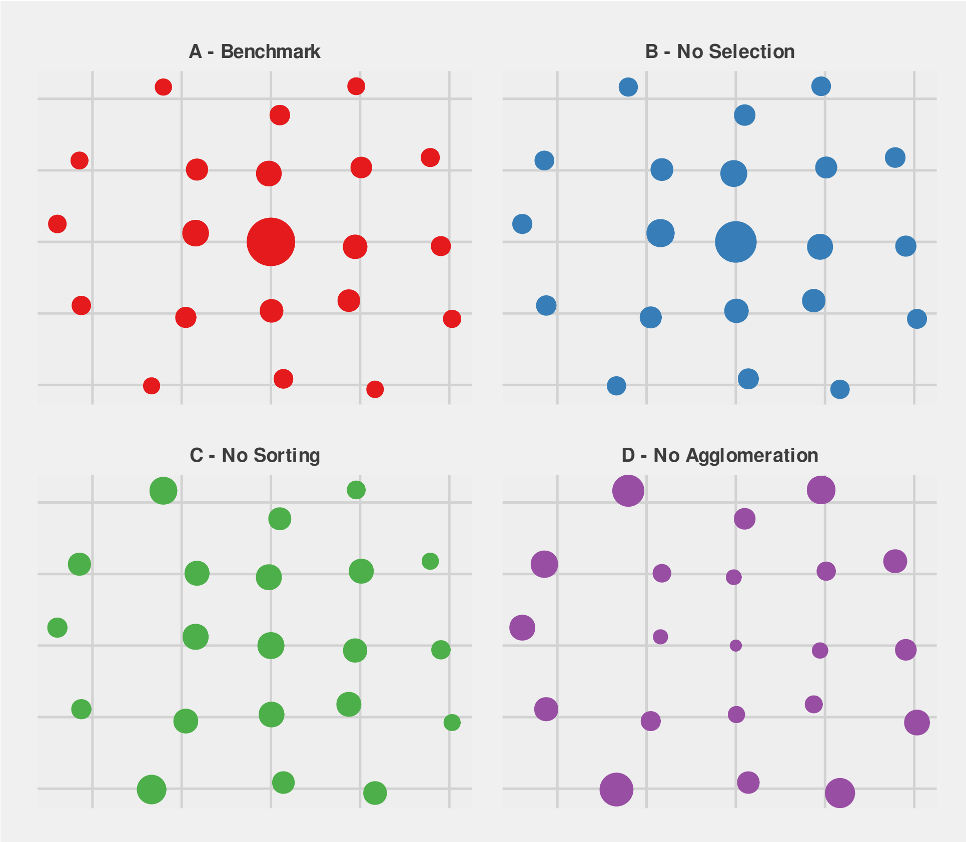
Implementation — We evaluate integrals using trapezoidal integration rule on discrete grids $h \in \mathbf{H}$ and $p \in \mathbf{P}$. Numerically, the algorithm converges at iteration n when $\Delta^{(n)} < 0.01$, with:

$$\begin{aligned}\Delta^{(n)} &= \sum_{j\in\mathcal{J}} |\theta_j^{(n)} - \theta_j^{(n-1)}| + \sum_{(j,h)\in\mathcal{J}\times\mathbf{H}} |u_j^{(n)}(h) - u_j^{(n-1)}(h)| \\ &+ \sum_{(j,h,p)\in\mathcal{J}\times\mathbf{H}\times\mathbf{P}} |m_j^{(n)}(h,p) - m_j^{(n-1)}(h,p)| + \sum_{(j,l,h,p)\in\mathcal{J}^2\times\mathbf{H}\times\mathbf{P}} |\mathcal{S}_{jl}^{(n)}(h,p) - \mathcal{S}_{jl}^{(n-1)}(h,p)|\end{aligned} \tag{21}$$

We parametrize the initial distributions of unemployment and matches to be uniform and the only heterogeneity in $\mathcal{S}_{jl}^{(0)}(h,p)$ is derived from $\chi(h,p)$.

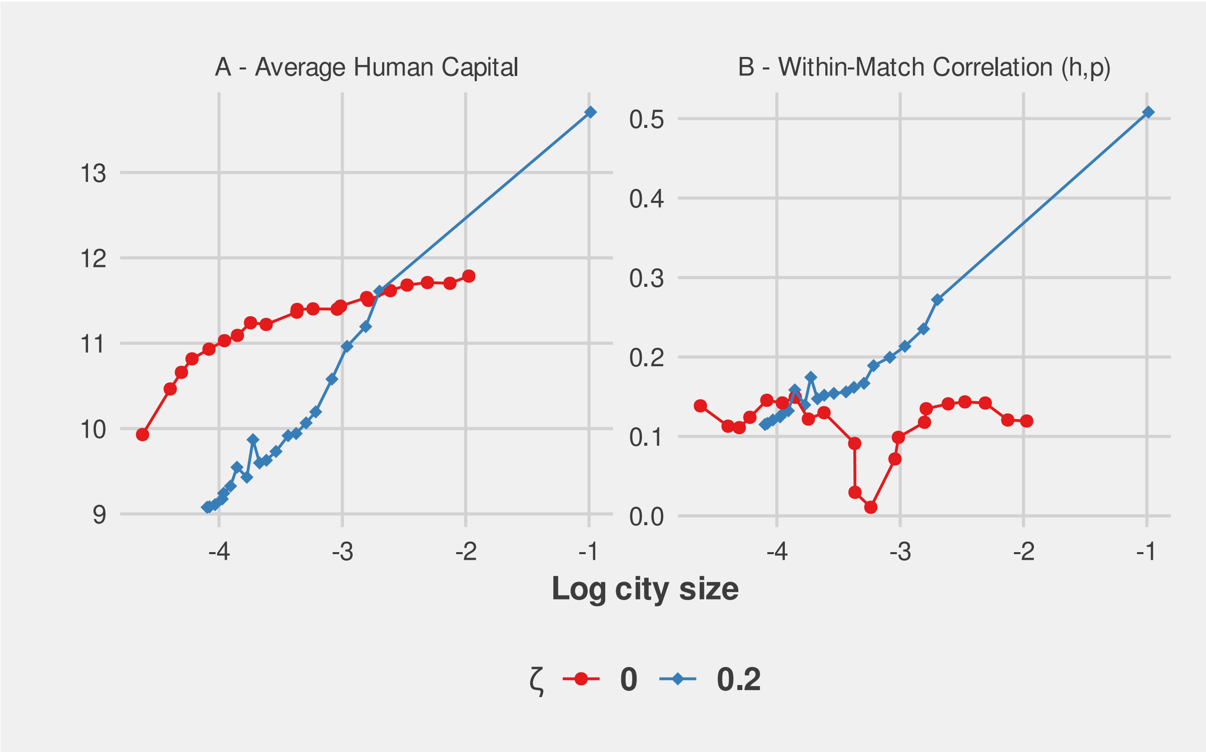
C Additional figures

Figure C.1: Spatial configurations: Selection, sorting and agglomeration



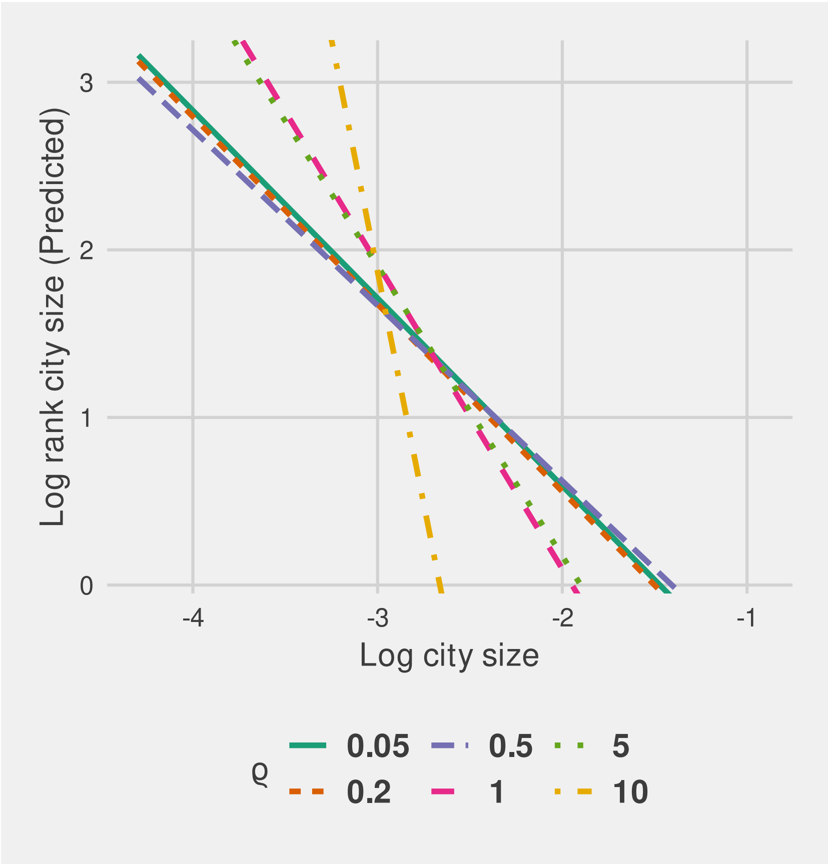
Notes: Maps of the distribution of the population for our benchmark simulation, a simulation without firm selection ($c = 150$), a simulation without worker sorting ($\sigma = 100$) and a simulation without agglomeration economies ($\zeta = 0.01$); all other parameters are set to benchmark values.

Figure C.2: Separating sorting from agglomeration



Notes: Mean skill level $\int h(\ell_j(h)/L_j)dh$ (Graph A) and within-match correlation between skill and productivity (Graph B) by city size ($\log L_j$), without or with on-the-job search; all other parameters are set to benchmark values, including elasticity of substitution $\sigma = 0.2$.

Figure C.3: Zipf slopes and spatial frictions



Notes: Predicted Zipf relationship (linear regressions) for different values of ϱ; all other parameters are set to benchmark values; the associated elasticities and R-squareds are -1.12 (97%) for $\varrho = 0.05$, -1.12 (92%) for $\varrho = 0.1$, -1.05 (88%) for $\varrho = 0.5$, -1.82 (92%) for $\varrho = 1$, -1.73 (92%) for $\varrho = 5$ and -5.54 (78%) for $\varrho = 10$.

CREST
CENTER FOR RESEARCH
IN ECONOMICS AND STATISTICS

CREST
Center for Research in Economics and Statistics
UMR 9194

5 Avenue Henry Le Chatelier
TSA 96642
91764 Palaiseau Cedex
FRANCE

Phone: +33 (0)1 70 26 67 00
Email: info@crest.science
https://crest.science/

The Center for Research in Economics and Statistics (CREST) is a leading French scientific institution for advanced research on quantitative methods applied to the social sciences.

CREST is a joint interdisciplinary unit of research and faculty members of CNRS, ENSAE Paris, ENSAI and the Economics Department of Ecole Polytechnique. Its activities are located physically in the ENSAE Paris building on the Palaiseau campus of Institut Polytechnique de Paris and secondarily on the Ker-Lann campus of ENSAI Rennes.