Local Labor Markets and Aggregate Productivity*

Paolo Martellini†

December 2022

Abstract

I propose a dynamic spatial equilibrium model that accounts for the heterogeneous labor market experience of workers across US cities. Productivity differentials between large and small cities emerge as an equilibrium outcome due to spatial sorting, increasing returns to scale in job search, and knowledge diffusion through local peer effects. The model delivers testable predictions with respect to selection into and returns to migration, which are supported by the data. I use this framework to quantify the aggregate implications of relaxing zoning regulation in large cities. I show how the resulting relocation of workers affects the size and composition of cities, the return from local interactions, and the spatial distribution of productivity. Failing to account for the impact of housing policy on local productivity would lead to overstating the magnitude of equilibrium income gains by a factor of 3.

JEL Codes: E24, J24, J31, R12

Keywords: Knowledge Diffusion, Search Frictions, Life-Cycle Income, Urban Wage Premium

*I thank Francesco Agostinelli, Anmol Bhandari, Jonathan Dingel, Alessandro Dovis, Michael Peters, Chris Tonetti, Michael Waugh, and audience members at several seminars and conferences for helpful comments and suggestions. I also thank the US Bureau of Labor Statistics for providing access to the NLSY79 confidential geocode data. All errors are my own.

†Department of Economics, University of Wisconsin-Madison. Email: martellini@wisc.edu.

1 Introduction

Vast wage differentials exist between cities, particularly between large and small ones. This observation naturally gives rise to two questions. First, what are the aggregate implications of local policies that trigger the reallocation of workers across cities? Second, how are the outcomes of such policies affected by the nature of wage differentials themselves? At one extreme, if the heterogeneous performance of cities is due solely to sorting of unequally productive workers, local policies have no effect on aggregate productivity. At the opposite extreme, if productivity is an exogenous characteristic of cities, the income gains from worker reallocation are potentially enormous. In this paper, I investigate the nature of labor productivity differentials between large and small US cities and their implications for the evaluation of local policies. I show that spatial wage premia are due both to sorting and to *endogenous* features of cities, which are in turn determined by city size and human capital composition. Consequently, local policies that trigger worker relocation across cities do affect aggregate productivity, but their quantitative effect depends on the equilibrium response of local productivity itself.

I build a spatial equilibrium life-cycle model of location choice and labor market dynamics that incorporates three sources of productivity differences between large and small cities. First, I allow for increasing returns to scale in the matching process between workers and firms. Increasing returns represent the idea that the concentration of a large number of workers and firms inside a city reduces the transportation and information frictions that are associated with the job search process and facilitate the formation of more productive matches. Second, I model human capital accumulation via a process of knowledge diffusion that is affected by the local economic environment. I allow for workers in large cities to experience more frequent interactions with one another and, to the extent that large cities host a greater fraction of high-skilled workers, to learn from better peers. Third, heterogeneous workers migrate over the life cycle and their unequal distribution across space generates spatial wage dispersion through sorting. Higher productivity in large cities is associated with higher house prices because of imperfectly elastic housing supply, hence generating the coexistence of cities of different size in equilibrium.

To identify the importance of these channels, I exploit heterogeneity in labor market outcomes between workers in large and small cities. In the data, a city-size wage premium is present at any level of labor market experience, but it significantly increases over the life cycle.

Guided by this evidence, the parameters of the knowledge diffusion technology are identified by cross-city differences in wage *growth*. Empirically, unemployment and job-to-job transition rates do not vary systematically with city size. This observation imposes a restriction on the exogenous distribution of potential matches such that increasing returns to scale in the search process induce workers and firms to be more selective with respect to the quality of matches they are willing to create. Under such restriction, matches in larger cities are endogenously better and proportionally so at all stages of labor market experience. Consequently, I show that the lower search frictions generate and are identified by the higher wage *level* in larger cities. A key contribution of this paper is to provide a sharp distinction between the role of knowledge diffusion and search efficiency in generating spatial wage differentials in a way that is consistent with labor market flows in small and large cities.

The data for the model calibration is drawn from the National Longitudinal Survey of Youth 1979 (NLSY79), which surveys a panel of US workers who entered the labor market in the 1980s and are observed for multiple consecutive years. I define large (small) cities as commuting zones with more (less) than 750,000 people in 1990. Through the lens of the model, exchanges of ideas via interactions between workers are 20% more frequent in large cities. The higher rates of knowledge diffusion in large cities leverage an equilibrium human capital distribution of peers that first-order stochastically dominates the distribution in small cities. As learning is allowed to vary by workers' education and human capital level, I find that knowledge diffusion is particularly beneficial to college graduates and, conditional on education, to workers with lower initial human capital. Consistent with increasing returns to scale in the labor market, I find that the contact rate between workers and firms in large cities is 73% higher than in small ones. This estimate is quantitatively in line with direct evidence I provide on the positive and statistically significant elasticity of the number of both applications per vacancy received by firms and applications to vacancies submitted by workers with respect to city size. It follows that exogenous TFP differences between cities—that are silent on the life-cycle *growth* in the city-size wage premium—are also unnecessary to generate its *level* under an empirically plausible estimate of the higher search efficiency in large cities.

By identifying spatial heterogeneity in the benefits from matching and knowledge diffusion, this paper delivers a theory of sorting across cities over the life cycle. The identification of the model described above is based on aggregate wage differences between large and small

cities and not on the experience of movers.[1] Thus, I validate the model according to its ability to reproduce the micro-level evidence on the wage of movers with respect to that of stayers in the origin cities and of incumbents in the destination cities. I find that the model quantitatively replicates the wage path of workers who move from small to large cities, which is steeper compared with that of stayers. Newcomers to large cities earn significantly less than incumbents, although the difference partially shrinks in the years after migration. I also show that movers from large to small cities are negatively selected in terms of pre-migration wages and that their income prospects further decline after moving, compared with those of stayers in large cities. Taken together, these results highlight how the model successfully addresses the issue of selection into—and the dynamic returns to—migration, which has proven to be a challenge to existing models that deal with the nature of spatial wage differentials.

In the last part of the paper, I explore the aggregate implications of local policies. In contrast to the traditional approach in that literature, in this paper spatial productivity differences emerge in equilibrium, making it an ideal setting to study policies that significantly change the size and composition of cities. I study the equilibrium response of the economy to a relaxation of zoning regulations in large cities in the form of an increase in housing supply elasticity. This counterfactual is motivated by the recent debate on whether place-based policies might increase aggregate income by inducing the relocation of workers toward some of the largest, most productive cities in the United States. The existing literature on the aggregate effects of local policies postulates static and exogenous gains from working in large cities, hence it is silent on the dynamic nature of the city-size wage premium and how such premium is affected by the policy itself. This paper incorporates fully endogenous local productivity differentials and dynamic benefits that are consistent with the evidence on life cycle wage growth in cities. I highlight the importance of these margins by contrasting the *equilibrium* response to the policy change with an alternative scenario in which local productivity was *exogenous* and unaffected by the policy. I show that the latter scenario overstates the percentage growth in total income by a factor of 3. When large cities further grow in size, more workers gain access to their productivity benefit. However, the workers who are attracted to large cities are, on average, both less skilled than those who locate there under strict housing regulation and more skilled than those who remain in small cities after the policy change. As

[1]The vast majority of workers never move from a small to a large city, nor vice versa. See the discussion in Section 3.5, where I show that both in the model and in the data, the identifying moments are virtually unchanged when computed on the whole sample or on the sample of non-movers only.

a result, the human capital composition of workers deteriorates in both types of cities, with additional negative consequences for local productivity due to the lower quality of peers.

Related Literature

This paper is closely related to the literature that studies productivity gains through knowledge diffusion (Lucas and Moll (2014); Perla and Tonetti (2014)), with applications in various contexts like learning from coworkers inside a firm (Herkenhoff et al. (2018); Jarosch, Oberfield, and Rossi-Hansberg (2021)) and through trade (Buera and Oberfield (2020); Perla, Tonetti, and Waugh (2021)). Contributing to this body of work, I explore the role of knowledge diffusion in accounting for the remarkable difference in life-cycle wage growth across cities. This paper shares key features with Lucas (2004), who models the long-run transition from a rural to an urban economy, the latter being characterized by a human capital–intensive technology and by learning from others. In contrast, I focus on an equilibrium in which small and large cities coexist, and I provide a quantification of the "external effect" of other workers on the process of human capital accumulation in a way that is empirically consistent with the labor market experience of workers in US cities.[2]

This paper also belongs to the literature on the determinants of wage growth over the life cycle, within and between jobs. Bagger et al. (2014) and Menzio, Telyukova, and Visschers (2016) augment standard job ladder search models with human capital accumulation through learning by doing. This paper explores how search efficiency and human capital accumulation are affected by a worker's location through increasing returns in matching and local knowledge spillovers, both of which are affected by city size and composition. Relatedly, Baum-Snow and Pavan (2012) find that search frictions are not significantly different between small and large cities, while they estimate a steeper wage profile in large cities.[3] Martellini and Menzio (2020) provide restrictions on the shape of the match quality distribution under which increasing returns to scale in the search process are consistent with the observed lack of variation in labor market flows with respect to city size. While the model in this paper satisfies

[2]The main specification in Lucas (2004) is such that everyone in the city learns only from the most skilled worker. Hence, he conjectures that a social planner would want only the "leader" to invest time in learning, in order to maximize the extent of knowledge spillovers. However, he concludes that "if one is to gain the ability to use the theory to discover ways to improve on the equilibrium, a better description of the social character of the learning process will be needed." (Lucas (2004), page S56)

[3]While abstracting from worker heterogeneity and human capital, Schmutz and Sidibé (2019) find that job-to-job transitions are a major source of wage growth within large cities in France.

those restrictions, Baum-Snow and Pavan (2012) does not, hence the different findings. I also provide external supporting evidence on the existence of lower search frictions in larger labor markets by using worker- and firm-level information on how the number of job applications varies along the city-size distribution. In addition, I replace their exogenous wage profile with an equilibrium learning model that can be used for policy analysis.

Investigating the origins of spatial wage differentials has been at the core of a vast urban economics literature, at least since the seminal paper by Glaeser and Maré (2001) documented the existence of an urban-rural wage premium in the US, which has both a level and a growth component (see also De La Roca and Puga (2017)).[4,5] Combes et al. (2012) find that agglomeration economies, in contrast to selection due to stronger competition, are responsible for the higher productivity of large cities. Davis and Dingel (2019) build a model in which large cities are locations in which high-ability workers cluster in order to share their knowledge. Although both papers address cross-sectional heterogeneity between cities, their static nature makes them silent on the life-cycle profile of the wage premium. This paper presents the first empirical equilibrium model with dynamic knowledge diffusion in cities. The mechanism I adopt is in the spirit of the two-period theoretical model in Glaeser (1999), in which homogeneous young workers learn from (skilled) old ones. By allowing for heterogeneity in human capital and learning ability in a quantitative life-cycle model, this paper speaks to the evidence on selection into and on the short- and medium-term return to migration.

A different but related literature takes the existence of spatial wage differentials as given and focuses on its aggregate implications. Hsieh and Moretti (2019) and Herkenhoff, Ohanian, and Prescott (2018) study the effect of relaxing land-use regulation in some large US cities (or states) on total output through to the relocation of workers toward more productive locations. They abstract from worker heterogeneity and human capital accumulation: the de-

[4]Eckert, Hejlesen, and Walsh (2022) address the endogeneity of workers' initial location decision by considering a sample of refugees who were randomly assigned to Danish cities. They find that while entry-level wages do not differ across space, accumulating experience in Copenhaghen is associated with higher wage growth.

[5]In their handbook chapter, Duranton and Puga (2004) list matching and learning as two of the three main sources of agglomeration economies, with input sharing as the third one. Combes, Duranton, and Gobillon (2011) review the issues involved in the identification of the city-size wage premium using reduced-form specifications, particularly with regard to the difficulties implied by mobility over the life cycle. The challenge lies in the existence of a systematic life-cycle component in both the frequency of labor market episodes—like job-to-job transitions and human capital accumulation—irrespectively of the worker's location, and in the decision to move in or out of large cities. Hence, a life-cycle model is key to recover the fundamental parameters that govern the pattern of wages and migration episodes observed in the data.

terminants of spatial productivity differentials are locations' exogenous TFP levels. I perform a policy experiment similar to theirs, and show that micro-founding the sources of spatial wage differentials reduces the equilibrium income gains that originate from worker spatial reallocation.

The results from the policy experiment emphasize the importance of connecting the literature that studies the origins of spatial wage differentials, with the literature on their macroeconomic implications. I contribute to the first by showing how to disentangle the role of matching and knowledge diffusion in the presence of worker sorting, and I bring that insight into the second literature by allowing for the benefits of working in large cities to be endogenous to the policy itself.

2 The Model

Overview

The economy is modeled in continuous time. There are two types of locations: small and large cities. Cities are inhabited by a continuum of workers with different ages, levels of education, and human capital, and by a continuum of identical firms. Inside each city, or local labor market, workers can be either employed or unemployed. Local labor markets are characterized by search frictions and heterogeneity in firm-worker match quality. Workers search both on and off the job. When they contact a firm, they observe the quality of the potential match and decide whether to form a new match. Workers also search across cities but must pay a moving cost if they migrate. Workers accumulate human capital through learning by doing and through knowledge diffusion with other workers in their same city. Both migration decisions over the life cycle and learning determine the equilibrium size and human capital composition of cities. City size is allowed to affect the extent of search frictions and the frequency of interaction between workers through increasing returns to scale in the labor market and in the process of knowledge diffusion, respectively. The human capital composition of a city determines the quality of peers other workers learn from. Workers consume a non-tradeable local good (e.g., housing) whose equilibrium price differs across cities. Local prices operate as a congestion force in workers' location choice.

Geography

The economy is represented by a star network with 1 city at the center and N cities connected to it via migration flows. Cities are ex-ante heterogeneous only with respect to their position on the network and their housing supply function, explained below. To simplify the exposition, and foreshadowing the equilibrium characterization that follows, I denote cities by $i \in \{small, large\}$, where *large* refers to the city at the center of the network.

Workers

Demographics. The economy is populated by a measure M of workers. Workers are indexed by the tuple (h, a, e). Human capital $h \in \mathcal{H} = \{h_1, h_2, ..., h_L\}$ is discrete and evolves endogenously over the life cycle. Age is denoted by $a \in \{y, o\}$, where y stands for young and o for old. Each worker is born young and turns old at (Poisson) rate ψ_y. When old, the worker exits the economy at rate ψ_o and is replaced by a new young worker in the same city. A worker's education is permanent and denoted by $e \in \{hs, col\}$. Newborns draw their education and initial human capital from a distribution with cdf $G_{0,i}(h, e)$ and probability mass function $g_{0,i}(h, e)$. The probability that a worker enters the economy with given education depends on the educational attainment of the worker being replaced. Conditional on e, the initial human capital distribution is independent of the worker's location. This implies that $G_{0,i}(h, e) = G_0^h(h|e)G_{0,i}^e$, where $G_{0,i}^e = G_0^e(\pi_i^{col,o})$ and $\pi_i^{col,o}$ is the equilibrium fraction of old college graduates in city i.[6]

Workers inelastically supply one indivisible unit of labor and maximize the present value of net flow income discounted at rate r. Net flow income is equal to $b_i h - q^e p_i$, for some constant b_i, if the worker is unemployed in city i or $\omega z h - q^e p_i$ if the worker is employed at wage $\omega z h$, described below. $q^e p_i$ is the flow cost of non-tradeable goods in city i.

Human Capital Accumulation. Young workers accumulate human capital through two channels. First, they experience "learning by doing". This form of learning is unaffected by the worker's location. From learning by doing, the human capital of an employed worker with

[6]Workers in the NLSY79 sample were administrated a test of cognitive ability known as the Armed Force Qualification Test (AFQT). The assumption I make on the workers' initial distribution is motivated by the observation that conditional on education, the average AFQT score upon entering the labor market is virtually identical in small and large cities—as already documented by De La Roca, Ottaviano, and Puga (2014) and Baum-Snow and Pavan (2012).

education e increases from h_ℓ to $h_{\ell+1}$ at the rate

$$\eta^e exp(-\eta h_\ell).$$

This formulation generates a decline in learning probability with respect to the worker's current human capital while allowing the level of the learning rate to vary by education.

The second form of human capital accumulation—which I call "knowledge diffusion", or "imitation"—is affected by geographic proximity with other workers (Jaffe, Trajtenberg, and Henderson (1993); Akcigit et al. (2018)). For example, highly populated cities may be characterized by more frequent interactions and exchanges of ideas. I define the dependence of the rate of interaction between workers on city size as the flow of ideas channel. In addition, the amount of learning that occurs through imitation depends on the type of individuals a worker interacts with. Therefore, the human capital composition of a city—or peer effects—is also a key determinant of the gains from knowledge diffusion. I model a flexible imitation technology by assuming that a worker with education e in city i experiences an increase in human capital from h_ℓ to $h_{\ell+1}$ at rate $\sigma_i \kappa(G_i(h_{\tilde{\ell}}), h_\ell, e)$, where $\sigma_i \equiv \sigma(M_i)$, for some function σ. The dependence of the learning rate on city size (measured by population), M_i, and on the equilibrium distribution of human capital in city i, $G_i(h_{\tilde{\ell}})$, represents the flow of ideas channel and peer effects, respectively. The function κ takes the form

$$\kappa(G_i(h_{\tilde{\ell}}), h_\ell, e) = \mathbb{E}_{G_i(h_{\tilde{\ell}})}[\eta^e_{ver} max\{h_{\tilde{\ell}} - h_\ell, 0\} + \eta^e_{hor} h_{\tilde{\ell}}].$$

The first term on the right-hand side is learning that occurs exclusively when interacting with more-skilled workers (*vertical imitation*), in the tradition of Lucas and Moll (2014) and Perla and Tonetti (2014). The second term implies that workers might learn from everyone, since even less-skilled workers have knowledge to transfer (*horizontal imitation*). Taken together, the technology of knowledge diffusion can be interpreted as follows. A worker of type (h_ℓ, e) in city i meets another worker of type $h_{\tilde{\ell}} \sim G_i(h_{\tilde{\ell}})$ in city i at rate σ_i, and she becomes type $h_{\ell+1}$ with probability $\eta^e_{hor} h_{\tilde{\ell}}$, if $\tilde{\ell} \leq \ell$ or with probability $\eta^e_{hor} h_{\tilde{\ell}} + \eta^e_{ver}(h_{\tilde{\ell}} - h_\ell)$, if $\tilde{\ell} > \ell$.[7,8]

[7] Since the support of the human capital distribution is bounded from above, workers of type h_L experience neither learning by doing nor knowledge diffusion. In practice, the set $\mathcal{H}$ is chosen so that only a negligible measure of workers ever achieve the human capital level h_L over the life cycle.

[8] I assume that employed and unemployed workers are equally likely to experience human capital accumulation through knowledge diffusion. This assumption can be relaxed by comparing the differential wage losses after unemployment episodes across cities. The fraction of unemployed workers being small, this measurement would likely be noisy and have minor implications for the aggregate results of this paper.

Although the focus of this paper is on how human capital accumulation is affected by a worker's location, education-specific learning by doing accounts for the spatial heterogeneity in learning that is due solely to the educational composition of cities. Absent this channel, the role of knowledge diffusion in large cities is overstated whenever college graduates are more likely to both live in large cities and experience faster wage growth in every city.

Firms

Each city is populated by a positive measure of homogeneous firms. Firms operate a constant returns to scale technology that transforms one unit of labor into zh units of output. The variable z denotes the quality of the match, which is the component of productivity that is specific to the firm-worker pair, and h is the human capital of the worker. Firms maximize the present value of their profits, $(1-\omega)zh$, discounted at rate r.

Local Labor Market

The labor market is characterized by search frictions. Workers can be either employed or unemployed.[9] An unemployed worker who lives in city i contacts a firm also located in city i at rate $\lambda_{0,i} \equiv \lambda_0(M_i)$. The dependence of λ_0 on city size is what I define the matching channel, as it captures the lower information and transportation frictions in accessing potential matches in larger cities. Employed workers in city i contact firms in their same city at rate $\lambda_{1,i} \equiv \lambda_1(M_i) = \rho\lambda_{0,i}$, where $\rho \in [0,1)$ is the relative search efficiency on the job. Upon meeting, the firm-worker pair draws a match quality $\hat{z} \sim F(\hat{z})$, where $\hat{z} \in [\underline{z}, \bar{z}]$, $\underline{z} > 0$, and $\bar{z} \leq \infty$. If the pair decides to form a match, it starts producing; otherwise, they keep searching. Jobs are exogenously destroyed at rate δ^e.

While in the first part of the paper I take the meeting rates λ as a given parameter, in Section 4, I introduce a zero-profit condition in the market for vacancies to allow firms' hiring behavior to endogenously respond to policy changes.

[9]The model abstracts from flows into and out of the labor force. In the empirical section, I pool unemployed workers and those out of the labor force into a single category that I simply refer to as unemployment. As stated in the quantitative section, I calibrate the model on a representative sample of young males, for whom fluctuations in labor force participation are less likely to be a concern.

Location Choice

Workers in city i also contact firms located in a different city, $-i$. The meeting rate between an unemployed worker in city $i = small$ and a firm in city $-i = large$ is denoted by $\lambda^*_{0,-i} \equiv \lambda^*_0(M_{-i}) = \rho^* \lambda_{0,-i}$. The meeting rate between an unemployed worker in city $i = large$ and a firm in each one of the N cities $-i = small$ is denoted by $\frac{1}{N}\lambda^*_{0,-i} \equiv \frac{1}{N}\lambda^*_0(M_{-i}) = \frac{1}{N}\rho^* \lambda_{0,-i}$ so that the overall cross-city meeting rate for workers in the large city is equal to $\rho^* \lambda_{0,small}$.[10] The parameter ρ^* captures the extent of search frictions across—rather that within—cities. This specification corresponds to the intuition that if a larger city generates more job offers for its current residents, it does so for workers in a different city as well, although possibly at a proportionally lower rate.

Upon meeting, the firm-worker pair observes the quality of the match $\hat{z} \sim F(\hat{z})$. In contrast to meetings that occur inside a given city, the worker also draws a *iid* migration cost $c \sim D(c)$. According to the realization of $(\hat{z}, c)$, one of the following events occurs: i) the worker migrates and the match is formed; ii) the worker migrates, but the match is not formed, in which case the worker becomes unemployed in city $-i$ (or, equivalently, the worker migrates, the match is formed and endogenously destroyed immediately); or iii) the worker does not migrate and remains in the current city and employment status.

It is worth motivating why the economy is populated by $N > 1$ small cities, even though I focus on an equilibrium in which they are all identical to one another. Under the classification I adopt in the quantitative section, the number of small cities in the US economy is much larger than the number of large cities. Therefore, while there are only two types of cities in equilibrium, setting $N > 1$ allows me to aggregate city-level outcomes when evaluating the macroeconomic impact of local policies.

Housing Market

Each city i is characterized by a supply function for housing,

$$p_i = p_{0,i} Q_i^{\gamma_i} \qquad p_{0,i} > 0, \gamma_i \in \mathbb{R},$$

[10] Recall that the economy is represented by a star network in which all the small cities are connected to a large city, but not to each other. Adding migration links between small cities would not qualitatively change the results of this paper. Perhaps the main difference would lie in an increase in small-city wages due to additional sampling of matching opportunities, which would require stronger increasing returns to scale in the search process in order to rationalize the observed city-size wage premium. This insight is more easily understood in light of the quantitative section that follows.

where γ_i is the (inverse) elasticity of housing supply.[11] The housing stock is owned by absentee landlords outside the economy. A worker of education e located in city i inelastically demands q^e units of housing; hence, she incurs a flow payment $q^e p_i$. While I refer to p_i as house price, or local price, for simplicity, p_i captures both non-tradeable consumption and the share of tradeable goods whose price is affected by local house prices.

Contracts

The contracts offered by firms to workers are assumed to be sufficiently flexible so that the outcome of the matching process is bilaterally efficient in the sense that the joint value of a match—that is, the sum of the present discounted value of the firm's profit and the worker's utility—is maximized. As a consequence, the allocation of workers across jobs and cities does not depend on the history of wages. Many contractual environments satisfy this convenient property (see Menzio and Shi (2011) for some examples). In this paper, I follow the approach in Bagger et al. (2014) and assume that a worker is paid a fraction ω of her productivity—a so-called "piece rate"—so that her wage is equal to $\omega z h$. The wage-setting mechanism is as follows. When a worker is hired, either from unemployment or through a job-to-job transition, the equilibrium piece rate is the unique value of ω that solves the Nash bargaining problem with worker's bargaining power β. Over the course of the employment relationship, the piece rate is revised upward to match the outside offers obtained via search on the job, if the latters are higher that the value the worker is currently receiving. Notice that migration entails a one-time cost c that must be subtracted from the value of any new potential match that requires a change of city.[12] Additional details are in Online Appendix A.

[11] The assumption of constant elasticity housing supply at the city level is common in the urban literature (see, among others, Hsieh and Moretti (2019)) and it allows me to map the effect of land use regulation on house prices through a single parameter, γ_i.

[12] To keep the model simple, I rule out the possibility of quitting a job while remaining in the same city. This assumption is quantitatively innocuous, as I find that only 0.03% of workers are employed in a match they would not have formed to begin with. Without this assumption, quitting would be observed in equilibrium after a worker becomes old or she accumulates human capital whenever i) the reservation match quality is increasing in age or in human capital for at least certain values of h, and ii) the worker's match quality is sufficiently close to its reservation value. Besides, if quitting is allowed, unemployment represents a credible outside option that might trigger an increase in the piece rate upon aging or learning. This possibility would complicate the analysis without adding any relevant insight. What makes this restriction quantitatively negligible is the fact that the reservation match quality is almost constant in age and in the level of human capital and that workers leave marginal matches at a sufficiently high rate.

2.1 Definition of a Stationary Equilibrium

I define a stationary equilibrium for this economy. As previously mentioned, I restrict attention to equilibria in which the N small cities are all identical to each other in terms of size and composition. Let $U(h_\ell, a, e, i)$ be the present discounted value of income of an unemployed worker of human capital h_ℓ, age a, and education level e who lives in city i. Let $S(h_\ell, a, e, i, z) \equiv V(h_\ell, a, e, i, z) - U(h_\ell, a, e, i)$ be the surplus of a match of quality z, where V is the sum of the present discounted value of utility to the worker and profit to the firm. Let $\Delta(h_\ell, a, e, i, -i, c) \equiv U(h_\ell, a, e, -i) - U(h_\ell, a, e, i) - c$ be the net gain in unemployment value upon migrating from i to $-i$ given migration cost c.

The value of unemployment satisfies the Hamilton-Jacobi-Bellman (HJB) equation,

$$\begin{aligned} rU(h_\ell, a, e, i) = b_i h_\ell - q^e p_i + \\ \sigma_i \kappa(G_i, h_\ell, e)[U(h_{\ell+1}) - U]\mathbb{1}\{a = y\} + \psi_a[U(o)\mathbb{1}\{a = y\} - U] + \\ \lambda_{0,i}\beta \mathbb{E}_F[\max\{S(\hat{z}), 0\}] + \\ \lambda^*_{0,-i}\mathbb{E}_{F,D}[\max\{\beta S(-i, \hat{z}) + \Delta(i, -i, c), \beta(S(-i, \hat{z}) + \Delta(i, -i, c)), \Delta(i, -i, c), 0\}]. \end{aligned} \tag{2.1}$$

For ease of notation, I omit the dependence of the value and policy functions on the right-hand side from the individual states that are the same as on the left-hand side. The LHS of Equation (2.1) is the annuitized value of unemployment. The first line on the RHS is the flow payoff of unemployment net of the house price. The second line shows the gains from knowledge diffusion for the young and the transition to old age/exit. The third line shows the option value of searching in the worker's current city, which is equal to the rate at which an unemployed worker meets a firm, multiplied by the fraction of surplus that accrues to the worker if the match is formed. Since the surplus of a match is strictly increasing in its quality, the decision to create a match generates a cutoff $R(h_\ell, a, e, i)$, which satisfies

$$S(h_\ell, a, e, i, R(h_\ell, a, e, i)) = 0. \tag{2.2}$$

The last line of Equation (2.1) describes the event in which the worker contacts a firm in city $-i$. According to the realization of ($\hat{z} \sim F$, $c \sim D$), the gain from a cross-city meeting is equal to the maximum between four terms, from left to right: i) moving as employed to city $-i$, having unemployment in city $-i$ as the outside option in the bargaining protocol (high $\hat{z}$, low c); ii) moving as employed to city $-i$, having unemployment in city i as the outside

option (high $\hat{z}$, high c); iii) moving as unemployed to city $-i$ (low $\hat{z}$, low c); or iv) not moving (low $\hat{z}$, high c). The max operator in the fourth line pins down the migration-cost threshold, $x(h_\ell, a, e, i, z, z^*)$ that a worker of type (h_ℓ, a, e) in city i, with match quality z is willing to pay to move to city $-i$ in a job of quality z^*. The cutoff migration cost satisfy

$$\Delta(h_\ell, a, e, i, -i, x(h_\ell, a, e, i, 0, 0)) = 0 \tag{2.3}$$

$$S(h_\ell, a, e, -i, z^*) + \Delta(h_\ell, a, e, i, -i, x(h_\ell, a, e, i, 0, z^*)) = 0 \tag{2.4}$$

where a match quality of 0 stands for unemployment.

The HJBE (2.5) describes the surplus of a firm-worker pair of quality $z \geq R(h_\ell, a, e, i)$,

$$\begin{aligned}
&rS(h_\ell, a, e, i, z) \\
&= (z - b_i)h_\ell + \psi_a[S(o)\mathbb{1}\{a = y\} - S] + \\
&[\sigma_i \kappa(G_i, h_\ell, e)(S(h_{\ell+1}) - S) + \eta^e exp(-\eta h_\ell)(S(h_{\ell+1}) - S + U(h_{\ell+1}) - U)]\mathbb{1}\{a = y\} - \\
&\delta^e S + \lambda_{1,i}\beta\mathbb{E}_F[\max\{S(\hat{z}) - S, 0\}] - \lambda_{0,i}\beta\mathbb{E}_F[\max\{S(\hat{z}), 0\}] + \\
&\lambda^*_{1,-i}\mathbb{E}_{F,D}[\max\{\beta S(-i, \hat{z}) + \Delta(-i, i, c) - S, \beta(S(-i, \hat{z}) - S + \Delta(-i, i, c)), \Delta(-i, i, c) - S, 0\}] - \\
&\lambda^*_{0,-i}\mathbb{E}_{F,D}[\max\{\beta S(-i, \hat{z}) + \Delta(i, -i, c), \beta(S(-i, \hat{z}) + \Delta(i, -i, c)), \Delta(i, -i, c), 0\}].
\end{aligned} \tag{2.5}$$

As most events are common to employed and unemployed workers, I highlight only the main differences between the terms in Equation (2.5) and their counterparts in Equation (2.1). The first term on the RHS is the flow payoff of the firm-worker pair minus the flow payoff of unemployment. The human capital accumulation process in the second line is analogous to the process for unemployed workers, except for the presence of learning by doing. The third line shows the loss of surplus from exogenous job destruction and the net gain from searching on—compared with off—the job in city i. The last two lines describe the gains from migration while employed net of the gains as unemployed. The decision in the fourth line pins down the migration cost thresholds when employed

$$\Delta(h_\ell, a, e, i, -i, x(h_\ell, a, e, i, z, 0)) - S(h_\ell, a, e, i, z) = 0 \tag{2.6}$$

$$S(h_\ell, a, e, -i, z^*) + \Delta(h_\ell, a, e, i, -i, x(h_\ell, a, e, i, z, z^*)) - S(h_\ell, a, e, i, z) = 0. \tag{2.7}$$

Labor market and location decisions, learning, and aging induce a distribution of workers, ϕ, over the state space. The distribution determines the frequency of meetings between workers and firms (matching channel), the rate of knowledge diffusion (flow of ideas channel), the

human capital distributions of cities (peer effects), and house prices.

The Kolmogorov Forward equation, KFE, (2.8) describes the law of motion of the measure of unemployed workers, $\phi(h_\ell, a, e, i, 0)$ where, once again, I omit the arguments of ϕ that are unaffected by a given event,

$$\begin{aligned} 0 = {} & \sigma_i[\kappa(h_{\ell-1})\phi(h_{\ell-1}) - \kappa(h_\ell)\phi(h_\ell)]\mathbb{1}\{a = y\} + \psi_y\phi(y)\mathbb{1}\{a = o\} - \psi_a\phi + \psi_o g_{0,i} M_i(o)\mathbb{1}\{a = y\} - \\ & \frac{F(R)}{N_i}\left[\lambda^*_{0,i} D(x(-i,0,0))\phi(-i,0) + \lambda^*_{1,i}\int_{\underline{z}}^{\bar{z}} D(x(-i,\hat{z},0))\phi(-i,\hat{z})d\hat{z}\right] - \\ & \lambda^*_{0,-i}\phi\left[F(R(-i))D(x(0,0)) + \int_{R(-i)}^{\bar{z}} D(x(0,\hat{z}))dF(\hat{z})\right] + \lambda_{0,i}\phi[1 - F(R)] + \delta\int_{\underline{z}}^{\bar{z}}\phi(\hat{z})d\hat{z}. \end{aligned} \tag{2.8}$$

The LHS is the derivative of the distribution with respect to time, which is 0 in a steady state. The first line on the RHS states that learning through imitation induces an outflow (inflow) of young workers of human capital h_ℓ ($h_{\ell-1}$). It also shows the inflow and outflow of workers due to aging and the entry of young workers into the labor market. Newborns are distributed according to $g_{0,i}$ and they replace old workers in city i—whose measure is denoted by $M_i(o)$—at rate ψ_o. The second line represents the inflow of workers from city $-i$, who are either unemployed or employed at some match quality $\hat{z}$ before moving to city i as unemployed. The term $1/N_i$ accounts for the fact that while all the mobility from small cities happens toward the single large city in the economy, workers in large cities are equally likely to receive an offer from any of the N small cities. Hence, $N_i = 1$ if $i = large$, while $N_i = N$ if $i = small$. The first term on the third line describes the migration flow in the opposite direction, that is, unemployed workers moving into city $-i$, either as unemployed or as employed at match quality $\hat{z}$. The last two terms of the equation show the flow out of unemployment by workers who accept a job offer within city i and the inflow into unemployment by workers whose jobs are destroyed.

Equation (2.9) is the KFE for the measure of employed workers at match quality z, $\phi(h_\ell, a, e, i, z)$,

$$
\begin{aligned}
0 = &\{[\sigma_i \kappa(h_{\ell-1}) + \eta^e exp(-\eta h_{\ell-1})]\phi(h_{\ell-1}) - [\sigma_i \kappa(h_\ell) + \eta^e exp(-\eta h_\ell)]\phi(h_\ell)\}\mathbb{1}\{a = y\} + \\
&\psi_y \phi(y)\mathbb{1}\{a = o\} - \psi_a \phi + \\
&\lambda_{0,i}\phi(0)f(z)\mathbb{1}\{z \geq R\} + \lambda_{1,i}\int_{\underline{z}}^{z} \phi(\hat{z})d\hat{z}f(z) - \lambda_{1,i}\phi(1 - F(z)) - \delta^e \phi + \\
&\frac{1}{N_i} f(z)\mathbb{1}\{z \geq R\}\left[\lambda^*_{0,i} D(x(-i, 0, z))\phi(-i, 0) + \lambda^*_{1,i}\int_{\underline{z}}^{\bar{z}} D(x(-i, \hat{z}, z))\phi(-i, \hat{z})d\hat{z}\right] - \\
&\lambda^*_{1,-i}\phi\left[F(R(-i))D(x(z, 0)) + \int_{R(-i)}^{\bar{z}} D(x(z, \hat{z}))dF(\hat{z})\right].
\end{aligned}
\tag{2.9}
$$

Equation (2.9) is analogous to Equation (2.8), except for few differences. First, employed workers accumulate human capital through learning by doing, not just through knowledge diffusion. Second, the measure of employed workers at match quality z increases because of both hiring from unemployment and on-the-job search from $\hat{z} < z$ (third line, first two terms). The measure decreases because of transitions to jobs with quality above z and to unemployment (third line, last two terms). The fourth line shows the inflow of workers from city $-i$ into jobs of match quality z in city i. The last line shows the outflow of workers toward city $-i$, either as unemployed or as employed at match quality $\hat{z}$.

Integrating over the steady-state distribution ϕ, we obtain the measure of workers of age a in city i, $M_i(a)$; the share of workers of age a in city i with a college degree, $COL_i(a)$; and the equilibrium human capital distribution in city i, $G_i(h_\ell)$. Last, the equilibrium price in the housing market is given by

$$
p_i = p_{0,i} Q_i^{\gamma_i} = [(q^{col} COL_i + q^{hs}(1 - COL_i))M_i]^{\gamma_i}. \tag{2.10}
$$

The equations in this section lead to the following definition.

Definition. A stationary equilibrium is a tuple $\{U, S, R, x, \phi, p\}$ of value and policy functions, equilibrium distribution, and house prices such that i) U solves the HJBE (2.1); ii) S solves the HJBE (2.5); iii) R satisfies the optimality condition (2.2), iv) x satisfies the optimality conditions (2.3), (2.4), (2.6), and (2.7); v) ϕ solves the system of KFEs (2.8)–(2.9); and vi) p satisfies (2.10).

2.2 Discussion

The model includes a number of simplifying assumptions that are motivated by both data availability and computational tractability. Before turning to the quantitative analysis of the

model, I briefly provide motivating evidence in support of two assumptions that are directly related to the nature of productivity differences between cities. The results in this section are based on data from the decennial census. Further details are in Online Appendix B.

First, one might conjecture that heterogeneity in occupational composition across cities could play a substantial role in accounting for the city-size wage premium and its life-cycle dynamics. The limited size of the panel dataset used in the calibration limits the number of worker characteristics whose spatial variation can be measured with reasonable accuracy. To gauge the importance of this omission, I verify whether the coefficient on city-size in a canonical wage regression is smaller once the worker's (3-digit) occupation is controlled for. I find that such a drop in the correlation between wages and city size (both in level and growth rate) is present in the data, but it is only slightly larger than the drop associated with controlling for workers' education (a coarser measure of workers' ability). In addition, the heterogeneity in the human capital composition of cities in the model subsumes worker heterogeneity above and beyond differences in education, including, for example, the heterogeneity across workers in different occupations.

Second, the learning technology features randomness in the type of workers that interact with one another within a city. While it is reasonable to believe that workers can partially direct their search for peers, I provide suggestive evidence against a perfect segregation by, say, level of education. I find that college graduates experience higher wages *growth* in larger cities, but particularly so in cities with a higher share of college graduates in total population.[13] A simple back-of-the-envelope calculation shows that "adding" one high school graduate to a city is associated with observing lower wages for college graduates: the increase in city size is more than offset by the reduction in the college share in the city. Through the lens of the model, if the knowledge diffusion process were perfectly excludable across education types, an increase in the number of high school graduates would have no correlation—or a positive one—with the wage of college graduates in that city.

3 Quantitative Analysis

The model does not have a closed-form solution due to the interaction between labor market dynamics, human capital accumulation, and location choice. I solve it numerically using

[13] Abstracting from wage growth, the same point in terms of wage *levels* has been made in recent static models by Rossi-Hansberg, Sarte, and Schwartzman (2019) and Diamond (2016).

a standard finite difference method in the spirit of Achdou et al. (2022).[14] After describing the data (Section 3.1) and specifying the parametric assumptions used in the calibration (Section 3.2), I discuss the identification of the model parameters (Section 3.3). A subset of model parameters are set externally. The remaining parameters are chosen to match key differences between labor market outcomes *within* large and small cities.[15] The results of the calibration are presented in Section 3.4. In Section 3.5, I validate the model by verifying its ability to replicate the wage difference between stayers and movers *across* cities in the years before and after migrating. Last, I decompose the observed wage premium into the contribution of sorting, matching, flow of ideas, and peer effects (Section 3.6).

3.1 Data

The main source of data for this paper is the NLSY79. The NLSY79 is a survey of young men and women who were between 14 and 21 years old on December 31, 1978. The survey comprises a "cross-sectional" subsample that is representative of the US population and other subsamples that target specific demographic groups. Interviews took place annually from 1979 until 1994 and biennially thereafter. For each respondent, the NLSY79 contains information on education, employment status, job transitions, wages, and location. In order to have a homogeneous sample and avoid dealing with issues related to labor force participation, I use only information on men from the cross-sectional subsample.[16] Further sample restrictions involve dropping individuals who entered the labor force before January 1, 1978, which is the first date for which labor market information is available, or have not worked for at least 20 years by 2012. I also drop individuals with a significant amount of missing information on education, job history, or location.[17] I map the respondents' county of residence into commut-

[14] Computing the equilibrium of the economy amounts to solving a standard system of HJB and KF equations. The individual states in Achdou et al. (2022) are replaced with human capital, match quality, age, and location. The competitive market prices they solve for are replaced with the relevant aggregate variables in this model: city size, house prices, and distribution of peers.

[15] Newborns in the model are all unemployed, while they are employed in the data. To bring the model closer to the data, I follow an approach similar to Lise and Postel-Vinay (2020). I simulate a pre-sampling period of six months in which I shut down all events that can happen to a worker except meeting a firm in the local labor market. I then drop the pre-sampling period and consider the worker's state at the end of it as the initial condition.

[16] Among other things, the model abstracts from the colocation problem of "power couples" who tend to disproportionally live in large cities (Costa and Kahn (2000)). While certainly related to worker mobility across cities, joint job search problems are outside the scope of this paper.

[17] Additional details about the sample are provided in Online Appendix C.

ing zones (CZs) following the methodology developed by Tolber and Sizer (1996).[18] Because of the small number of observations associated with each CZ, I then group all the CZs in the sample into two size categories according to their total population in 1990. One group contains CZs with less than 750,000 individuals, and the other group contains CZs with more than 750,000 individuals. I refer to these groups as small and large cities, respectively. The threshold is chosen to guarantee both substantial heterogeneity between groups and similar group size in the NLSY79. The final sample contains information on the labor market experience of 1,532 men who are observed for 20 consecutive years since the month they started working after completing formal education.

3.2 Parametric Assumptions

Local Labor Market. The meeting rate between unemployed workers and firms, both located in city i, is given by the constant elasticity function

$$\lambda_{0,i} \equiv \lambda_0(M_i) = \chi M_i^{\chi_M}, \qquad \chi > 0,\ \chi_M \in \mathbb{R}.$$

The parameter χ_M represents the degree of returns to scale in the search process.

Match Quality Distribution. Upon meeting, a firm-worker pair samples a match quality

$$z \sim Pareto(\underline{z}, \alpha) \qquad \underline{z} > 0,\ \alpha > 1,$$

with cdf $F(z)$, where $\underline{z}$ is the lower bound of the support of the distribution, and α is the tail coefficient. The choice of a Pareto distribution is motivated in Section 3.3.

Human Capital Accumulation. The rate of interaction between workers in city i is equal to

$$\sigma(M_i) \equiv \tau M_i^{\tau_M}, \qquad \tau > 0,\ \tau_M \in \mathbb{R}.$$

Analogously to χ_M, τ_M captures the degree of returns to scale in knowledge diffusion.

Migration Cost. Upon migrating, workers pay a cost $c \sim Logistic(\mu_c, \sigma_c)$, with cdf $D(c)$. The parameters μ_c and σ_c represent the mean and standard deviation of the distribution, respectively. Similarly to the Normal, the Logistic distribution is symmetric about the mean and can be fully described by two parameters. The property of having a closed-form conditional

[18] To compute certain aggregate variables (e.g., city size) I use larger datasets like the decennial census. When information on Public Use Metrodata Areas (PUMAs), but not county, is available, I use the crosswalk between PUMAs and CZs developed by David Dorn (https://www.ddorn.net/data/Dorn_Thesis_Appendix.pdf)

expectation makes the Logistic distribution computationally convenient to model endogenous migration decisions with a fixed cost.

Initial Human Capital Distribution. Let $\hat{h}|e \sim LogN(\mu^e, \sigma^e)$ be a random variable with pdf $\hat{g}_0^h(\hat{h}|e)$. Let π^e be the fraction of workers in the economy with education e and $\pi^{e,e'}$ be the probability that a worker with education e is replaced by one with education e'. Newborns' human capital and education are distributed according to

$$g_{0,i}(h_\ell, e) = g_0^h(h_\ell|e) g_{0,i}^e,$$

where

$$g_0^h(h_\ell|e) = \frac{\hat{g}_0^h(h_\ell|e)}{\sum_{\ell=1}^{L} \hat{g}_0^h(h_\ell|e)}, \tag{3.1}$$

$$g_{0,i}^{col} = [\pi^{col,col} COL_i(o) + \pi^{hs,col}(1 - COL_i(o))]\hat{\pi}. \tag{3.2}$$

Equation (3.1) states that the conditional distribution of initial human capital is assumed to be approximately log-normal (Huggett, Ventura, and Yaron (2006)) on a finite set of points with parameters μ^e and σ^e that are allowed to vary by education. As Equation (3.2) describes, for each old worker who exits the economy in city i, the education of the newborn in the same city is governed by the transition probability $\pi^{e,e'}$. The normalizing factor $\hat{\pi}$ guarantees that the total fraction of college graduates in the economy is equal to π^{col}.[19]

3.3 Identification

The model is calibrated at monthly frequency in order to replicate the behavior of some high-frequency events, like unemployment-to-employment and job-to-job transitions. Even though all parameters are either set externally or jointly calibrated within the model, I provide an intuitive identification argument that clarifies how specific empirical moments inform certain model parameters. Overall, 25 parameters are internally calibrated using 30 targeted moments (Tables 3.1 and 3.2 below).

Human Capital Accumulation and Aging. The rate of human capital accumulation through knowledge diffusion is $\sigma(M_i)\kappa(G_i(h_{\tilde{\ell}}), h_\ell, e)$, where $\sigma(M_i) = \tau M_i^{\tau_M}$ and $\kappa(G_i(h_{\tilde{\ell}}), h_\ell, e) =$

[19] The normalizing factor is given by

$$\hat{\pi} = \frac{\pi^{col} \sum_i N_i M_i(o)}{\sum_i N_i M_i(o)[\pi^{col,col} COL_i(o) + \pi^{hs,col}(1 - COL_i(o))]}.$$

$\mathbb{E}_{G_i(h_{\tilde{\ell}})}[\eta^e_{ver} max\{h_{\tilde{\ell}} - h_\ell, 0\} + \eta^e_{hor} h_{\tilde{\ell}}]$. Since κ is a constant returns to scale function of the parameters η^e_{hor} and η^e_{ver}, I normalize $\tau = 1$. Hence, the set of parameters related to human capital accumulation is given by $\{\tau_M, \eta^{hs}_{hor}, \eta^{col}_{hor}, \eta^{hs}_{ver}, \eta^{col}_{ver}, \eta^{hs}, \eta^{col}, \eta\}$. To calibrate these eight parameters, I define eight groups of workers. First, I split the sample by education group (*hs* and *col*). Then, within each category, workers are divided into those with wages in the top and bottom half of the wage distribution in the first year of employment. Last, for every level of experience, each of the four groups is partitioned into workers located in large cities and those located in small ones.

The average wage *growth* rate for workers in each group provides a moment that is used in the calibration. Learning by doing is described by the parameters η^{hs}, η^{col}, and η. Intuitively, η^{hs} and η^{col} determine the average wage growth by education group, and η reproduces the curvature of wage profiles, irrespectively of a worker's location. The knowledge diffusion technology is described by η^{hs}_{ver}, η^{col}_{ver}, η^{hs}_{hor}, η^{col}_{hor}. Vertical imitation, η^{hs}_{ver} and η^{col}_{ver}, is identified from the faster growth rate of wages—heterogeneous across cities—for workers who are in the bottom half of the initial wage distribution compared with those in the top half. Horizontal imitation, η^{hs}_{hor} and η^{col}_{hor}, is pinned down by differences across cities in wage growth for workers who start in the top half of the wage distribution. The parameter τ_M leverages η^e_{ver} and η^e_{hor} in determining the overall higher wage growth in large cities. Notice that the mapping between the learning parameters and wage growth is affected by the equilibrium human capital distribution in large and small cities. The better the quality of peers, the faster wage growth is, even under constant returns in meeting ($\tau_M = 0$).

Since young (old) workers age (retire) at rate ψ_y (ψ_o), imposing $\frac{1}{\psi_y} + \frac{1}{\psi_o} = 40 \times 12$ guarantees that workers retire after 40 years of work on average. The share of life spent as young, $\frac{\psi_o}{\psi_o + \psi_y}$, is also the time when human capital is accumulated and it is pinned down by the ratio between the average wage in the second 10 years and that in the first 10 years of labor market experience. Intuitively, for a given total wage growth, the ratio is higher the more time workers spend as young.

Local Labor Market. The job destruction rate, δ^e, is taken directly from the data. It is equal to the average probability of becoming unemployed in a given month across the entire sample period. I pin down the elasticity of the meeting rate in the local labor market with respect to city size, χ_M, by targeting the average *level* of the wage premium between large and small cities. The intuition is that a higher meeting rate makes firm-worker pairs more selective with

respect to the type of matches they are willing to form, which implies that the reservation match quality is higher in larger cities. Recall that workers' types need not be the same across cities. Hence, higher match quality is responsible only for the portion of the average wage premium that is not accounted for by differences in human capital composition.

One could conjecture the existence of lower search frictions in larger cities to be ruled out by the lack of systematic differences in unemployment and job-to-job transitions rates along the city-size distribution. In an economy with homogeneous workers and a single location, Martellini and Menzio (2020) prove that if and only if $z \sim Pareto(\underline{z}, \alpha)$, the presence of lower search frictions in a larger market, here $\chi_M > 0$, is consistent with observing the same labor market flows in markets of different size. Quantitatively, I find that the intuition in Martellini and Menzio (2020) carries through to the much richer environment of this paper. An alternative hypothesis that could also account for the observed labor market flows in large and small cities is the existence of constant returns to scale in the search process, $\chi_M = 0$, combined with an arbitrary match quality distribution. As I show below, if $\chi_M = 0$, the model understates the average wage premium in the economy and needs to resort to other mechanisms outside the ones proposed in this paper. More importantly, constant returns to scale would be inconsistent with direct estimates of χ_M I obtain from measuring the heterogeneity in the job search behavior of workers and firms along the city-size distribution (see discussion in Section 3.4).

Since χ determines the level of the firm-worker meeting rate, this parameter is identified by the average unemployment rate in the economy. The relative efficiency of searches on the job, ρ, is identified by the average number of jobs held by workers over the life cycle.

Match Quality Distribution and Bargaining Power. It is well known that if $z \sim Pareto(\underline{z}, \alpha)$, $\underline{z}$ cannot be identified separately from χ.[20] Hence, I normalize $\underline{z} = 1$. The slope of the distribution, α, and workers' bargaining power, β, are jointly identified by the average wage growth in a job-to-job transition and by returns to job tenure. As a measure of returns to tenure, I use the wage difference between equally experienced workers who spent more, compared with less, than a fifth of their working years at a given firm (very similar results are found using other cutoffs). The intuition is that the model can match a high average wage growth in job-to-job transitions by posing either a thick tail of the match quality distribution, that is, a small α, or a high value of workers' bargaining power, β. However, while in the former case returns

[20] Let $F(z) = 1 - \underline{z}/z^{\alpha}$. The rate at which an unemployed worker of type (h, a, e) accepts a job in city i is equal to $\chi M_i^{\chi_M}[1 - F(R(h, a, e, i))] = \chi M_i^{\chi_M} \underline{z}/R^{\alpha}(h, a, e, i)$. Clearly, labor market flows can identify only $\chi\underline{z}$.

to tenure would be large—since the dispersion in match quality would also be large—in the latter, returns to tenure would be small since workers would receive most of the job surplus upon transitioning to the new firm.

Location Choice. Workers in city i contact firms in city $-i$ at a rate that is ρ^* times as large as that of workers who live in city $-i$. Migrating entails a cost $c \sim Logistic(\mu_c, \sigma_c)$. The values of ρ^*, μ_c, and σ_c are jointly identified by relative city size, the hazard rate of moving from a small to a large city (or vice versa), and the share of workers in large cities with a college degree.[21] Since idiosyncratic migration motives—captured by σ_c—are symmetric across cities, high values of σ_c tend to equalize the equilibrium relative city size but also increase the probability that the realized migration cost is sufficiently low for a worker to choose to migrate. Both decreasing μ_c and increasing ρ^* positively affect the migration hazard rate. However, since migration entails a lump-sum cost and college graduates earn higher incomes, lower values of μ_c disproportionately facilitate the access of high school graduates to large cities.

Initial Human Capital Distribution. I normalize the mean of initial log-human capital for high school graduates to $\mu^{hs} = 0$, and choose μ^{col} to match the initial mean wage gap between college and high school graduates in the economy. The standard deviations of the initial human capital distributions, σ^{hs} and σ^{col}, are pinned down by the inter-quartile range of initial wages by education. The value of π^{col} is equal to the fraction of college graduates in the NLSY79 data. I use information on parental education for workers in the sample to compute the transition matrix $\pi^{e,e'}$ for newborns.

Housing Cost. The house price function in city i is described by a level, $p_{0,i}$, and an elasticity parameter, γ_i. Workers with education e consume q^e units of housing. The values of γ_i are obtained from the estimates in Saiz (2010).[22] I normalize $q^{hs} = 1$ and choose values of $p_{0,i}$ and q^{col} so that each educational group has the average expenditure share of non-tradeable goods computed by Moretti (2013).

Remaining Parameters. The monthly discount rate, r, is set to 1.25%, which corresponds to an annual value of 15%, as in Herkenhoff et al. (2018). This value is higher than what is

[21] For given relative city size and hazard rate of moving from a small to a large city, the hazard rate of moving in the opposite direction does not provide additional identifying information, since it is pinned down by the law of motion for city size, combined with the assumption of stationarity.

[22] Saiz (2010) reports the elasticity of housing supply for 269 MSAs in 2000. I assign to γ_{large} (γ_{small}) the population-weighted inverse elasticity for the 65 (204) metropolitan statistical areas (MSAs) with a population above (below) 750,000. Splitting the sample in Saiz (2010) according to the size of the commuting zones that mostly overlap with those MSAs—when such CZs can be clearly identified, as is the case for virtually all large cities—generates almost identical results.

usually adopted in traditional macroeconomic models. In contrast to standard concave utility functions, a linear utility has an infinite elasticity of substitution, which implies that it accommodates wide fluctuations in flow payoffs and might lead workers to temporarily accept negative wages in exchange for significant wage growth. Therefore, the value of r in this environment is meant to capture not only the rate of time preference but also, at least in part, the inverse of the workers' intertemporal elasticity of substition. The parameters b_i are chosen to match a flow payoff from unemployment equal to 60% of the average wage in city i.

Last, the (relative) number of small cities, N, is equal to 9.5, which is the ratio between the numbers of small and large cities in the 1990 US Census.

3.4 Calibration Results

3.4.1 Parameter Values

In this section, I discuss how the key calibrated parameters shed light on the mechanisms behind spatial wage differentials and workers' location choice. The full list of parameters is reported in Table 3.1.

The degree of returns to scale in the technology of knowledge diffusion, τ_M, is equal to 0.067. This value implies that at the equilibrium relative city size, $M_{large}/M_{small} = 15.7$, workers in large cities experience $15.7^{\tau_M} = 1.20$ times more interactions with one another than those in small cities.[23] A higher frequency of meetings in large cities leverages an equilibrium distribution of peers that first-order stochastically dominates the distribution in small cities. The relative importance of vertical and horizontal imitation is embodied in the parameters η^e_{ver} and η^e_{hor}. The calibration implies that the human capital of a college graduate at the 75th percentile of the human capital distribution in large cities grows at 2.8% per year, compared with 1.8% for a worker in the same position of the human capital distribution of a small city. The same comparison at the 25th percentile is 6% versus 4.5%. Notice that this is a conservative representation of the heterogeneity in learning opportunity across cities as for any quantile of the distribution the associated human capital value is higher in large cities.[24] A similar pattern applies to high school graduates, though the gaps are quantitatively smaller. College gradu-

[23] Alternative definitions of city size leave the results virtually unchanged. For example, defining size as working-age (male) population implies a ratio of $M_{large}/M_{small} = 16.4$ (16.5), which changes $15.7^{\tau_M} = 1.202$ to $16.5^{\tau_M} = 1.206$.

[24] For example, the human capital of a worker at the 25th percentile of the small city distribution would increase at 9.6% per year if he were in a large city.

ates also accumulate more human capital than high school graduates irrespective of their location, owing to a superior learning-by-doing technology ($\eta^{col} > \eta^{hs}$). Recall that learning is a prerogative of young workers. According to the calibrated value of $\psi_o/(\psi_y + \psi_o) = 0.46$, workers learn during their first $0.46 \times 40 = 18.4$ years of working life on average.

Turning to the characteristics of the search-and-matching technology, I find an elasticity of the contact rate between workers and firms with respect to city size, χ_M, equal to 0.2, which corresponds to increasing returns to scale ($\chi_M = 0$ under constant returns). The value of χ_M suggests that the frequency of meetings between workers and firms (both located in a large city) is $15.7^{\chi_M} = 1.73$ times higher than the meeting rate inside a small city. At the same time, the tail coefficient of the Pareto distribution of match quality is $\alpha = 3.6$ which implies that a match at the 75th percentile of the distribution is 36% more productive than a match at the 25th percentile. As explained in the identification section, lower search frictions leverage the dispersion in match quality and allow for the formation of more productive jobs. Not all the productivity advantages of creating better matches are passed on to the workers through higher wages as workers capture $\beta = 20\%$ of the gains from trade.

The degree of returns to scale in the search technology is calibrated by targeting the average wage premium between small and large cities at the equilibrium level of sorting on human capital. Therefore, it is of particular interest to compare the value of χ_M with direct evidence on the heterogeneity in search frictions across cities. Interpreting a firm-worker contact in the model as an application to a vacancy in the data, I compare the calibrated value of χ_M with two empirical measures of the same statistic. Martellini and Menzio (2020) report an estimate of the elasticity of the number of applications per vacancy received by *firms* with respect to the population of the commuting zone where the firm is located. The estimated elasticity, provided to us by Ioana Marinescu, is 0.52. On the other end of the labor market, the 1982 wave of the NLSY79 contains information on the most recent job search of all the *workers* who were employed when surveyed. Using information on workers' commuting zones and the number of employers they had contacted, divided by the number of weeks they had been looking for a job, I compute a (new) elasticity of 0.12. Notably, this model delivers a measure of the contribution of better match quality to the city-size wage premium while also being quantitatively consistent with both labor market flows in small and large cities, and direct estimates on worker and firm search behavior.

With regard to location choice, the calibrated value of ρ^* implies that workers contact firms

in a different city at a rate that is 8% of the within-city meeting rate. The average migration cost, μ_c, is equal to about two months of median income in the economy. In their seminal paper on interstate migration, Kennan and Walker (2011) estimate a much larger average cost equal to more than $300,000. Beyond comparing two different types of location choices, an additional difference between their paper and mine is that I replace a frictionless labor market with one characterized by search frictions. As pointed out by Schmutz and Sidibé (2019) and Heise and Porzio (2022), the presence of search frictions greatly reduces the size of the migration cost that is necessary to replicate the observed mobility rate. Furthermore, the model in this paper replicates, without targeting, the average wage gain of about 9% experienced by workers upon changing location.[25,26]

3.4.2 Model Fit

Table 3.2 reports the model-generated moments next to their empirical counterparts.

Through the lens of the model, the characteristics of the knowledge diffusion process generate variations in wage growth across locations, educational category, and relative position in the initial wage distribution. In line with the existence of vertical imitation, college graduates with low initial wages experience faster wage growth in large cities, 6.3% per year, than in small cities, which are at 4.7%. Learning from better peers also benefits college graduates that have high initial wages; they experience 4% wage growth per year compared to 3% in small cities. Even though high school graduates are characterized by a much flatter experience gradient, they display a similar pattern in terms of variations in wage growth across cities and initial wages (Figure D.1. in the Online Appendix).

The average wage growth in a job-to-job transition, which is key to identifying the tail of the match quality distribution, is closely matched at about 11%, and it is also very similar across cities. In addition, the model successfully replicates the declining experience profile of

[25] I compute the percentage wage change in the three months after migration compared with the three months before. The average value in the model is 9%, which is remarkably close to the empirical 9.1%. However, in the data, moving from a small to a large city is associated with a wage gain virtually identical to the one from moving in the opposite direction. In the model, wage gains are equal to 16% and 6%, respectively. An asymmetric mobility cost, or a systematic preference for living in large cities, would likely reduce this gap.

[26] The reason the estimates in Kennan and Walker (2011) are also consistent with the observed wage gain at migration is that, contrary to this paper, they allow for heterogeneity in location-specific preferences across people. They show that correcting for the utility gain from living in a preferred location generates actual moving costs paid by movers that are close to zero, if not slightly *negative*.

Table 3.1. Parameter Values

Internally Calibrated	Description	Value
σ_{hs}	initial hum cap hs, std	0.28
μ^{col}, σ^{col}	initial hum cap col, mean and std	0.3, 0.47
χ	meeting rate	0.07
χ_M	returns to scale matching	0.20
ρ	relative search eff. on the job	0.4
ρ^*	relative search eff. across cities	0.08
μ_q	migration cost (mean)	2 month wage
σ_q	migration cost (std)	1.6 month wage
α	shape match quality distrib	3.6
β	workers' bargaining power	0.2
τ_M	returns to scale knowledge	0.07
η_h^{hs}, η_h^{col}	horizontal imitation	0, 0.14
η_v^{hs}, η_v^{col}	vertical imitation	0.3, 0.7
η^{hs}, η^{col}, η	learning by doing	0.05, 0.08, 0.850
$\psi_o/(\psi_y+\psi_o)$	share of young	0.46
$p_{0,large}$, $p_{0,small}$	house price (intercept)	4.6e-5, 0.007
q^{col}	housing demand (col vs hs)	1.68
b_{large}, b_{small}	flow payoff unempl.	(1.27,1.13) $\times$ $\underline{z}$
Externally Calibrated	**Description**	**Value**
μ_{hs}	initial hum cap hs, mean	0
δ^{hs}, δ^{col}	job destruction	0.0135, 0.005
r	discount rate	0.0125
π^{col}	college share	0.2521
$\pi^{hs,col}$, $\pi^{col,col}$	intergen. transition rate to college	0.175, 0.62
γ_{large}, γ_{small}	(inverse) house supply elasticity	1.47^{-1}, 2.46^{-1}
N	number of small cities	9.5
$1/\psi_y + 1/\psi_o$	average life span	40 $\times$ 12 months

Table 3.2. Empirical and Model-Generated Moments

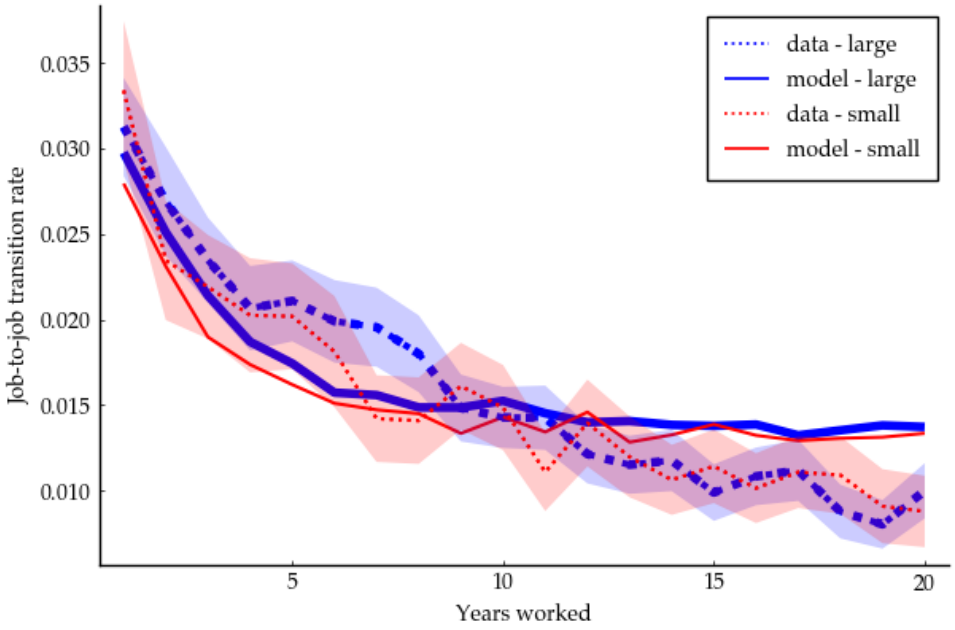
Moment (large/small)	Data	Model
Labor Market		
unemp. rate (%)	5.8/5.9	5.8/5.9
n. of jobs	6.9/7.2	7.1/7.2
mean wage gap (%)	29.3	29.2
EE wage growth (%)	10.7/11.8	10.7/11.2
return to tenure (%)	23.7	25.2
flow payoff unempl. (% of mean wage)	60/60	61.6/60
Cities		
city size	15.7	15.6
college share (%)	29/19	29/20
hazard rate small→large/large→small (%)	0.31/0.18	0.31/0.17
exp share non-trad. (col, %)	58/58	55/55
exp share non-trad. (hs, %)	58/58	58/57
Wage growth (%)		
bottom-col-large/small	6.3/4.7	6.7/4.6
top-col-large/small	4.0/3.1	4.0/3.0
bottom-hs-large/small	4.3/3.6	4.4/3.0
top-hs-large/small	2.6/1.3	2.7/1.8
wage 11–20 vs. 1–10 yr. of exp.	30.6	30.4
Initial Wage		
col wage premium (%)	38.7	37.3
$75^{th}/25^{th}$ pctile, hs	1.63	1.62
$75^{th}/25^{th}$ pctile, col	1.88	1.90

job-to-job transitions, or the employment-to-employment (EE) transition rate (Figure 1). After "climbing the job ladder" in early career stages, the probability of further improving on the current match declines and so does the frequency of job-to-job transitions. Both in the model and in the data, the EE rate is remarkably similar across cities of different sizes.[27] This find-

[27]Bleakley and Lin (2012) document that the probability of changing employers for young workers

Figure 1. Job-to-Job Transition Rate

Monthly job-to-job transition rate within large (thick blue line) and small (red thin line) cities, in the model (solid line) and in the data (dotted line).



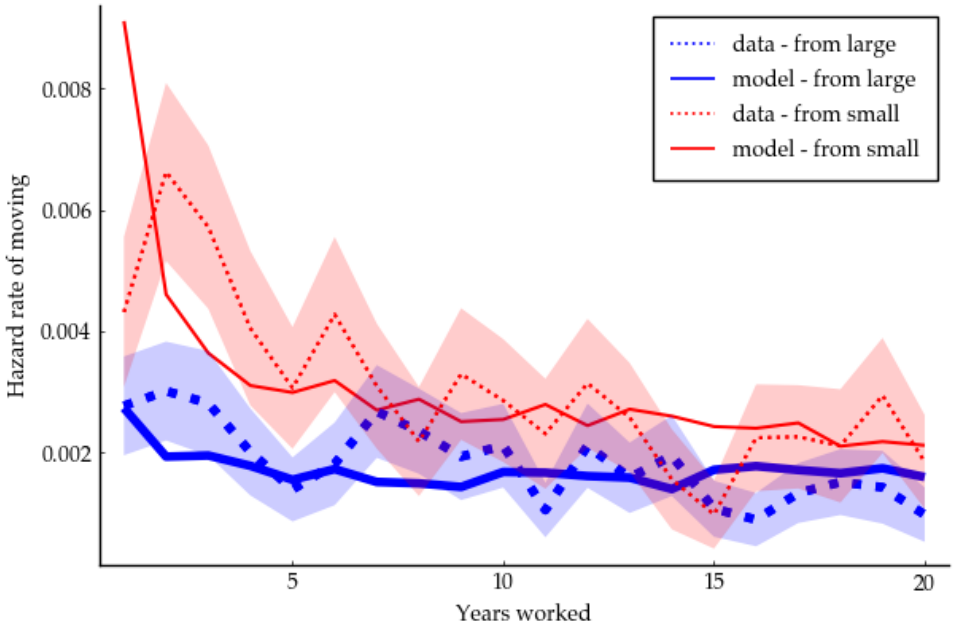
ing is entirely consistent with the existence of lower search frictions in large cities. On the one hand, employed workers in larger cities contact firms more frequently, contributing to a higher observed number of transitions. On the other hand, the endogenous distribution of employed workers over match qualities is also better in large cities, so that, conditional on a meeting, workers have a lower probability of improving on their current match. Under a Pareto match quality distribution, these two opposing forces cancel each other out, delivering a pattern of job-to-job transitions that is line with the data. A similar intuition can explain why the unemployment rate is virtually identical in large and small cities ($\approx 6\%$): the higher meeting rate between firms and workers is offset by the equilibrium increase in the reservation match quality in large cities.

Turning to the location choice, the average hazard rates of moving from a small to a large city and vice versa are equal to 0.31% and 0.18% per month, respectively. Figure 2 depicts life-

is somewhat declining in city density (they find a relationship of opposite sign when looking at occupational switching, which is their main focus). Their findings could differ from mine for two main reasons. First, they consider density instead of size. Second, they measure transitions at a much lower frequency—three years instead of a month—which is likely to combine actual EE transitions with unobserved EUE transitions and to be affected by location changes between observations.

Figure 2. Migration Hazard Rate

Hazard rate of moving out of a large (thick blue line) and a small (thin red line) city, in the model (solid line) and in the data (dotted line).



cycle profile of migration patterns. The model is consistent with the fact that migration flows into large cities are considerably larger than those in the opposite direction. Large cities offer two main advantages: a higher rate of human capital accumulation and lower search frictions in the labor market. Furthermore, the model generates a steady decline in migration rates over the life cycle: learning happens only at a "young" age, while older workers are likely to be employed in good matches and rarely start a new job, including in a different city.

Importantly for the study of the aggregate implications of local policy, the model closely matches the empirical relative city size. The large city in the model represents US commuting zones with a population of more than 750,000 people in 1990 and an average of 2.2 million, compared with 140,000 people for small cities. The model also replicates the spatial distribution of education. The share of college graduates in large cities is tightly matched at 29%, compared with 19% in small cities. College graduates benefit relatively more from the presence of better peers and, because of production complementarities between human capital and match quality, from lower search frictions.

3.5 Model Validation: Migration Episodes

In the previous section, I showed that the model is able to account for the labor market experience of workers in small and large cities. The calibration highlights the existence of increasing returns to scale in labor search and, to a lesser extent, in the rate of knowledge diffusion. Learning in larger cities is further fostered by the presence of better peers.

The calibration uses only information on aggregate wage differences between large and small cities and not specifically on movers. In fact, the identifying information is overwhelmingly provided by workers who never move from a small to a large city or vice versa which represent 78% of the NLSY79 sample.[28] Yet small and large cities are not isolated from each other, and some workers do in fact change locations over the life cycle. It is natural to ask whether the migration incentives generated by the matching and knowledge diffusion channels are consistent with the observed labor market experience of movers. Of particular interest, those incentives are shaped by the different portability of human capital, which is tied to a worker, and matching capital, which is associated with a job within a city. In this section, I test the model's ability to replicate the relative wage of movers with respect to both stayers in the pre-migration city and incumbents in the destination city.

I apply the methodology introduced in the context of job displacements by Jacobson, LaLonde, and Sullivan (1993) to the study of migration episodes between small and large cities. I estimate the following specification:

$$log(w_{it}) = \beta X_{it} + \sum_{k=-10^{+},..,-1,1,..,10^{+}} \delta_k d_{itk} + \varepsilon_{it}.^{29} \tag{3.3}$$

The dependent variable is the inflation-adjusted log-hourly wage of worker i in year t. The controls include a quadratic term in experience, an indicator for college education, and year dummies. The key coefficients of interest are given by the vector of δ_k associated with the dummy variables d_{itk}, which are equal to 1 if worker i in time t is in her k^{th} year after migrating to a city of different size. Negative values of k stand for years before migration takes place.[30]

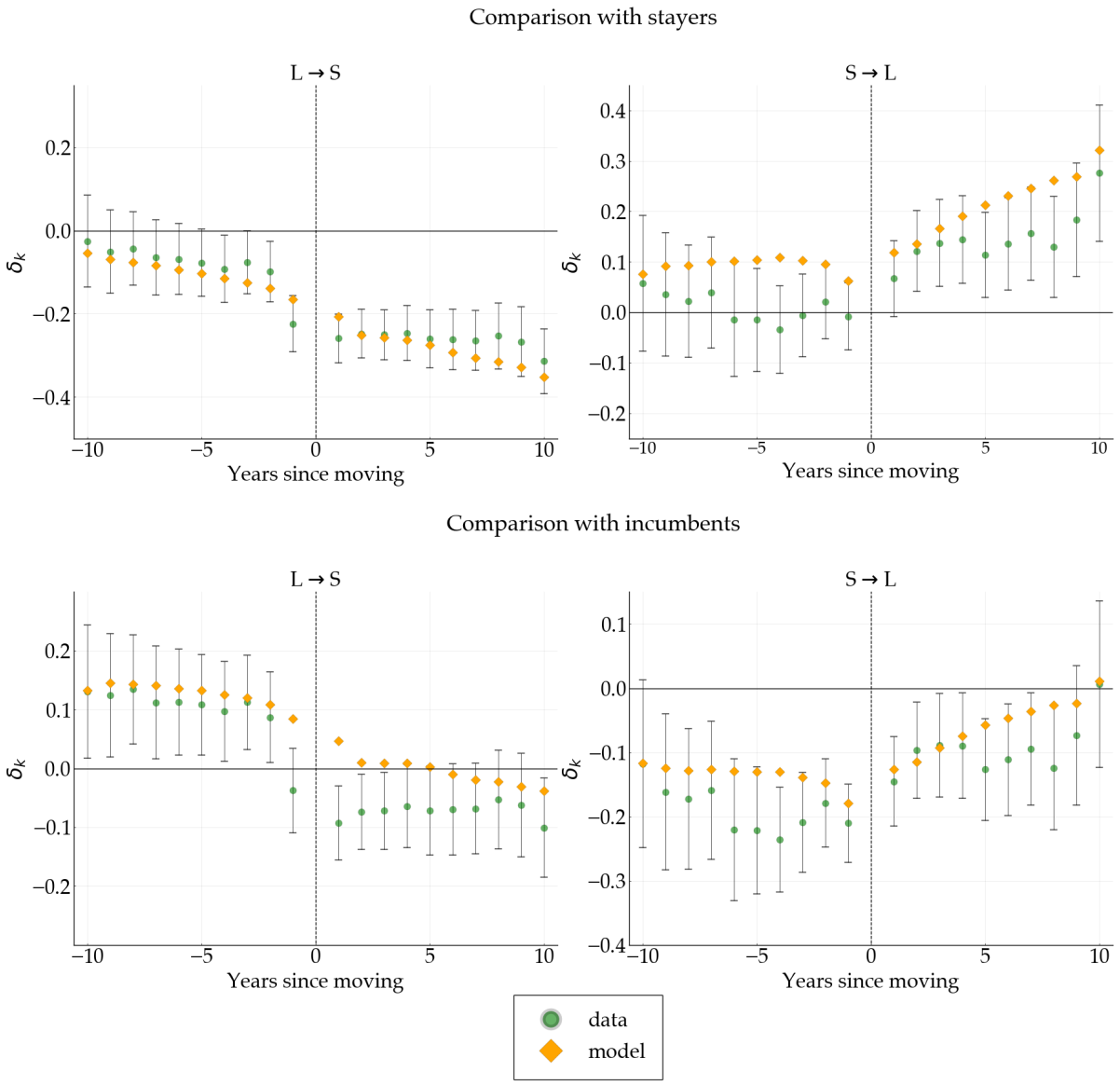
[28] To make this point precise, I re-compute the targeted moments on the subsample of workers who never move, both in the data and in the model. The empirical moments obtained from the set of non-movers are very close to their counterparts computed on the entire sample and are also remarkably similar to those generated by non-movers in the model (see Online Appendix D)

[29] Glaeser and Maré (2001) introduce a similar regression in the context of migration into and out of urban areas. One of the contributions of this paper is providing a micro-founded equilibrium theory that is empirically consistent with such reduced-form evidence.

[30] The value of $\delta_{10^{+}}$ ($\delta_{-10^{+}}$) stands for 10 or more years after (before) the migration episode.

Figure 3. Wage of Movers and Non-Movers

Wage of movers versus stayers (top panels) and versus incumbents (bottom panels). Left: from large city. Right: from small city. Data: green circles. Model: orange diamonds.



I run four different versions of regression 3.3, which are distinguished by the type of city a worker migrates from and by the selected control group, that is, the set of observations for which $d_{itk} = 0$. In Figure 3, I plot the coefficients generated from the NLSY79 sample (green circles) and those obtained using model-generated data (orange diamonds).

The top left panel shows the wage of workers who move from a large into a small city compared with the wage of those who remain in large cities. Leaving a large city is associated with lower pre-migration wages and a much larger, but substantially flat, wage drop in the years after moving to a small city. The top right panel repeats the same exercise with respect to the opposite migration flow. Migrants do not differ much from stayers until the moving year. After that, the wage of those who move to large cities starts to diverge, and it reaches a 28% premium after 10 years or more. Except for somewhat overestimating the extent of sorting out of small cities, the predictions from the model match the empirical evidence over the entire time window around migration episodes.

Recall that the calibration targets the aggregate wage difference between small and large cities. Since the model is able to replicate the relationship between the wages of movers and stayers, it should not be surprising that the model-based comparison between movers and incumbents is in line with the data as well (bottom panels of Figure 3). In the bottom left panel, I show that migrants to small cities earn somewhat less than incumbents, even if they were earning significantly more while working in large cities. This finding supports the existence of a productivity benefit from being in large cities that workers lose upon migrating out. From the bottom right panel, we observe how prospective migrants out of small cities earn less than those who already live in large cities and continue to do so even after migration takes place. This evidence is in line with the idea that most of the benefit from living in a large city does not accrue on impact. Consistently with the dynamic return to experience in large cities, the gap with respect to the incumbents slowly shrinks in the 10 years or more after migration.

In the model as in the data, there is a positive correlation between job tenure and wages. Recent migrants are relatively new at their job; hence, part of the discrepancy between their wage and non-movers' is simply due to shorter job tenure. To illustrate this point, I augment regression (3.3) with a quadratic term in the tenure at the current job and compute new values for the coefficients δ_k. The corresponding plots are reported in Online Appendix D. Both in the model and in the data, controlling for job tenure increases the wage gain at migration (or lowers the losses) but only in the first few years after moving.

3.6 Matching, Knowledge Diffusion, Sorting: Decomposition

In this section, I quantify the contribution to the city-size wage premium over the life cycle of matching, knowledge diffusion, and sorting. The decomposition proceeds as fol-

lows. I set the aggregate characteristics of cities at their baseline equilibrium values, treating them as *fixed parameters*, and I remove the possibility of migration between cities. I define $\Theta_i = \{\lambda_{0,i}, \lambda_{1,i}, \sigma_i, G_i\}$, for $i = \{small, large\}$, to be the set of aggregate characteristics of city i in the baseline economy, and I create four scenarios, s, each denoted by Θ_i^s.[31] Because of the assumption of no migration across cities, the policy functions in city i depend only on workers' idiosyncratic states (and match quality, if employed) and on Θ_i^s, but they do not depend on Θ_{-i}^s. That is, large and small cities are treated as separate economies. In all scenarios, I set $\Theta_{small}^s = \Theta_{small}$, and vary the characteristics of large cities, Θ_{big}^s, one by one. The thought experiment is comparing the life-cycle city-size wage premium in a world in which only one of the aggregate characteristics of large cities differs from its counterpart in small cities.

In the first scenario, I isolate the contribution of sorting. I solve the decision problem of workers and firms and simulate life-cycle wage paths under $\Theta_{big}^{NS} = \Theta_{big}$. I label this scenario *no-sorting, NS*, as the only departures from the baseline economy are imposing the same initial distribution of workers across cities—in terms of both human capital and education—and removing all migration flows between them. The left panel of Figure 4 shows the city-size wage premium in the data (black dotted line), the baseline economy (thick black line), and the no-sorting scenario (thin blue line). The vertical distance between the baseline and the no-sorting scenario captures the contribution of sorting through migration over the life cycle.

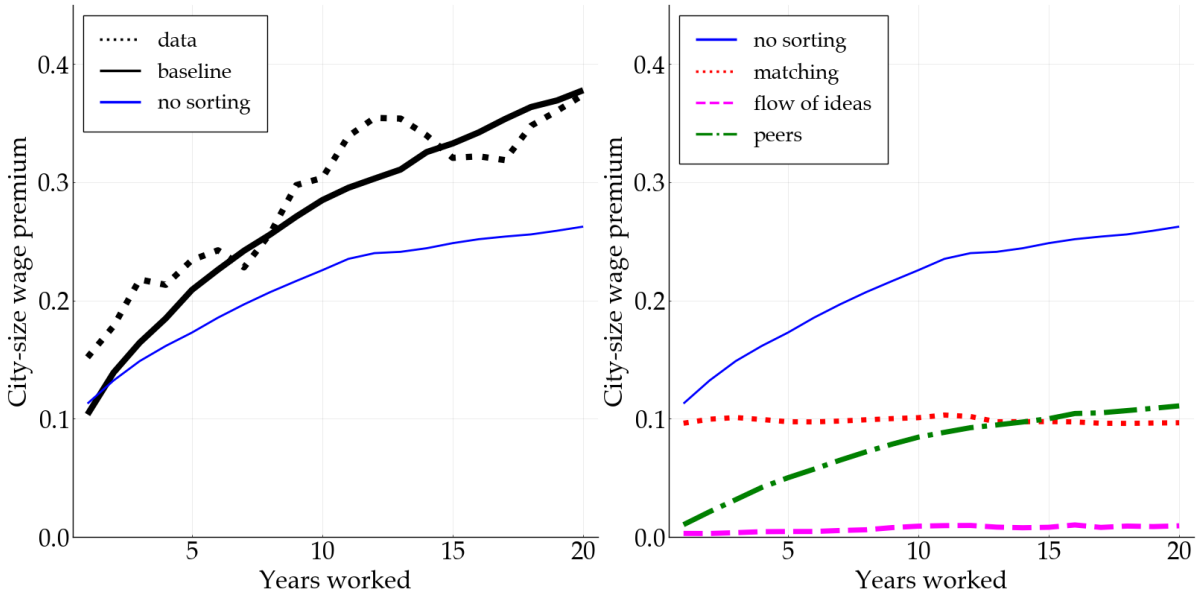
The remaining city-size wage premium is represented by the blue thin line reported on both panels of Figure 4, and it is due to heterogeneity in the equilibrium characteristics of small and large cities. The red dotted line on the right panel of Figure 4 corresponds to an economy in which only the *matching* channel is operative, $\Theta_{large}^M = \{\lambda_{0,large}, \lambda_{1,large}, \sigma_{small}, G_{small}\}$. In this scenario, large cities feature higher firm-worker meeting rates (on and off the job) but the rate and quality of interaction between workers are set equal to those in small cities. The pink dashed and green dashed-dotted lines repeat the same exercise with regard to the *flow of ideas* channel—which implies higher rate of interaction between workers in large cities, $\Theta_{large}^F = \{\lambda_{0,small}, \lambda_{1,small}, \sigma_{large}, G_{small}\}$—and with regard to *peer effects*, thanks to which the distribution from which workers learn in large cities dominates the one in small cities, $\Theta_{large}^P = \{\lambda_{0,small}, \lambda_{1,small}, \sigma_{small}, G_{large}\}$.[32]

[31] I omit house prices from the aggregate variables that characterize a city. If migration is ruled out, house prices are equivalent to a lump-sum tax paid by every worker in the economy—although potentially different across cities—but they have no impact on wages.

[32] In the flow of ideas and peer effect scenarios, I also set the value of b in large cities equal to its value in small cities. Recall that the flow payoff from unemployment in city i is $b_i h$. In the model, wages are

Figure 4. Decomposition of the City-Size Wage Premium

Percentage wage premium. Left panel: data (black dotted line), baseline model (black thick line), 'no-sorting' scenario (blue thin line). Right panel: 'no-sorting' (blue thin line), 'matching' (red dotted line), 'flow of ideas' (pink dashed line), 'peer effect' (green dashed-dotted line) scenarios.



Starting from the left panel of Figure 4, we observe how the contribution of sorting slowly rises with experience, and it reaches 12 pp after 20 years. I further isolate the contribution of sorting on observable education, up to 7 pp, from sorting on unobservable human capital, up to 5 pp. To perform this additional decomposition, not shown in the graphs, I re-simulate the *no sorting* scenario, weighting the observations in each city and at each experience level according to the experience-specific educational composition of cities in the baseline economy. The modest role of sorting on human capital is consistent with the findings by, among others, Eeckhout, Pinheiro, and Schmidheiny (2014) and De La Roca and Puga (2017). While in the static environment of Eeckhout, Pinheiro, and Schmidheiny (2014) lack of sorting on average is due to production complementarity between workers at the top and the bottom of the skill distribution, in this paper it is the result of two opposing forces. On the one hand, high-skill workers benefit more from locating in large cities because of complementarity between

determined by the human capital of the worker and the match quality of the job. Given that the value of b_i aims to replicate a flow payoff of unemployment that is a constant fraction of wages, it is natural to set it equal across cities in those scenarios in which the only systematic difference between workers in small and large cities is in terms of human capital.

human capital and match quality. On the other hand, low-skill workers have more scope for learning in large cities, thanks to vertical imitation in knowledge diffusion.

The right panel of Figure 4 shows that matching has a *level* effect on the wage premium of about 10 pp, but it plays no role in the growth of the premium over the life cycle. Once again, if the distribution of match qualities is Pareto, increasing returns to scale in the search process generate better matches at any level of experience. At the same time, a Pareto distribution is the only one that is consistent with increasing returns to scale, even in the absence of size-dependent differences in unemployment and job-to-job transition rates. The flow of ideas and peer effects are negligible for newborns, as human capital accumulation occurs over time, but they are responsible for the entire growth of the wage premium that is not accounted for by sorting, up to 15 pp, with peer effects being by far the predominant force.[33]

The decomposition highlights one of the key contributions of this paper: measuring the role of (endogenous) learning and matching opportunities in accounting for cross-city differences in wage level and growth while being consistent with the empirical evidence on labor market flows in small and large cities.

It is worth noticing that besides isolating the role of sorting, the absence of migration flows in all scenarios is necessary to correctly measure the contribution of any specific treatment (matching, flow of ideas, and peer effects) at its baseline equilibrium value. To see why, consider the matching scenario. If workers were allowed to migrate, the resulting wage premium would necessarily represent a combination of both the actual object of interest—higher average match quality in large cities—and the sorting behavior induced by matching itself.

4 Closing the Model: Endogenous Market Tightness

The stationary equilibrium of the model includes an exogenous specification for the rate at which a worker contacts a firm in the local labor market and across cities, λ and λ^*, respectively. In evaluating the aggregate implications of local policies, I allow firms to adapt their behavior to changes in the profitability of jobs in different cities by introducing vacancy creation in the tradition of Mortensen and Pissarides (1994).

I assume that the number m_i of meetings between a firm and an unemployed worker, both located in city i, is given by the product between $M_i^{\chi_M}$ and a constant returns to scale function

[33] Recall that flow output is equal to $h_\ell z$. A residual wage premium of about 1 pp is due to the interaction between matching and knowledge diffusion.

m,

$$m_i = M_i^{\chi_M} m(s_i, v_i) = M_i^{\chi_M} s_i^{1-\zeta} v_i^{\zeta},$$

where v_i is the number of vacancies, and s_i is the actual measure of workers seeking jobs in city i. The meeting rate between a firm and an employed worker when both are located in city i is equal to ρm_i, while it is equal to $\rho^* m_i$ ($\rho\rho^* m_i$) if the unemployed (employed) worker is in city $-i$ at the moment of the contact with a firm in city i. It follows that

$$s_i = [u_i + \rho(1 - u_i)]M_i + \rho^*[(u_{-i} + \rho(1 - u_{-i}))]M_{-i}, \tag{4.1}$$

where u_i is the unemployment rate in city i. Since search is random and m has constant returns to scale, an unemployed worker in city i contacts a firm, also in city i, at rate

$$\frac{m_i}{s_i} = M_i^{\chi_M} \theta^{\zeta}, \tag{4.2}$$

where $\theta = v_i/s_i$ is assumed to be the same across cities. From the specification introduced in Section 3, the contact rate (4.2) is also equal to $\lambda_{0,i} = M_i^{\chi_M}\chi$. Petrongolo and Pissarides (2001) document a wide range of empirical estimates of ζ and suggest a value of 0.5, which implies $\theta = \chi^{\frac{1}{\zeta}}$. The unit cost of creating a vacancy, k_i, is pinned down by the zero profit condition in the market for vacancies

$$k_i = M_i^{\chi_M} \theta^{\zeta-1}(1 - \beta)\mathcal{S}(i), \tag{4.3}$$

where $\mathcal{S}(i)$ is the expected job surplus from contacting a worker and $(1 - \beta)$ is the share that accrues to the firm. A formal description of $\mathcal{S}(i)$ follows directly from equation (2.5) and it is reported in Online Appendix A.

Applying Equation (4.3) to both large and small cities, I find $k_{large}/k_{small} = 1.6$. Because of increasing returns to scale in the search process and a better human capital composition, firms in large cities are willing to pay a 60% higher vacancy creation cost. A higher value of k reduces firms' incentives to create vacancies in a given city in a similar fashion to how higher house prices represent a congestion force in workers' location choice.

5 Local Housing Policy and Aggregate Productivity

Extensive research has investigated the sources of the dispersion in productivity across US cities. Far less explored is how the nature of such spatial heterogeneity is central to evaluating the aggregate implications of local policies that trigger the relocation of workers across cities

and, in particular, toward more-productive ones. I illustrate this point by focusing on a type of policy that has been at the forefront of the academic and policy debate in recent years: relaxing housing regulation in large US cities. In Section 5.1, I describe the policy experiment and show how this paper is related to the existing literature on this topic. In Section 5.2, I propose an identity that allows me to decompose the change in total income after the policy implementation into pre-policy observables and equilibrium responses of local productivity in small and large cities. Section 5.3 presents the results.

5.1 Housing Regulation: Background

City-level (inverse) elasticities of housing supply, γ_i, have been estimated by Saiz (2010). Averaged across large and small cities, they are equal to 1/1.47 and 1/2.46, respectively. A growing literature has documented that part of the higher inverse elasticity in some large US cities—the fact that house prices grow more rapidly as the city expands—is not due to physical constraints but is the outcome of tighter land use regulation.[34]

Hsieh and Moretti (2019), henceforth HM, explore the aggregate implications of strict housing regulation in the most productive large US MSAs. They find that setting housing supply regulation in New York, San Francisco, and San Jose to the level of the median US city would generate an increase in GDP of about 3.7% under imperfect worker mobility (8.9% with perfect mobility). In their model, housing regulation leads to higher house prices, which prevent productive cities from expanding and generating more output. Herkenhoff, Ohanian, and Prescott (2018), henceforth HOP, develop a neoclassical growth model with multiple regions. They find that GDP would be 12% higher if land use regulation in all US states moved halfway toward that of the least restrictive state, Texas, or by 3.7% if regulation in California and New York returned to their 1980 levels. HM and HOP model spatial heterogeneity in productivity as an exogenous differences in total factor productivity (TFP) between locations.[35] In addition, they impose homogeneity between workers and abstract from life-cycle

[34]Saiz (2010) estimates the contribution of regulation to the housing supply elasticity of 269 US cities. Glaeser and Gyourko (2018) compute the magnitude of house prices in excess of construction costs in US MSAs and find that it is negatively correlated with the number of building permits issued in the same MSA.

[35]Both papers consider an extension of their model in which TFP is an exogenous, increasing, constant elasticity function of local population (HM) or output (HOP), but they do not take a stand on the foundations of such agglomeration force. Furthermore, they find that the type of agglomeration force they consider further amplifies the income gains from the relaxation of housing policy.

considerations.

This paper introduces three new margins to the debate. First, exogenous heterogeneity in productivity is replaced by agglomeration forces in the form of increasing returns to scale in the labor market and in the process of knowledge diffusion. Second, in line with the increase in the city-size wage premium over the life cycle, part of the productivity differential between large and small cities is dynamic and stems from differences in the rate of human capital accumulation. Third, heterogeneity in human capital between workers to contribute to the process of knowledge diffusion in cities through spatial sorting.

All these aspects—increasing returns, knowledge diffusion, sorting—are endogenous to the policy environment and account for key features of labor market outcomes across cities that are missing from the existing literature on housing and other place-based policies.[36]

5.2 Interpreting the Policy Outcome

I introduce an identity that highlights the role of endogenous local productivity in shaping the equilibrium response to a policy change. Throughout the analysis, I focus on the steady-state total income in the economy before and after the implementation of the policy, assuming a working life of 40 years for all workers.

I denote by j^0 (j^1) the value of any variable j before (after) the policy implementation and by $\bar{j}$ the mean of j. Let $\Delta M = M_\ell^1 - M_\ell^0 = N(M_s^0 - M_s^1)$ be the equilibrium change in the measure of workers located in large cities. The difference between total income, denoted by Y, after, compared with before, the change in policy is equal to

$$Y^1 - Y^0 = \underbrace{\Delta M(\bar{y}_\ell^0 - \bar{y}_s^0)}_{\text{(D)}} + \underbrace{M(\bar{y}_s^1 - \bar{y}_s^0)}_{\text{(S)}} + \underbrace{M_\ell^1[(\bar{y}_\ell^1 - \bar{y}_s^1) - (\bar{y}_\ell^0 - \bar{y}_s^0)]}_{\text{(P)}}, \tag{5.1}$$

where y_i represents income in city i. To interpret the terms identity (5.1), we can think of the income produced by a worker as being the sum of two components: the income she would produce in a small city and, for a worker located in a large city, the additional income difference between large and small cities.

The change in total income can be decomposed into the sum of three terms. First, the

[36]This paper exhibits a more stylized geography than the traditional urban literature. HM model an economy with more than 200 MSAs, while HOP consider seven groups of states. Combining the mechanisms I analyze with a richer geography would allow the model to account for additional heterogeneity in productivity and house prices across cities (or between states), but would present significant computational challenges and data requirements and it is left for future work.

direct effect (D) is the change in total income given by the average productivity gap currently observed multiplied by the additional measure of workers located in large cities after the policy implementation. It is easy to verify that if average income in small and large cities is policy invariant, that is, $\bar{y}_\ell^1 = \bar{y}_\ell^0$ and $\bar{y}_s^1 = \bar{y}_s^0$, identity (5.1) simplifies to

$$Y^1 - Y^0 = \Delta M(\bar{y}_\ell^0 - \bar{y}_s^0). \tag{5.2}$$

If such invariance is not satisfied, the total income change includes two additional terms. First, $M(\bar{y}_s^1 - \bar{y}_s^0)$ is the change in the average income of small cities multiplied by the entire population in the economy, (S). Second, $M_{large}^1[(\bar{y}_\ell^1 - \bar{y}_s^1) - (\bar{y}_\ell^0 - \bar{y}_s^0)]$ is the change in the city-size premium, multiplied by the size of large cities after the policy, (P).

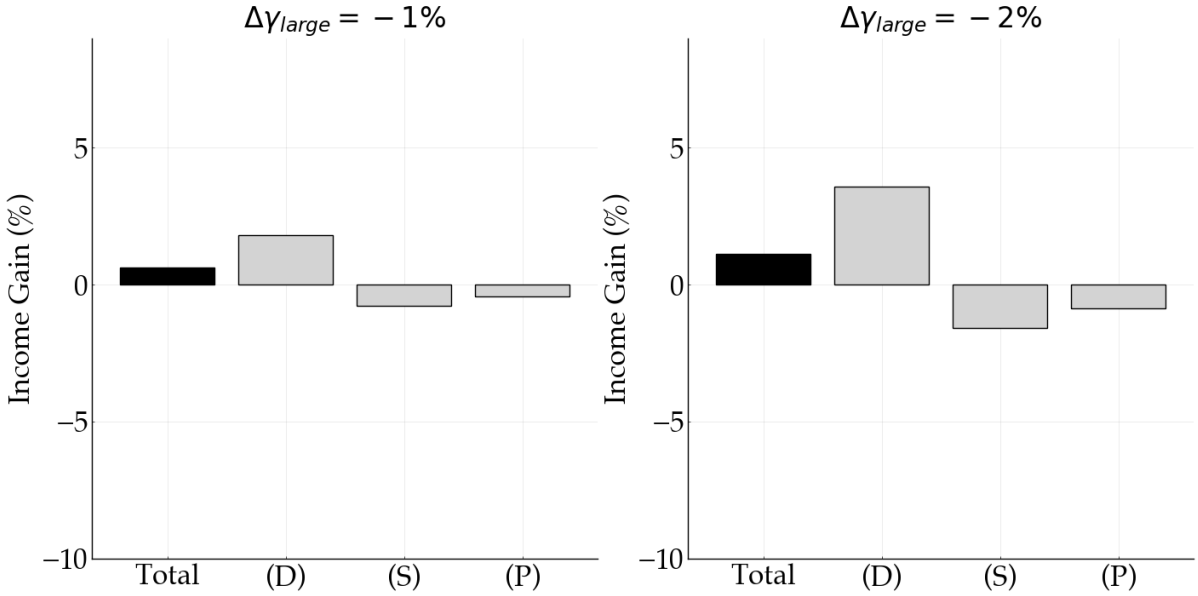
5.3 Results

I consider the steady-state equilibrium response of the economy to a 1% and a 2% increase in the elasticity of housing supply in large cities, while I leave the same parameter in small cities unchanged. Specifically, I replace $\gamma_{large} = 1/1.474$ with $\gamma_{large} = \{1/1.489, 1/1.504\}$. To put the magnitude of this policy in context, the average increase in the elasticity of housing supply in all large cities obtained from setting their zoning regulation equal to the average across all US cities would be equal to 8%.[37] Under looser housing regulation, large cities gain in population: they become 8.8% and 17.3% larger according to the size of the policy change—well within the range considered in existing work. The increase in size is accompanied by a change in their educational and human capital composition. The share of workers in large cities who hold a college degree drops from 29% in the baseline economy to 28% and 27% as γ_{large} shrinks. The equilibrium human capital distributions deteriorate in *both* large and small cities. This finding is in line with the fact that workers who live in large cities under higher housing supply are more skilled than those who remain in small cities, but less skilled than those who live in large cities in the baseline economy. Such change in sorting further affects human capital accumulation through its effect on the knowledge diffusion process within cities. Next, I explore the aggregate implications of the change in city size and composition, as measured by variations in total income.

[37] In their leading counterfactual experiment, HM relax zoning regulation only in New York, San Francisco, and San Jose. They also obtain an increase in housing supply elasticity in those three cities of about 8% (both they and I follow the procedure in Column (2) of Table V in Saiz 2010). As the policy I consider is applied more broadly to all large cities, I implement a smaller change in regulation.

Figure 5. Housing Regulation and Aggregate Income Gains

Percentage change in total income (black bar) and in its components (grey bars). Left panel: $\Delta\gamma_{big} = -1\%$. Right panel: $\Delta\gamma_{big} = -2\%$.



The outcome of the housing policy can be analyzed by using identity (5.3),

$$\underbrace{\frac{Y^1 - Y^0}{Y^0}}_{\text{Total}} = \underbrace{\frac{\Delta M(\bar{y}_\ell^0 - \bar{y}_s^0)}{Y^0}}_{\text{(D)}} + \underbrace{\frac{M(\bar{y}_s^1 - \bar{y}_s^0)}{Y^0}}_{\text{(S)}} + \underbrace{\frac{M_\ell^1[(\bar{y}_\ell^1 - \bar{y}_s^1) - (\bar{y}_\ell^0 - \bar{y}_s^0)]}{Y^0}}_{\text{(P)}}, \tag{5.3}$$

which includes the same terms as (5.1), each divided by Y^0 to obtain percentage deviations from the baseline. Figure 5 shows the overall effect (black bars) and its components (grey bars). The direct effect is equal to +1.8% (+3.6%), while the total effect is +0.6% (+1.1%), when γ_{large} is lowered by 1% (2%). The direct effect (D) is clearly positive: since large cities are more productive, their expansion increases total income. The size of the direct effect is quantitatively in line with the previous literature that assumes exogenous productivity differences between cities. However, because of endogenous changes in the productivity of cities, the other two terms create a wedge between the direct and the total effect of the policy. As previously mentioned, the composition of small cities deteriorates compared with the baseline economy, and so do peer effects in those cities. In addition, increasing returns to scale have an adverse effect on the average match quality and on the frequency of interactions in a city that shrinks in size. All these forces contribute to the negative sign of the second term, (S). Because of weaker sorting and worse peers, income in large cities *declines* as well. The con-

tribution of agglomeration forces (i.e., matching and flow of ideas) is positive in a city that expands, but not enough to offset the human capital loss in large cities: the income premium (P) between large and small cities is slightly lower in the after-policy—compared with the pre-policy—equilibrium. All in all, even accounting for the equilibrium response of agglomeration forces, peer effects, and sorting, relaxing housing restrictions in large cities does generate non-negligible income gains. Yet, assuming that local productivity was policy-invariant would overstate the percentage increase in total income by a factor of 3.

6 Conclusions

The geography of economic activity in the United States is characterized by significant and persistent productivity differentials between cities, particularly between large and small ones. In this paper, I build a model in which this productivity gap emerges as an equilibrium outcome due to spatial sorting, increasing returns to scale in the labor search process, and heterogeneity in the (endogenous) composition of peers. The model replicates the observed labor market experience of workers in large and small cities, and it is tested with respect to its predictions for the experience of movers in the years before and after a move.

In the presence of spatial dispersion in productivity, a recent debate has drawn attention to the aggregate implications of local policies that restrict workers' access to large and productive cities. I perform a policy experiment in which housing regulation in large cities is relaxed and show that micro-founding the nature of local productivity is key to evaluating the policy outcome. The existing literature on the aggregate effects of local policies postulates static and exogenous gains from working in large cities, hence it is silent on the dynamic nature of the city-size wage premium and how such premium is determined in equilibrium. I find that increasing housing supply in large cities would further expand their size, but it would also change their composition and the productivity benefit from living in those cities, mainly through a weakening of sorting and peer effects. Assuming that local productivity was invariant to the policy would overstate the actual income gain by a factor of 3.

This paper leaves some important questions open for future research. First, I abstract from firm heterogeneity and firm sorting across locations (Lindenlaub, Oh, and Peters 2022). Through the lens of this paper, a high-quality firm is one with better workers, and consequently, better peers. The distinction between worker composition and exogenous firm types

is salient insofar as firms' location decisions might feature a different elasticity to policy changes compared with their workers' decisions. Second, I model human capital as being one-dimensional, with workers always benefiting from interacting with one another and increasing their stock of knowledge. In reality, cities are heterogeneously specialized in different occupations, which presumably require a certain degree of occupation-specific human capital. Capturing this richer heterogeneity while preserving the dynamic structure of the model is a challenging task, but it could provide insights into the differential impact of technological innovations and policy changes across cities and sectors of the economy.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2022. "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach". *Review of Economic Studies* 89 (1): 45–86.

Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stefanie Stantcheva, and Valerio Sterzi. 2018. "Dancing with the Stars: Innovation through Interactions". NBER Working Paper No. 24466.

Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. "Tenure, Experience, Human Capital, and Wages: A Tractable Equilibrium Search Model of Wage Dynamics". *American Economic Review* 104 (6): 1551–1596.

Baum-Snow, Nathaniel, and Ronni Pavan. 2012. "Understanding the City Size Wage Gap". *Review of Economic Studies* 79 (1): 88–127.

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.

Buera, Francisco J., and Ezra Oberfield. 2020. "The Global Diffusion of Ideas". *Econometrica* 88 (1): 83–114.

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

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

Costa, Dora L., and Matthew E. Kahn. 2000. "Power Couples: Changes in the Locational Choice of the College Educated, 1940–1990". *The Quarterly Journal of Economics* 115 (4): 1287–1315.

Davis, Donald R, and Jonathan I Dingel. 2019. "A Spatial Knowledge Economy". *American Economic Review* 109 (1): 153–170.

De La Roca, Jorge, Gianmarco I P Ottaviano, and Diego Puga. 2014. "City of Dreams". Working Paper.

De La Roca, Jorge, and Diego Puga. 2017. "Learning by Working in Big Cities". *Review of Economic Studies* 84 (1): 106–142.

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, edited by J. V. Henderson and J. F. Thisse* 4:2063–2117.

Eckert, Fabian, Mads Hejlesen, and Conor Walsh. 2022. "The Return to Big City Experience". *Journal of Urban Economics* 130.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny. 2014. "Spatial Sorting". *Journal of Political Economy* 122 (3): 554–620.

Glaeser, Edward L. 1999. "Learning in Cities". *Journal of Urban Economics* 46 (2): 254–277.

Glaeser, Edward L., and Joseph Gyourko. 2018. "The Economic Implications of Housing Supply". *Journal of Economic Perspectives* 32 (1): 3–30.

Glaeser, Edward L., and David Maré. 2001. "Cities and Skills". *Journal of Labor Economics* 19 (2): 316–342.

Heise, Sebastian, and Tommaso Porzio. 2022. "Labor Misallocation Across Firms and Regions". NBER Working Paper 30298.

Herkenhoff, Kyle F., Jeremy Lise, Guido Menzio, and Gordon M. Phillips. 2018. "Production and Learning in Teams". NBER Working Paper No. 25179.

Herkenhoff, Kyle F., Lee E. Ohanian, and Edward C Prescott. 2018. "Tarnishing the Golden and Empire States: Land-Use Restrictions and the U.S. Economic Slowdown". *Journal of Monetary Economics* 93:89–109.

Hsieh, Chang-Tai, and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation". *American Economics Journal: Macroeconomics* 11 (2): 1–39.

Huggett, Mark, Gustavo Ventura, and Amir Yaron. 2006. "Human Capital and Earnings Distribution Dynamics". *Journal of Monetary Economics* 53 (2): 265–290.

Jacobson, Louis S., Robert J. LaLonde, and Daniel G. Sullivan. 1993. "Earnings Losses of Displaced Workers". *American Economic Review* 83 (4): 685–709.

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): 577–598.

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg. 2021. "Learning from coworkers". *Econometrica* 89 (2): 647–676.

Kennan, John, and James R. Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions". *Econometrica* 79 (1): 211–251.

Lindenlaub, Ilse, Ryungha Oh, and Michael Peters. 2022. "Firm Sorting and Spatial Inequality". NBER Working Paper 30637.

Lise, Jeremy, and Fabien Postel-Vinay. 2020. "Multidimensional Skills, Sorting, and Human Capital Accumulation". *American Economic Review* 110 (8): 2328–2376.

Lucas, Robert E. Jr. 2004. "Life Earnings and Rural-Urban Migration". *Journal of Political Economy* 112 (S1): S29–S59.

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

Martellini, Paolo, and Guido Menzio. 2020. "Declining Search Frictions, Unemployment and Growth". *Journal of Political Economy* 128 (12): 4387–4437.

Menzio, Guido, and Shouyong Shi. 2011. "Efficient Search on the Job and the Business Cycle". *Journal of Political Economy* 119 (3): 468–510.

Menzio, Guido, Irina A Telyukova, and Ludo Visschers. 2016. "Directed Search Over the Life Cycle". *Review of Economic Dynamics* 19:38–62.

Moretti, Enrico. 2013. "Real Wage Inequality". *American Economic Journal: Applied Economics* 5 (1): 65–103.

Mortensen, Dale T, and Christopher A Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment". *Review of Economic Studies* 61 (3): 397–415.

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

Perla, Jesse, Christopher Tonetti, and Michael Waugh. 2021. "Equilibrium Technology Diffusion, Trade, and Growth". *American Economic Review* 111 (1): 73–128.

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.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Felipe Schwartzman. 2019. "Cognitive Hubs and Spatial Redistribution". NBER Working Paper 26267.

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

Schmutz, Benoît, and Modibo Sidibé. 2019. "Frictional Labour Mobility". *Review of Economic Studies* 86 (4): 1779–1826.

Tolber, Charles M., and Molly Sizer. 1996. "U.S. Commuting Zones and Labor Market Areas: A 1990 Update". *Economic Research Service Staff Paper 9614*.