On the Geography of Inequality: Labor Sorting and Place-Based Policies in General Equilibrium *

Santiago Truffa[†] Alexis Montecinos [‡] Diogo Duarte [§]

October 14, 2019

Abstract

We study how cities' amenities and restrictions on the housing supply contribute to aggregate wage inequality and affect housing prices through the sorting of heterogeneously skilled workers. We develop a general equilibrium model where workers differ along a continuum of skills and compete for limited housing. Our analysis suggests that spatial sorting accounts for 7.5% of the aggregate wage dispersion, increases average housing prices by 20–40% in constrained cities, and makes the economy 1.9% more productive. In addition, we evaluate a place-based policy that aims to expand the supply of houses in 1% in constrained cities and find that it improves aggregate productivity between 0.2% and 0.4%. However, the place-based policy has an unintended consequence of aggravating aggregate wage inequality by the same magnitude.

JEL classification: D58, D44, F16, J24, R13
keywords: Sorting, Inequality, Housing

*We are extremely grateful to Ernesto Dal Bó, William Fuchs and John Morgan for their continuous guidance and support. We would also like to thank Victor Couture, Cecile Gaubert, Rui de Figueiredo, William Grieser, Enrico Moretti, Gonzalo Maturana, Steve Tadelis, Joachim Voth, Reed Walker and Noam Yuchtman as well as numerous seminar and conference participants for their helpful discussions and comments.

[†]ESE Business School, Universidad de los Andes Chile. E-mail: struffa.ese@uandes.cl

[‡]Sawyer Business School, Suffolk University. E-mail: amontecinosbravo@suffolk.edu

[§]College of Business, Florida International University. E-mail: diogo.duarte@fiu.edu

1. Introduction

Rising wage inequality has been a defining feature of the US economy over the last few decades. This increase in inequality has been accompanied by changes in the organization of the economic activity of cities. This reorganization has generated severe urban differentiation: cities that concentrate a higher fraction of high-skilled workers also feature higher wages and housing prices. Changes between cities have gone hand-in-hand with changes within cities, as more productive places have also become more unequal.

While prior studies have shown the importance of physical characteristics of cities to the spatial variation in wages and housing prices, there is still disagreement about the key drivers of those differentials. For instance, Roback (1982) and Gyourko and Tracy (1991) attribute differences in housing prices across cities mainly to variation in interurban amenities and other local traits, such as intercity fiscal differentials. By contrast, other studies argue that those price differentials can mainly be explained by limited housing supply in large metropolitan areas, such as New York, Los Angeles, and San Francisco. These so-called *superstar cities* attract high-income families, shifting the income distribution and rising the price of land. According to Gyourko et al. (2013), metropolitan statistical areas (MSAs) with low land availability tend to be more productive due to the sorting of high-skilled workers into these places. These same cities have been among the least likely to add new housing in the last couple of decades, contributing to the exacerbation of regional divergences.

Our paper sheds light on this discussion by developing a spatial general equilibrium model that features both of the two key drivers of differentials contemplated by the literature: amenities and limited housing supply. Our model links these city characteristics to the sorting of workers with a continuum of skills to evaluate how city fundamentals contribute to wage inequality, housing prices, and output growth through the spatial distribution of the population.

To the best of our knowledge, this is the first spatial equilibrium model to jointly analyze the impact of amenities and limited housing supply to wage dispersion, house prices, and growth.

The main findings are as follows. First, if workers account for both cities' amenities and limited housing supply when deciding to reallocate, spatial sorting makes the aggregate economy 1.9% more productive, with places featuring a tighter housing market, such as superstar cities, experiencing an increase in productivity of 30% on average. The rationale behind this result is straightforward: when we allow for spatial sorting, most productive workers tend to cluster in large cities that require more talent for technology-intensive tasks. As a result of agglomeration externalities, workers' marginal productivity increases, which contributes to higher overall productivity.

Second, the influx of high-skill workers to large cities with limited housing supply put further pressure on the demand for housing and ultimately raises their price. Our model estimates an average increase of 20–40% because of talent sorting on these metro areas. We show that this effect is particularly detrimental to house prices in smaller cities, where housing supply is not constrained and can experience a decline by up to 30% due to the drain of talented workers.

Third, we show that spatial sorting has critical consequences to wage inequality when cities' amenities and shortage of housing impact workers' mobility decisions. We find that spatial sorting accounts for 7.5% of the aggregate wage dispersion and wage inequality is between 20% to 40% higher in those cities with a tighter housing markets. The reasoning goes as follows. Workers care, by assumption, about their disposable income (i.e., wages net of housing costs) and amenities provided by the city they live in. Since high-skilled workers have higher disposable, they become more sensitive to the cities' amenities. As they cluster in MSAs due to the high demand for high-skill tasks, they outbid low-skilled in the housing market and drive them out of the metropolitan area, which raises the average salary within the city. As a result, wage dispersion increases. In addition, in equilibrium, those low-skilled workers who stay in supercities are compensated with higher salaries as well since the supply of those workers diminishes significantly. Our model shows that this mechanism can generate a differential 30–50% in the salaries of the workers in the 20th percentile of the skill distribution who live in large MSAs and those in the same group living in smaller cities.

The fourth contribution of our study is to analyze how a place-based policy could potentially ameliorate or aggravate inequality in large urban areas. We argue that our model is particularly suitable for such a task because our framework is quantitatively tractable and rich enough to replicate patterns in the dispersion of talent and wages of the Current Population Survey (CPS) for 54 MSAs in the US, for the year 2011. It features the existence of a unique equilibrium, which allows us to numerically quantify the equilibrium effects of place-based policies. Thus, by simply turning spatial sorting off and re-evaluating the model, we obtain a clear measurement of the impact of sorting effects on equilibrium quantities. In particular, we focus on a housing policy aimed to expand the housing supply in constrained cities by 1%. We find that such a policy could make the economy 0.2–0.4% more productive. However, we show that relaxing housing constraints in superstar cities has the unintended consequence of increasing aggregate wage inequality by the same magnitude.

Regional inequality is a complex topic and one that for several reasons requires general equilibrium reasoning. First, in spatial equilibrium, workers must not prefer other locations to their current one. This indifference condition implies that population, wages, and housing prices in all cities should be considered simultaneously, making them interdependent. Second, wages and housing prices are inseparable from local agglomeration externalities and the distribution of skills available in the city (Combes et al., (2008), Gennaioli et al. 2013, Acemoglu and Dell (2010), Van Nieuwerburgh and Weill (2010)). Investigating this association requires a model in which the labor productivity of each worker is endogenous to the location decisions of all workers. Third, to understand how city characteristics relate to regional inequality, we must disentangle how they affect agglomeration externalities and skill sorting separately. The reason is that a change in city characteristics (e.g., increasing the housing supply) could have an impact on local labor productivity by increasing density and by changing the skill composition of the city.

Our framework can be described as follows. We develop a general equilibrium model that allows for the sorting of heterogeneous workers with a continuum of skills in the presence of endogenous agglomeration externalities. In this model, workers care about their disposable income and the level of local amenities. Workers can freely choose where to live, but to access city's amenities, they must consume one unit of housing on that particular location. Since the number of houses per city is limited, workers compete for houses in bidding wars. By competing for limited housing, workers impose on each other a pecuniary congestion cost that depends on their skills through the bids they are willing to make. The strength of this effect ultimately depends on the tightness of the housing market (i.e., how many buyers show up to compete for a house), directly linking the features of a city's housing market to the characteristics of its talent pool.

Workers in cities perform non-tradable tasks that combine in a single final good. Each task requires only labor and workers are differentiated in their labor productivity, which depends on workers' skills. Since workers can freely move between cities, skill distributions are determined by a spatial equilibrium condition. The equilibrium of this economy yields, for each city, an endogenous distribution wages, and skills which are ultimately determined by fundamental city characteristics.

The calibrated model is consistent with several stylized facts. For instance, the model features the sorting of high-skilled workers into land-constrained cities. Given the higher supply of skilled workers, these cities specialize in high-productivity sectors, generating an endogenous correlation between low land availability and city productivity. Concurrently, land constrained cities have the property of having higher wage inequality driven by the interaction of endogenous relative prices and local agglomeration externalities. Furthermore, the calibrated model does a reasonable job of predicting moments that had not been targeted in the estimation, such as relative average house prices and city size.

Literature Review

This paper connects two important subjects in the urban economics literature: cities' characteristics (such as amenities and limited housing supply) and the heterogeneity in workers' skills. While extant research has shown that the physical geography of cities relates to their economic outcomes (Saiz 2010, Ganong and Shoag 2013, Hornbeck and Moretti 2015, Seegert 2016), a separate strand of the literature has focused on studying what determines the sorting of heterogeneous agent.[1] Eeckhout et al. (2010), Behrens and Robert-Nicoud (2014) and Puga et al. (2015) develop (theoretical) discrete agent models that explore how locations' fundamentals connect to productivity and inequality in cities. [2]

[1] (Glasear 2008, Shapiro 2012, Couture 2015, Albouy 2015, Albouy and Seegert 2016) have empirically shown the importance of amenities in accounting for sorting patterns. We build on this literature, and we quantify the tradeoff between amenities versus restrictions on the housing supply. A related literature has explored the sorting of heterogeneous firms (Gaubert 2015, Desmet and Rossi- Hansberg 2013 and Behrens et al. 2013, Serrato and Zidar 2014) to study the welfare implications of taxes and firm incentives. We complement this literature by focusing on the worker side. Further work is required to join these two threads in the literature.

[2] Frameworks that divide the workforce into discrete categories are empirically sensitive, since results depend on dichotomous definitions of what type of worker qualifies for each type of category. Indeed, Baum, Snow, Freedman, and Pavan (2015) show that if we change the definition of high-skilled worker to a worker with some college education, some of the results shown by Diamond (2015) no longer hold.

Our paper closely relates to Davis and Dingel (2012) in using a continuum of skill types, but departs from this study by providing a micro-founded housing market in which a restricted housing supply implies that, in equilibrium, a city may feature excess demand for housing (as in the superstar cities framework)[3]. As we can fully characterize and compute the unique equilibrium of the model using differential equations, we provide a tractable framework that is useful for quantitative policy evaluation. This analysis is exclusive to our setting.

This paper also relates to the literature that examines how city-level outcomes aggregate. Hsieh and Moretti (2015) use a Rosen–Roback framework to analyze the role of cities in aggregate growth. We contribute to this literature by providing a theory in which city productivity is endogenous to the interaction between sorting and local agglomeration externalities. A parallel literature has been studying wage inequality in cities (Baum-Snow and Pavan, 2012). Our model is also able to generate vibrant patterns in the distribution of wages both within and between cities. Furthermore, it predicts that the relationship between wage inequality and the level of local wages should follow a power function, which is consistent with the data.

Finally, this paper also speaks to a growing literature that seeks to evaluate the aggregate consequences of place-based policies. Despite many examples of local program evaluations, it is hard to assess the general equilibrium effects of these types of policies [4]. We contribute to this literature by quantifying the aggregate potential consequences of local policy changes by means of the spatial sorting of heterogeneous workers.

This paper is organized as follows. Section 2 presents the model and theoretical results. Section 3 discusses the data, describes the empirical estimation, and presents the main empirical results. Section 4 discusses policy implications and then concludes.

[3] To do so, we follow recent literature that models the housing market with bidding wars. For a review see Han and Strange (2015).

[4] A notable exception is Kline and Moretti (2014), who develop a methodology to estimate their aggregate effects.

2. Model

We consider an economy that contains $N > 1$ cities. Cities have a heterogeneous endowment of housing supply S^i and amenities a^i. Non-tradable services are produced within each city.

2.1. Production within a City

The economy is populated by a continuum of workers with skill $s \in [\underline{s}, \bar{s}]$ and these workers can move freely between cities. Denote $v^i(s) \geq 0$ as the endogenous supply of workers with skill s in city $i \in \{1, ..N\}$.

Cities produce one final good, and producing that final good requires the aggregation of intermediate tasks which will be indexed by their skill intensity $\sigma \in \Sigma = [\underline{\sigma}, \overline{\sigma}]$.

Production tasks are performed only through human capital, and workers vary in their productivity in these tasks. In particular, let $A(s, \sigma) > 0$ be the productivity of a worker of skill s in task σ [5]. $Y^i(\sigma) \geq 0$ is the endogenous output of task σ in city i and is given by,

$$Y^i(\sigma) = \int_{s \in S} A(s, \sigma) L^i(s, \sigma) ds,$$

where $L^i(s, \sigma) \geq 0$ is the endogenous number of workers with skill s who work on task σ in city i.

The output of the final good is given by a Dixit-Stiglitz production function:

$$Y^i = \kappa^i \left(\int_{\sigma \in \Sigma} [Y^i(\sigma)]^{\frac{\varepsilon-1}{\varepsilon}} d\sigma \right)^{\frac{\varepsilon}{\varepsilon-1}}$$

[5] To capture the idea that high skill workers have a comparative advantage in more complex tasks, we follow Costinot and Vogel (2010) and assume that productivity is log supermodular.

where, $\varepsilon > 1$ is the constant elasticity of substitution between tasks and κ^i is a city-level productivity shifter that captures the effect of agglomeration externalities.

Total profit for the final good is given by,

$$\Pi^i = \kappa^i \int_{\sigma \in \Sigma} [Y^i(\sigma)]^{\frac{\varepsilon-1}{\varepsilon}} d\sigma - \int_{\sigma \in \Sigma} p^i(\sigma) Y^i(\sigma) d\sigma$$

where $p^i(\sigma)$ is the endogenous price of task σ in city i.

Finally, total profits for intermediate tasks are given by,

$$\Pi^i(\sigma) = \int_{s \in S} [p^i(\sigma) A(s,\sigma) - w^i(s)] L^i(s,\sigma) ds,$$

where $w^i(s)$ is the endogenous wage for a worker with skill s in city i.

2.2. Housing market

We follow a directed-search model as in Albrecht, Gautier and Vroman (2013) (AGV, hereafter) to portray the housing market. Let $\theta^i = B^i/S^i$ be the tightness of the housing market in city i, where B^i is the total number of workers that bid for houses in city i, and S^i is the total amount of houses for sale.

The game has several stages:

1. Buyers randomly arrive to compete for a house.

2. Each buyer has a private valuation x, which will be the buyer's wage in city i. Buyers do not observe the number of other visitors to the house.

3. As buyers arrive at a house they compete for it following a first-price auction (with an unknown number of competitors).

With a random number of buyers, an individual buyer's optimal bid $b(x)$ is the weighted average of the buyer's optimal bids conditional on competing with n other competitors, with $n = \{0, 1, 2, 3, ...\}$, and has expression given by

$$b(x) = \frac{\sum p_n F(x;n) b(x;n)}{\sum p_n F(x;n)},$$

where p_n is the probability of a buyer competing with n other buyers, $b(x;n)$ is the optimal bid, and $F(x;n)$ is the distribution of types in the market. We assume B^i and S^i are large enough so that, in the limit, the arrival rate of buyers visiting a particular seller follows a continuous Poisson process with parameter θ^i. This characterization yields the endogenous housing prices shown in Appendix 1.

The tightness of the housing market also determines the probability of obtaining a house in the city. Since buyers' valuations for a house are sampled from a common distribution, we have a "Poisson race" among different players who arrive following the same arrival rate θ^i. Thus, $e^{-\theta^i(1-F^i(w^i(s)))}$ is the probability that a buyer of skill s makes the highest bid, and thus wins the auction. Notice that θ^i depends on the total number of workers who arrive in equilibrium to produce in city i. We assume that workers have rational expectations, thus they correctly anticipate the value of θ^i in equilibrium before moving into a city.

2.3. Workers' Preferences

Workers' have homothetic preferences over their disposable income x_i and the quality of local amenities a^i. Homothetic preferences can be represented as $U = Ta^i \log(x_i)$, where T is a deep preference parameter capturing the tradeoff between disposable income and the quality of local amenities. For convenience, we use the monotonic transformation $U = \exp(U) = x_i e^{Ta^i}$, as all the properties of this utility representation hold for monotonic transformations. In this case, the term e^{Ta_i} is interpreted as a utility shifter that depends on the amenities a^i and on the deep preference parameter T.

To enjoy the amenities of a particular city, workers need to own a house in that city. If they cannot afford a house, they receive a reservation utility $\underline{u}(s)$ from living in the suburbs. Workers are risk-neutral in whether they get to live in the suburbs or not.

The disposable income x_i is given by wages net of housing costs. Thus, by letting $b^i(w^i(s))$ represent the optimal bid of a worker of skill s for a house in city i, the disposable income is written as

$$x_i = w^i(s) - b^i(w^i(s)).$$

Hence, the expected value that a buyer of skill s receives from producing in city i is

$$U(w^i(s)) = (w^i(s) - b^i(w^i(s)))e^{Ta_i}e^{-\theta^i(1-F^i(w^i(s)))} + \underline{u}(s)(1 - e^{-\theta^i(1-F^i(w^i(s)))}) \tag{1}$$

In this context, using auctions to characterize the competition for limited housing can significantly simplify the problem. In particular, as the optimal bid for a house follows from a first price auction, then

$$b^i(w^i(s)) = w^i(s) - \frac{F^i(w^i(s))}{f^i(w^i(s))},$$

where $f^i(\cdot)$ is the first derivative of $F^i(\cdot)$. This implies that the worker's disposable income can be written as

$$x_i = w^i(s) - b^i(w^i(s)) = \frac{F^i(w^i(s))}{f^i(w^i(s))}. \tag{2}$$

Since an important fraction of labor rents accrue to land prices, the disposable income of a worker in city i is equal to her virtual surplus (i.e., informational rents). This means that the level of wages disappears and her utility depends only on her relative wage in city i. Substituting (2) in (1), we write the utility as

$$U^i(w^i(s)) = \frac{F^i(w^i(s))}{f^i(w^i(s))} e^{-\theta^i(1-F^i(w^i(s)))} e^{Ta_i} + \underline{u}(s)(1 - e^{-\theta^i(1-F^i(w^i(s)))}). \tag{3}$$

This characterization allow us to further simplify the problem as shown next.

Since wages are a monotone function of talent, we perform a change of variables to work in the space of types and not on the space of wages. In other words, there has exist a function V^i, such that $V^i(s) = F^i(w(s))$. Consequently, we rewrite (3) as

$$U^i(s) = \frac{V^i(s)}{v^i(s)} e^{-\theta^i(1-V^i(s))} e^{Ta_i} + \underline{u}(s)(1 - e^{-\theta^i(1-V^i(s))}),$$

where $v^i(\cdot)$ is the first derivative of $V^i(\cdot)$.

Notice the relevance of this change of variables and how it simplifies the overall problem. Now, the utility of a worker of living in city i is only a function of fundamental city characteristics (i.e. amenities a_i) and the endogenous distribution of skills s in that city (i.e. $V^i(s)$).

This transformation allows us to solve for the equilibrium in two steps. First, we solve for the spatial equilibrium for each type of worker as a function of the city attributes. After solving for the spatial distribution of talent, we impose the clearing conditions to obtain wages and housing prices in each city. Those steps are described next in detail.

Spatial Equilibrium: Sorting

Since there is free mobility, the utility of a worker of ability s must be equal across space.

Assumption We restrict attention to skills distributions on a close interval $[\underline{s}, \bar{s}]$ in which all cities share the same support of skills.

This assumptions simply tells us that, in equilibrium, all workers must be indifferent between all cities. In fact, this is consistent with empirical distributions of talent. Although we see differences between cities in the fraction of high-skilled to low-skilled workers, we still observe a positive mass of workers at every level of talent. Moreover, if we restrict attention to two cities, it is straightforward to show that, for any pair of non-overlapping skill distributions, this configuration is never in equilibrium, since the lowest-skill worker in the high-skilled city will always have an incentive to move to the low-skilled city, where she is the most skilled worker.

Assumption Given the costs of commuting to the center of the city, individuals capture only a fraction ϕ of the utility for living in city i.

This assumption is important to us because we can rewrite the worker's utility as

$$U^i(s) = \frac{V^i(s)}{v^i(s)} e^{Ta_i}[e^{-\theta^i(1-V^i(s))}(1-\phi)+\phi],$$

and use this expression to recover the reservation utility $\underline{u}(s)$ from the data. It is adopted purely for empirical purposes.

In essence, ϕ captures workers' sensitivity of not finding space in the city. If workers could commute at no cost, then ϕ would equal 1 and therefore housing restrictions would not matter. Given this preference representation, changes in T and ϕ are observationally equivalent for empirical purposes, since both speak to the tradeoff between disposable income and amenities. For this reason, we make the identifying assumption $\phi = 0$. Theoretical results do not change if $0 < \phi < 1$.

For each worker of skill s, it must be the case that her expected utility across cities is the same, i.e.

$$\mathbb{E}U^1(s) = \mathbb{E}U^2(s) = ... = \mathbb{E}U^N(s), \qquad \forall s \in [\underline{s}, \bar{s}]. \tag{4}$$

Since these equalities are preserved under any monotonic transformation, we can express (4) as the following linear combination of log functions:

$$\ln[\mathbb{E}U^i(s)] = \frac{1}{N-1}\sum_{j\neq i}\ln[\mathbb{E}U^j(s)]$$
$$\ln[\frac{V^i(s)}{v^i(s)}e^{-\theta^i(1-V^i(s))}e^{Ta_i}] = \frac{1}{N-1}\sum_{j\neq i}\ln[\frac{V^j(s)}{v^j(s)}e^{-\theta^j(1-V^j(s))}e^{Ta_j}],$$

for all type $s \in [\underline{s}, \bar{s}]$ and city $i \in [1, .., N]$.

Rearranging the system above, we obtain

$$\ln[v_i] = \ln[\prod_{j\neq i} v_j^{1/N-1}]\frac{V_i}{\prod_{j\neq i} V_j^{1/N-1}}e^{-\frac{1}{N-1}\sum_{j\neq i}\theta^j(1-V^j(s))+\theta^i(1-V^i(s))}e^{T[a_i-\frac{1}{N-1}\sum_{j\neq i}a_j]}]. \tag{5}$$

While the system in (5) can neither be easily solved numerically nor characterized by a closed-form solution, a linear approximation of the v_j terms yields to a tractable system which we can solve explicitly. The idea is to rearrange the system of differential equations such that all v_j terms are on the left hand side and all functions V_j and primitives of the model are on the right hand side as follows:

$$v_j = H(\vec{V}, \gamma),$$

where $\vec{V} = (V_1, V_2, ..., V_N)$, γ represent the primitive parameteres of the model, and H the expression in terms of the V_j terms and γ on the right hand side of the equality after the rearrangement.

To obtain this representation, we approximate the geometrical average of probability density functions (PDFs) v_j, where $0 < v_j << 1$, by its arithmetic average. Defining the geometric average of the probability density functions of different cities from i as

$$G_i \equiv \prod_{j \neq i} v_j^{1/N-1},$$

and taking a second-order Taylor expansion of G_i, it follows that

$$G_i \approx \Lambda_i - \sigma^2/2,$$

where $\Lambda_i = \frac{1}{N-1} \sum_{j \neq i} v_j$ and $\sigma^2/2$ is the sample variance.[6]

Since we have a closed economy, the sum of densities must be equal to the overall number of workers in this economy:

$$\sum_{j \neq i} v_j(s) + v_i(s) = v(s), \ \forall s \in [\underline{s}, \bar{s}], \tag{6}$$

where $v(s)$ is the exogenous aggregate PDF.

Thus, using expression (6) and the exogenous aggregate PDF $v(s)$, we rewrite the approximate geometric average as

$$G_i = \frac{1}{N-1}(v - v_i) - \sigma^2/2. \tag{7}$$

By substituting the approximate geometric average shown in (7) into (5) and exponentiating

[6]Note that we are dismissing the third-order terms, which are very close to zero in the case of probability densities.

both sides of the equation, we arrive at

$$v_i = [\frac{1}{N-1}(v - v_i) - \frac{\sigma^2}{2}][\frac{V_i}{\prod_{j\neq i} V_j^{1/N-1}} e^{-\frac{1}{N-1}\sum_{j\neq i}\theta^j(1-V^j(s))+\theta^i(1-V^i(s))} e^{T[a_i - \frac{1}{N-1}\sum_{j\neq i} a_j]}]. \quad (8)$$

Lastly, isolating v_i on the left hand side, we obtain the following system of ordinary differential equations (ODEs)

$$v_i = \frac{1}{(1 + \frac{1}{N-1}\frac{V_i}{\prod_{j\neq i} V_j^{1/N-1}} e^{-\frac{1}{N-1}\sum_{j\neq i}\theta^j(1-V^j(s))+\theta^i(1-V^i(s)} e^{T[a_i - \frac{1}{N-1}\sum_{j\neq i} a_j]})}$$
$$[\frac{1}{N-1}v - \frac{\sigma^2}{2}][\frac{V_i}{\prod_{j\neq i} V_j^{1/N-1}} e^{-\frac{1}{N-1}\sum_{j\neq i}\theta^j(1-V^j(s))+\theta^i(1-V^i(s)} e^{T[a_i - \frac{1}{N-1}\sum_{j\neq i} a_j]}].$$

[This next passage is unclear to me: Two observations are in order. First, from the above expression, we consider the problem of solving for the endogenous PDF as equivalent to the problem of solving for a mixing probability, since $\beta_i(s) \equiv v_i(s)/v(s)$ [7]. Thus, if we are to determine the endogenous PDF in equilibrium, we could determine the mixing strategy that players should follow in equilibrium. Second, as stated above, the above expression can be represented as

$$v_i = H(\vec{V}, \gamma)$$

which means that each probability density is a function of the vector of all probability distributions and a set of parameters γ.]

Since there is a continuum of skills and types, the multiplicative structure of $H(\vec{V}, \gamma)$ and the continuous differentiability of the functions composing it, it can be shown that the model

[7] Workers must be indifferent between all cities, thus, in equilibrium, each worker must choose a mixing probability $b(s)$ by which each worker of skill s randomizes between cities. If all workers follow the same mixing strategies, then, in equilibrium, we should expect to see the distributions predicted by such mixing probabilities.

has a unique spatial equilibrium. We provide the proof of the existence and uniqueness for the two-cities case in Appendix A.

Competitive Equilibrium

Once we have solved the spatial equilibrium, we can compute the competitive equilibrium for each independent city. We present next a definition and characterization of a competitive equilibrium in this economy.

Proposition 2.1. (Competitive Equilibrium)

In a competitive equilibrium, all firms maximize their profits and markets clear. This equilibrium is characterized by a continuous and strictly increasing matching function $M^i : S \rightarrow \Sigma$ such that

(i) $L^i(s,\sigma) > 0$ *if and only if* $M^i(s) = \sigma$*, and*

(ii) $M^i(\bar{s}) = \overline{\sigma}$ *and* $M^i(\underline{s}) = \underline{\sigma}$.

The matching function and wage schedule in each city is characterized by the following pair of differential equations

$$\frac{dM^i}{ds} = \frac{A(s, M^i(s))V^i(s)}{[p^i(M^i(s))/E^i(M^i(s))]^{-\varepsilon} \int_{s \in S} w^i(s) v^i(s) ds},$$

$$\frac{d \ln w^i(s)}{ds} = \frac{\partial \ln A[s, M^i(s)]}{\partial s},$$

where $v^i(s)$ *is the endogenous probability density function of salaries and* $V^i(s)$ *the cumulative density function for city* i *that solves the spatial equilibrium.*

The competitive equilibrium of the economy implicitly assumes that workers act as representative consumers and, regardless of where they live, consume all their income in the city. Analogously, our absentee landlords also behave as representative consumers.

3. Empirical Analysis

Data

We use the Current Population Survey (CPS) for March 2011. The CPS provides the wages for each MSA as well as the number of years of completed education, which we use as a proxy for talent.[8] The relative value of amenities in each MSA a^i is proxy by the hedonic parameters computed by Albouy (2015).[9] The measure comprises two parts: an *endogenous productivity* component (dependent on the workforce skill composition) and a *quality of life* component. We use the second component of this amenity index, since it is exogenous to the sorting of talent.

The last exogenous measure required by our model is a proxy for the [new?] housing stock in each MSA. Using the housing supply elasticities estimated by Saiz (2010), we assume that the supply of new houses in each MSA is proportional to the product of the housing elasticity and the [existing/current?] housing stock [is the data on housing stock also coming from Said 2010?].[10]

[8]Bacolod et al 2009 point out some limitations of using education as a proxy for talent. To check the robustness of our result, we use alternative talent data, as described in Appendix 4[COULDN"T FIND THIS APPENDIX]. Our findings are robust with respect to this alternative talent data as well.

[9]The author develop a methodology that derive hedonic measures of local productivity and local amenities from data, such as local wages, housing prices, and taxes. The quality of life measure positively correlates to measures of natural amenities relating to climate and geography.

[10]This measure stems from satellite-generated data on terrain elevation and the presence of water bodies to estimate the amount of developable land in each MSA. We focus on the part of the elasticity that is determined by geographical restrictions.

The last two datasets used in our calibration are the Zillow price index for the year of 2011 and the population size per MSA from the 2010 Census. While the calibration of the model does not target the moments of these metrics, they are used to assess the quality of fit for non-targeted variables, ultimately checking if the model can generate realist values for these endogenous quantities.

Calibration

There are two main sets of parameters in our model: (i) general parameters that have already been estimated in literature, such as the elasticity of substitution between services or skills ε, and the local agglomeration externalities η, and (ii) new parameters that are exclusive to our theoretical specification, such as the relative taste for amenities T and the complementarity between workers' skills and job complexity A.

To calibrate the parameters in the first group, we simply follow the literature. First, we fix the elasticity of substitution between services at $\varepsilon = 2$. Several authors, such as Katz and Murphy (1992) and Ciccone and Peri (2005), have estimated the elasticity of substitution between skilled and unskilled workers in the range of $[1,2]$. More recently, Hsieh and Klenow (2009) use an elasticity of substitution between manufacturing goods equal to 3. Our results are robust to the specification of this elasticity for values in the interval $[1,3]$.

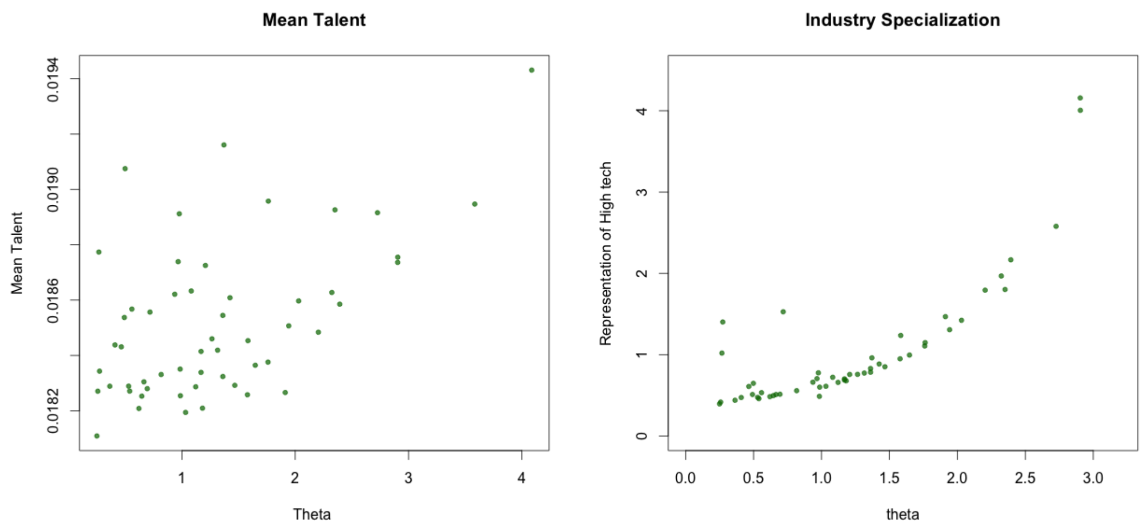
For the agglomeration externalities parameter η, we follow Moretti and Klein (2014), who show that the elasticity of agglomeration externalities with respect to density is constant, and fix it at $\eta = 0.08$. In our framework, agglomeration externalities are modeled as a city-level productivity shifter $\kappa^i = g_i(B^i)$, where $g_i(\cdot)$ is a function of the endogenous number of workers B^i coming to produce in city i. The productivity shifter follows a power function and satisfies $g_i(B^i) = (B^i)^\eta$.

There are two main implications of adopting this functional form for representing agglomeration externalities. First, it implies that a city's density makes workers more productive, because the productivity shifter is a monotonic increasing function of the city size. Second, it affects workers differently due to the curvature of the power function. High-skilled workers experience higher productivity gains from these local shifters relative to low-skilled workers.

To calibrate the parameters in the second group, we perform an indirect inference estimation. We exploit the fact that our model can be solved in two separate parts (i.e., the spatial sorting of workers and the competitive equilibrium in each city) and split the estimation procedure into two parts as well. In the first part, we solve the spatial sorting problem and recover only the parameter capturing the taste for amenities. Once we achieve an endogenous distribution of skills for each MSA, we solve for the distribution of wages in each city afterward, because the cities produce non-tradable services. From these wage distributions, we recover the second parameter representing the skill–technology complementarity. Our indirect inference estimation produces a taste of amenities $T = 7$ and a job complexity $A = 1.36$. Appendix C contains a detailed description of the estimation procedure.

Results

The calibrated model allows us to investigate not only the relationship between the endogenous quantities, such as the distribution of wages, housing prices, and the distribution of talent by MSA, but also how these variables respond to the city fundamental characteristics, such as the tightness of the housing market. As discussed in Section 2.2, we measure the housing market tightness of city i with the equilibrium ratio between buyers and sellers θ^i. It is worth noticing that θ^i reflects both dimensions of the city characteristics (i.e. amenities and limited housing supply), since high-theta cities are the result of high amenities and a restrictive housing supply.



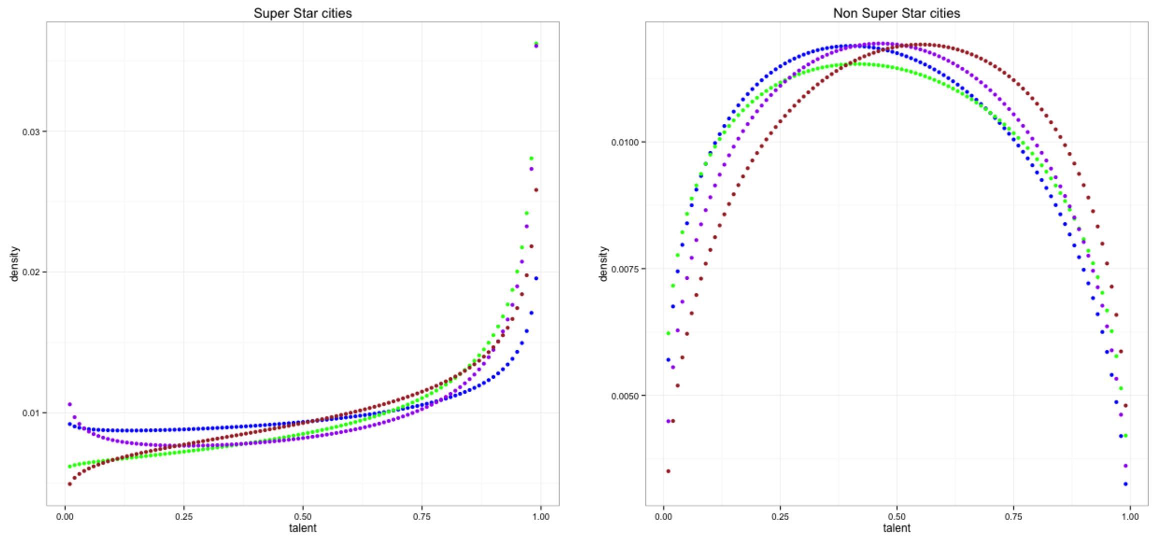
(a) Mean talent

(b) Industry Specialization

Figure 1: *Sorting and Specialization.*

Figure 1(a) shows that, in equilibrium, more talented workers tend to concentrate in high-theta cities (i.e. superstar citites). The average talent of workers increases as we move to cities featuring a tighter housing markets. Although all cities produce all goods in equilibrium, the fraction of these goods [each good relative to the aggregate output?] depends on the endogenous supply of talent available in each local economy. As a result, cities in our model are able to specialize.

Figure 1(b) plots the degree of specialization for each city, [measured by what? What is the representation of High Tech here? Is it HighTechOutput / AggregateOutput?]. The model predicts that superstar cities tend to specialize in tasks that require more talent and are more technology intensive.

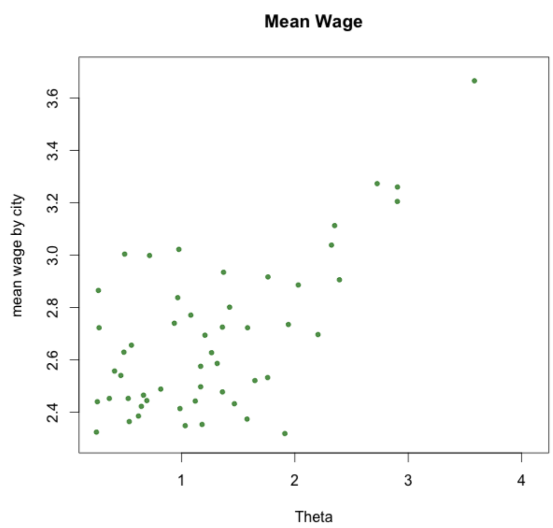


(a) Superstar cities

(b) Non-superstar cities

Figure 2: *Talent Distribution.* The figure shows the talent distribution for high- and low-theta cities. Panel 2(a) presents the probability density function for four high-theta cities, with $\theta \in \{?, ?, ?, ?\}$. Panel 2(b) presents the probability density function for four low-theta cities, with $\theta \in \{?, ?, ?, ?\}$. The support of the talent distribution is the interval $[0, 1]$.[Put legend with the colors]

Figure 2 shows the skill distributions predicted by the model for a subset of high-theta cities in Panel 2(a) and low-theta cities in Panel 2(b). Panel 2(a) reveals that superstar cities feature talent polarization, with U-shaped skill distributions. Thus, the skill distribution of high-theta cities displays "fat tails" for both high- and low-skilled workers. In fact, the density of the high-theta cities is the highest for highly skilled workers (i.e. with $s \in [0.75, 1]$). On the other hand, the skill distribution of low-theta cities presented in Panel 2(b) has an inverted U-shape. The plot indicates that non-superstar cities tend to have a large density of average skilled workers, while lacking very high (and low) skilled workers. [Can we show the EMPIRICAL distribution of talents for 2 superstar and 2 nonsuperstar to show that the model generates the correct shape?]



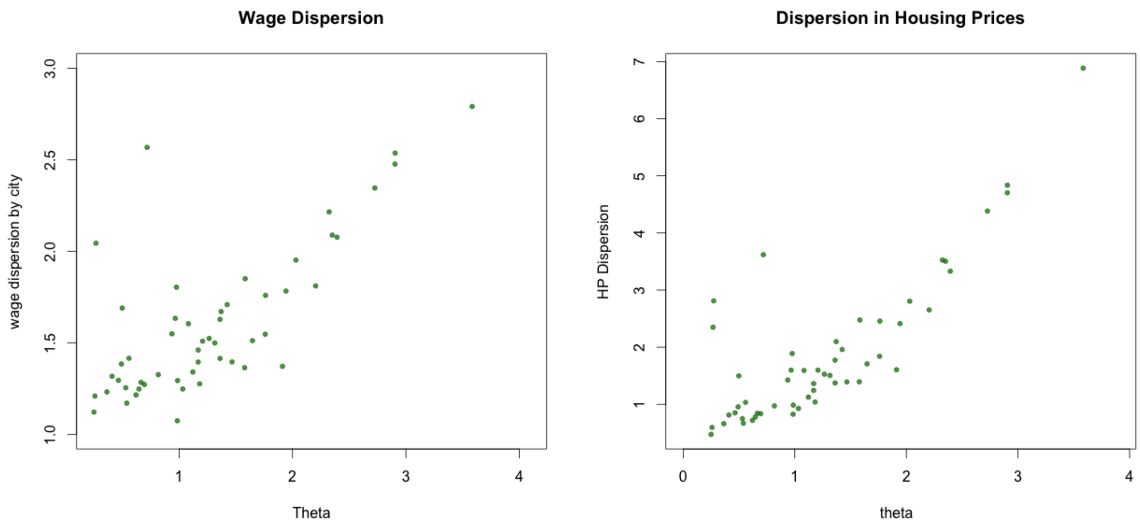
(a) Average wage

(b) Housing price

Figure 3: Wage and Housing Prices in Superstar Cities

The heterogeneity among skill distributions in superstar and non-superstar cities has profound consequences for two key economic variables: wage and housing price. Figure 3(a) shows the impact of θ in the mean wage for each of the 54 MSAs investigated. As illustrated, virtually all cities with low amenities and high housing supply (i.e., $\theta < 2$) have an average salary in the interval $[2.2, 3]$. However, in cities with high amenities and low housing supply (i.e. $\theta \geq 2$), the mean wage varies in $[3, 3.6]$.[dollars? in which units are the wages represented?] In essence, the graph indicates that these differences in talent distribution render significant differences in local productivity, measured by workers' salaries.

A similar pattern is observed in the average housing price, displayed in Figure 3(b). With the exception of 5 MSAs in the top left corner of the graph, housing prices display almost a monotonically increasing relationship with θ. While the MSA with the highest θ has the highest average housing price, the city with the lowest θ has the lowest average housing price. Overall, the graphs show that sorting not only generates between-city differences, but also cause important variation within cities.

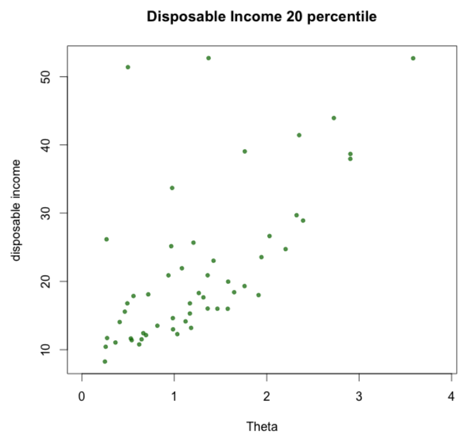


(a) Wage dispersion

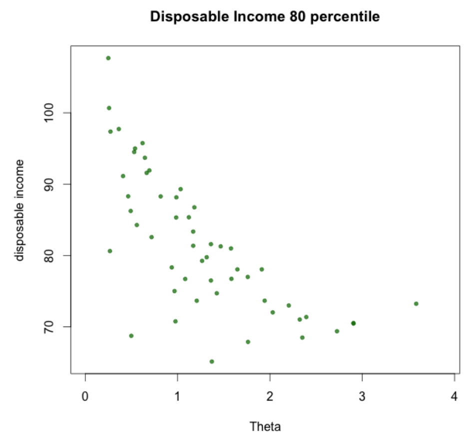
(b) Housing price dispersion

Figure 4: *Wage and Housing Price Dispersion.* [Do we have the empirical graph?]

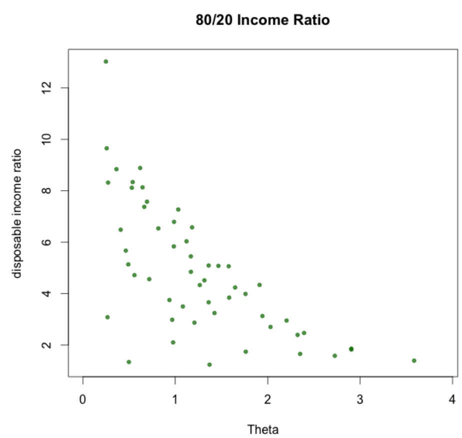
Figure 3 illustrates that city characteristics affect not only the level but also higher moments of wage and housing prices. Panel 4(a) shows how wage dispersion varies with θ. As illustrated, superstar cities (i.e. high-theta cities) have the largest wage dispersion and, consequently, are more unequal places. Panel 4(b) shows that the dispersion of house prices is also higher in supercities and the relationship between the house price and θ is practically monotonic. Together, the graphs indicate that, although workers are indifferent between cities in equilibrium, their disposable income varies considerably in the cross section.



(a) Disposable income 20th percentile worker by city type.



(b) Disposable income 80th percentile worker by city type.



(c) Disposable income ratio between 80th percentile and 20th percentile worker by city type.

Figure 5: Disposable income by city type

Another critical question that our framework allows us to answer is how amenities and limited housing supply (i.e. θ^i) affect the distribution of workers' disposable income across cities. To put it differently, we investigate if the 20% poorest (richest) workers in a low-theta city earn more or less than the 20% poorest (richest) workers in high-theta cities. Figure 5 presents the answer to this question.

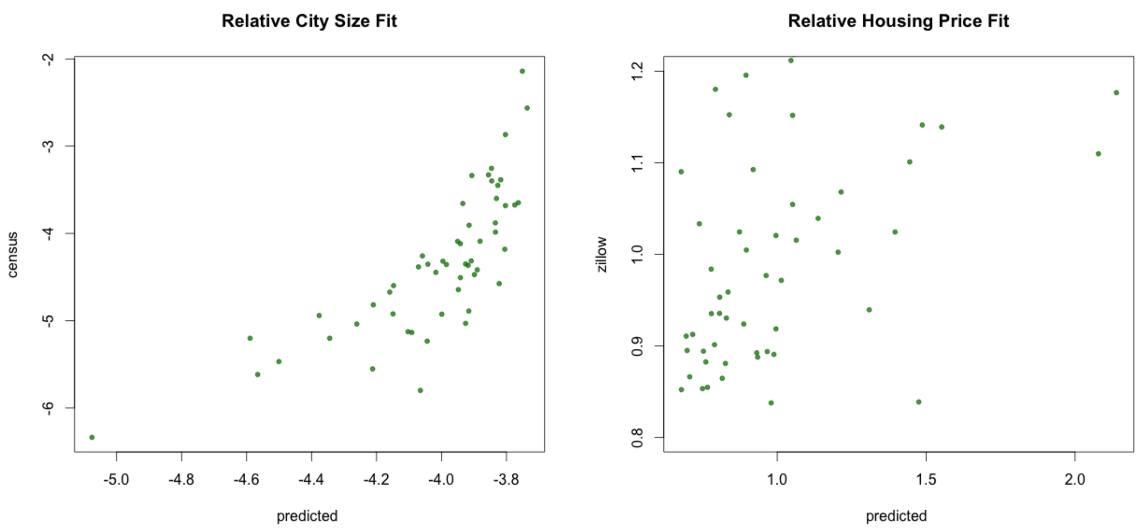
[Unclear passage: Figure 5(a) shows that the disposable income of workers in the 20th percentile of the talent distribution in low-theta cities is significantly lower than its counterpart in high-theta cities. The economic rationale for this result goes as follows. Low-skilled workers must be compensated for living in superstar cities, since they have very low chances of finding a house within the MSA. If they are pushed to the suburbs, they are not able to enjoy the city's amenities. Thus, the low supply of low-skilled workers makes their relative talent very valuable in high-theta cities, which reflects in high relative prices and wages.]

Figure 5(b) shows that the effect of θ^i on the disposable income of the 20% richest workers in each city is the opposite. Workers in the 80th percentile of the talent distribution have lower disposable income in supercities (i.e. high-theta cities). The reason is that they tend to pay a higher fraction of their wages for living in the city center. As a result, they become very sensitive to amenities and are willing to sacrifice some income to enjoy the urban living.

To understand which of these two groups of workers are affected the most by variations in θ^i, we plot in Figure 5(c) the ratio between the disposable income of workers in the 80th and 20th percentile. As the graph shows, this ratio is a decreasing function of θ^i, which indicates that the effects θ^i on the wages of low-skilled workers outweigh the changes caused in the disposable income of high-skilled workers. [Unclear passage: This shows that despite the fact that superstar cities are more unequal places because of their bimodal talent distributions, they tend to diminish differences in disposable income between high-and low-skilled workers.]

External Validity

We investigate the ability of the calibrated model of reproducing patterns that are not targeted in the original exercise. In particular, we evaluate if the model's prediction for two key variables: the relative size of cities and their relative housing prices. [why the quantities here are "relative"? Are they relative to what?]

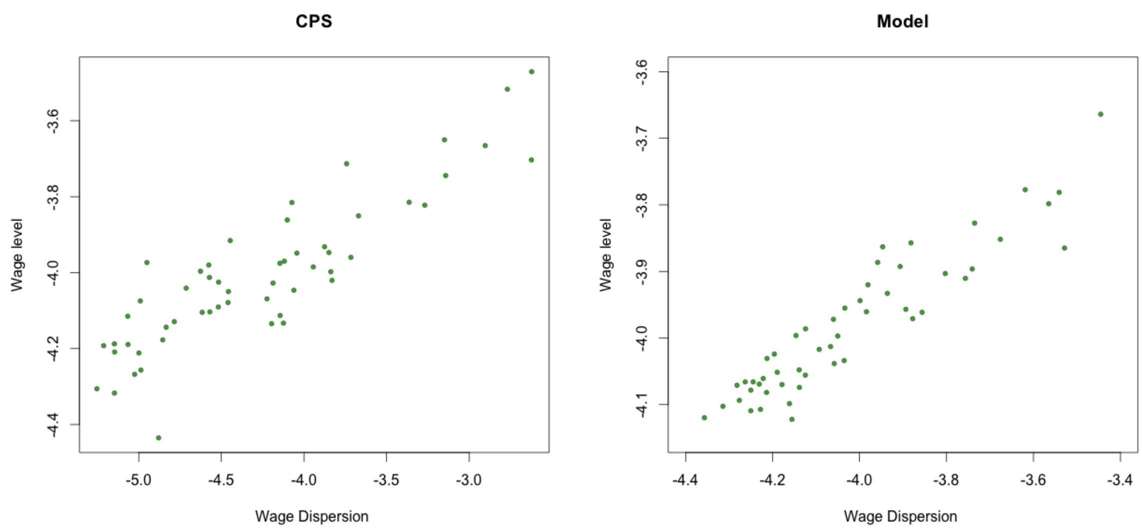


(a) Observed (Census 2010) versus predicted city sizes.

(b) Observed (Zillow index) versus predicted average housing prices.

Figure 6: Non-targeted Variables

Figure 6 shows the plots the predicted-versus-real for the city size and average housing prices. If the model perfectly matches the data, all the observations should align along the 45 degree line. Figure 6(a) compares the city sizes implied by the model and the empirical data from the 2010 Census. It shows that the model does a decent job predicting the real size of cities, tending to slightly overestimate them. On the other hand, Figure 6(b) compares the house prices implied by the model and the 2011 Zillow price index. The graph shows that model tend to underestimate the housing prices across cities and display an inferior performance relative to the previous variable.



(a) Empirical Correlation

(b) Theoretical Correlation

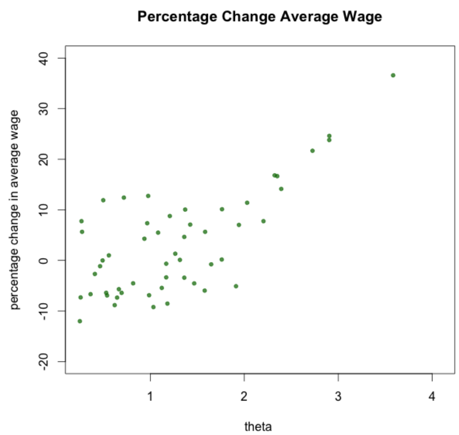
Figure 7: *Correlation.* This figure shows the correlation structure between wage level and wage dispersion predicted by the model versus the correlation observed in the data. Graphs are in logarithmic scale. Panel (a) plots the relation between average wages and wage dispersion for the March 2011 CPS. Panel (b) plots the relation between average wages and wage dispersion as predicted by the model.

Another critical pattern reproduced by the model is the correlation structure between the wage level and its dispersion. Using the 2011 CPS data, we document in Panel 7(a) that there is a high correlation between wage level and wage dispersion. To put it differently, the data indicates that places with higher wage inequality tend to have higher average salaries. Panel 7(b) shows that our model replicates this pattern quite well. It is worth noticing that, as the previous variables (city size and housing prices), the correlation between average wage level and its dispersion is not a targeted variable in our calibration exercise. Both quantities—wage level and dispersion—are calibrated independently, as described in Appendix C.

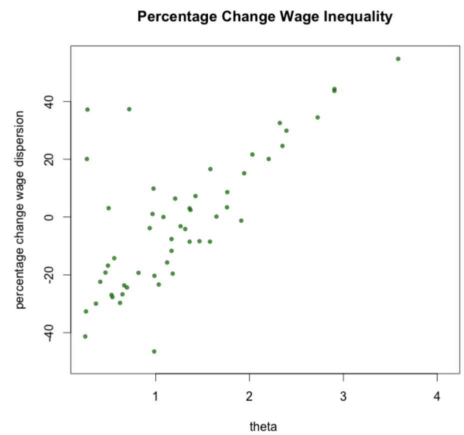
Unpacking Sorting Effects

Since the location decisions of workers are not driven by local productivity differences, but rather by city characteristics (i.e. amenities and limited housing supply), our framework allows us to separate what fraction of the wage dispersion is due to the sorting of heterogeneous agents between cities and what fraction of the wage dispersion is due to local agglomeration externalities. To obtain the first fraction of the wage dispersion, we fix the size of each city and impose the same skill distribution everywhere. We the model now lacking sorting effects, we recalculate the dispersion in wages and housing prices both between and within cities. Our numerical results shows that in the absence of sorting effects the total wage dispersion drops by 7.5% and housing price dispersion drops by 5.7%. While wage inequality and dispersion in housing prices decline, the overall economy experiences a fall in aggregate productivity, with the total GDP dropping 1.9%. The reason for this striking difference is that when we allow for the sorting of heterogenous agents, most productive workers tend to cluster in larger cities. In those large metropolitan areas, the agglomeration externalities boosts the workers' marginal productivity, which results in a higher aggregate output.

To better visualize the effects of talent sorting on the average wage, wage dispersion, and housing prices across the 54 MSAs, we plot in Figure 8 the changes in these variables in an economy where talent sorting is present relative to an economy where sorting is absent.



(a) Predicted change in the average wage.



(b) Predicted change in wage dispersion.

(c) Predicted change in average housing price.

Figure 8: *Sorting effect by city type*

Figure 8(a) shows the changes in the average wage per city. The graph shows that superstar cities experience an average wage increase of around 20–40% with sorting, whereas non-superstar cities oscillate between −10% and 10%. This result indicates that sorting effects can generate significant wage variations across cities.

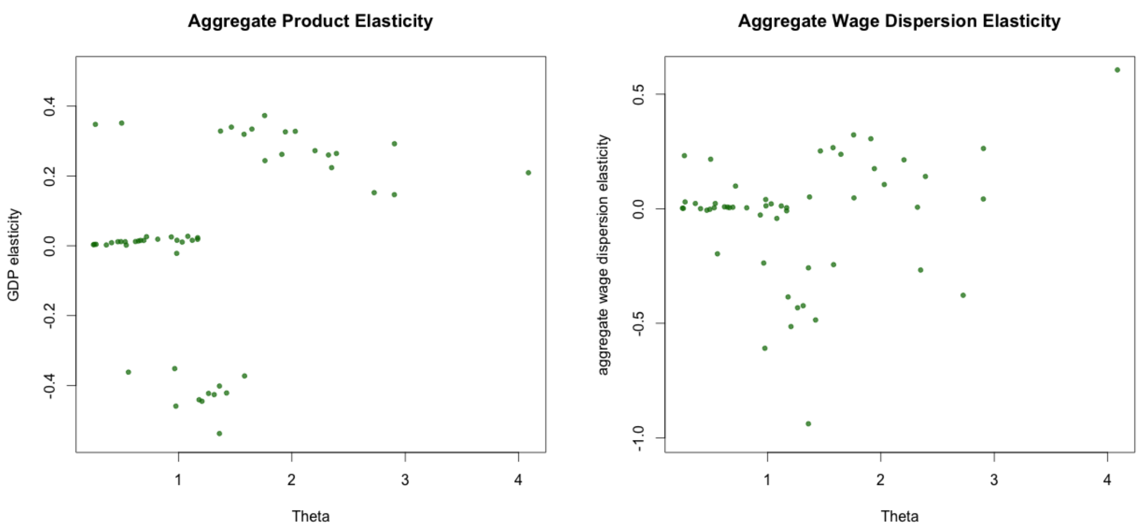
Figure 8(b) plots the changes in wage dispersion within cities as a function of city characteristics (i.e. θ^i). As illustrated, sorting have a different impact on low- and high-theta cities. For cities with $\theta^i > 3$ (i.e. high amenities and low housing supply), the wage inequality within cities is between 20% and 40% higher due to talent sorting. On the other hand, cities with $\theta^i < 2$, talent sorting induces a predominantly negative effect on wage dispersion, indicating that wage inequality is reduced in those metropolitan areas due to the accumulation of low-skilled workers. Thus, while the sorting of heterogenous agents contributes to the wage inequality within superstar cities, it promotes wage equality in non-superstar cities.

Lastly, Figure 8(c) shows the effects of talent sorting for the changes in housing prices. For cities with $\theta^i < 2$, the effects of talent sorting on housing prices are unclear, since house prices varying from -30% to 20% in our experiment. However, the effects of talent sorting for superstar cities are immediately understood. As the graph indicates, for cities with $\theta^i > 2$ (i.e. high amenities and restricted housing supply), talent sorting increases average housing prices by 20–40%.

We finish this section presenting the effects of agglomeration externalities on wage level and dispersion. To assess that, we set the agglomeration externality parameter to zero and re-evaluate the equilibrium in an economy that features sorting in the absence of agglomeration externalities. We then compute the change in average wages and wage dispersion relative to the benchmark economy that features talent sorting and agglomeration externalities. Our numerical experiment shows that changes in average wages between cities is, on average, 17% higher in the presence of agglomeration externalities, while changes within cities wage dispersion is, on average, 7% higher. Overall, these two forces seem to complement each other.

Place-Based Policies: National Housing

In this section, we evaluate the consequences of changes in the local housing supply to aggregate variables, such as productivity and wage inequality. In particular, we independently "shock" each city with an increase in housing supply equivalent to 1% of the national housing stock and evaluate the response to the aggregate output and wage dispersion. Figure 9 presents the results of this experiment.



(a) Percentage change in average wages. (b) Percentage change in wage dispersion.

Figure 9: Aggregate Effects of Shocking Housing Supply by City Type.

Figure 9(a) shows the effects of the increase of 1% in the housing supply of a particular city on the aggregate output (i.e. GDP). As illustrated, for the majority of the cities where the housing constraints are not binding (i.e. $\theta^i < 1$), the housing shock does not generates significant effects on the GDP. In these cities, aggregate effects are very close to zero. Interestingly, for cities where the housing constraints are binding but the shortage of housing is not as severe (i.e. $1 < \theta^i < 1.5$), aggregate output decreases by approximately 0.4%. It is worth noticing that expanding the housing supply does not change the number of workers in the economy, but rather the characteristics of the local housing markets. The reallocation of workers induced by those changes in city characteristics can generate higher or lower productivity as the result of the interaction between skill sorting and endogenous local agglomeration externalities. By relaxing housing constraints in these cities, we make them more attractive to skilled workers. As a result, some skilled workers stay away from larger urban centers and the productivity gains to the overall economy coming from agglomeration externalities is partially lost, resulting in a decline of the GDP.

Finally, the graph shows that expanding the supply of houses in high-theta cities (i.e. $\theta^i > 2$), the economy grows between 0.2–0.4% more. By increasing the housing supply in superstar cities, we expand the accessibility of these productive environments to more workers. As these cities grow, they also become more productive. Yet, the sorting of skilled workers moderates these effects.

Although it is the case that expanding superstar cities generates gains in aggregate productivity, this expansion comes at the cost of higher wage inequality. Figure 9(b) shows the change in ex-post aggregate wage dispersion after a housing shock. As before, a housing supply shock in places where the housing constraints are not binding has no effect on aggregate wage inequality. As we move to high-theta cities, we find that inequality is higher for cities featuring tighter housing constraints. In particular, these cities experience an influx of low-skilled workers, creating a negative congestion in the city they move to. As the relative supply of high-skilled workers decreases in large cities, their salaries increase. At an equal rate, low-skilled workers have their salaries depressed by the expansion of their relative talent supply.

Overall, while relaxing housing constraints in superstar cities generates important gains in aggregate production, these policy changes also entail adjustments in the composition of the local labor force that ultimately generates the unintended consequence of increasing wage inequality.

4. Conclusion

In this paper, we examine how city characteristics, such as housing supply and local amenities, affect the sorting of heterogeneous agents between cities. We developed an urban macro model in which cities have a restricted supply of houses, and workers with a continuum of skills compete for limited space through bidding wars. In this context, the pecuniary congestion costs that heterogeneous workers impose on each other operate as an endogenous driver for gentrification. The model has a unique equilibrium that can be calibrated to match different moments of the talent and wage distribution for a cross-section of US cities.

Overall, our numerical simulations stress that the sorting of heterogeneously skilled workers can generate sizable aggregate effects on productivity and inequality, mostly driven by the interaction between sorting and local agglomeration externalities. In particular, we find that the sorting of heterogeneous workers accounts for 7.5% of the total variation and generates considerable differences between cities. Sorting mostly affects cities that feature tighter housing markets, making them between 20% and 40% more productive. In the absence of sorting, aggregate production falls by 1.9% due to the loss in agglomeration externalities resulting from the reallocation of high-skilled workers from large urban areas to small cities.

Finally, we use the calibrated model to evaluate place-based policies. We find that policies designed to improve labor mobility, such as the expansion of housing in constrained urban centers, can have unintended consequences to wage inequality, given the sorting of heterogeneously skilled workers in the presence of local agglomeration externalities. In particular, we estimate that the expanding housing supply in cities with tighter housing markets increases productivity between 0.2% and 0.4%. This increase is mitigated by sorting as high-skilled workers tend to relocate to cities where they are less productive. Although these policies aim to reduce inequality by facilitating the spatial mobility of workers, they can produce an unexpected increase in aggregate wage inequality by the same magnitude.

Bibliography

1. Acemoglu, Daron and Autor, David “Skills, Tasks and Technologies: Implications for Employment and Earnings” 2011, Handbook of Labor Economics Volume 4

2. Acemoglu, Daron and Dell, Melissa, “Productivity Differences Between and Within Countries” 2010, American Economic Journal: Macroeconomics

3. Albrecht, James, Gautier, Pieter and Vroman, Susan,. “Directed Search in the Housing Market” 2016, Review of Economic Dynamics

4. Albouy, David , “What are Cities Worth? Land Rents, Local Productivity, and the Total Value of Amenities” 2016, Review of Economics and Statistics

5. Albouy, David and Seegert, Nathan, “The Optimal Population Distribution Across Cities and The Private-Social Wedge" 2016, Working Paper

6. Baum-Snow, Nathaniel , Freedman, Matthew , Pavan, Ronni “Why has urban inequality increased” 2014, Processed, Providence, Brown University

7. Baum-Snow, Nathaniel , Pavan, Ronni “Inequality and City Size” 2013, The Review of Economics and Statistics, vol. 95, issue 5, pages 1535-1548

8. Bacolod,Marigee ,Blum, Bernardo and Strange, William “Skills in the city” 2009, Journal of Urban Economics

9. Behrens, Kristian, Gilles Duranton, and Frederic Robert-Nicoud “Productive cities: Agglomeration, selection, and sorting” 2014, Journal of Political Economy, Vol. 122, No. 3, p. 507-533

10. Kristian Behrens and Frederic Robert-Nicoud “Agglomeration theory with heterogeneous agents” 2015, Handbook of Regional and Urban Economics, vol. V, pp. 171-245

11. Busso, Matias, Jesse Gregory and Patrick Kline, “Assessing the Incidence and Efficiency of a Prominent Place Based Policy” 2013, American Economic Review,103(2): 897-947

12. Ciccone, Antonio and Peri, Giovanni “Identifying human-capital externalities: Theory with applications”, 2005, The Review of Economic Studies, vol 73, pp. 381-412

13. Combes, Pierre-Philippe & Gobillon, Laurent, "The Empirics of Agglomeration Economies," 2014, IZA Discussion Papers 8508, Institute for the Study of Labor (IZA).

14. Combes, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, Pierre-Philippe, Gilles Duranton, and Laurent Gobillon, “Spatial wage disparities: Sorting matters!” 2008, Journal of Urban Economics 63 (2), 723–742

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

16. Costinot, Arnaud and Vogel, Jonathan “Matching and Inequality in the World Economy”, 2010, Journal of Political Economy, vol. 118, issue 4, pp. 747-786

17. Couture, Victor “Valuing the Consumption Benefits of Urban Density”, 2015, Working paper

18. Davidoff, Thomas “Supply Constraints are Not Valid Instrumental Variables for Home Prices Because They are Correlated with Many Demand Factors” 2016, forthcoming, Critical Finance Review

19. Davis, Donald R., and Jonathan I. Dingle “A spatial knowledge economy”, 2013, NBER Working Paper

20. Davis, Donald R., and Jonathan I. Dingle “The Comparative Advantage of Cities” 2015, NBER Working Paper

21. Diamond Rebecca “The Determinants and Welfare Implications of US Workers’ Diverging Location Choices by Skill: 1980-2000”, 2015, Working paper

22. Duranton, Gilles and Puga, Diego, "Micro-foundations of urban agglomeration economies,"

2004, Handbook of Regional and Urban Economics, in: J. V. Henderson & J. F. Thisse (ed.), Handbook of Regional and Urban Economics, edition 1, volume 4, chapter 48, pages 2063-2117 Elsevier

23. Eeckhout, Jan, and Philipp Kircher, "Assortative matching with large firms: Span of control over more versus better workers," 2012, Working Paper

24. Eeckhout, Jan, Pinheiro, Roberto and Schmidheiny, Kurt "Spatial sorting: Why New York, Los Angeles and Detroit attract the greatest minds as well as the unskilled," 2010, Working Paper

25. Gaubert, Cecile "Firm Sorting and Agglomeration", 2015, Working Paper

26. Ganong, Peter and Shoag, Daniel "Why Has Regional Income Convergence in the U.S. Declined?" 2013, HKS Working Paper No. RWP12-028

27. Gennaioli, Nicola, Rafael LaPorta, Florencio Lopez-de-Silanes, and Andrei Shleifer "Human Capital and Regional Development." 2013, Quarterly Journal of Economics 128 (1): 105-164

28. Guerrieri, Veronica, Hartley, Daniel, Hurst, Erik, "Endogenous Gentrification and Housing Price Dynamics" 2013, Journal of Public Economics, 100, 45-60

29. Glaeser, Edward, Kolko, Jed and Saiz, Albert "Consumer city", 2001, Journal of Economic Geography 27-50

30. Glaeser, Edward L. and Mare, David C "Cities and skills" 2001, Journal of Labor Economics, 316-342

31. Glaeser, Edward and Joshua Gottlieb, "The economics of place-making policies," 2008, Brookings Papers on Economic Activity, 155–239

32. Gyourko, Joseph, Mayer, Christopher and Sinai,Todd "Superstar Cities" 2013, American Economic Journal: Economic Policy, American Economic Association, vol. 5(4), pages 167-99

33. Han, Lu and Strange, William "Bidding Wars for Houses" , 2014, Real Estate Economics; Issue: 41(3)

34. Hsieh, Chang-tai, Klenow, Peter "Misallocation and Manufacturing TFP in China and India" 2009, The Quarterly Journal of Economics 124 (4): 1403-1448

35. Hsieh, Chang-tai and Moretti, Enrico "Why Do Cities Matter? Local Growth and Aggregate Growth" 2015, NBER Working Paper

36. Hornbeck, Richard and Moretti, Enrico "Who Benefits From Productivity Growth? The Local and Aggregate Impacts of Local TFP Shocks on Wages, Rents, and Inequality" 2015, NBER Working Paper

37. Katz, Lawrence and Murphy, Kevin "Changes in Relative Wages, 1963–1987: Supply and Demand Factors" 1992, The Quarterly Journal of Economics, vol. 107, issue 1, pages 35-78

38. Kline, Patrick and Moretti, Enrico "Place Based Policies with Unemployment" 2013, American Economic Review, Papers and Proceedings, 103 (3)

39. Kline, Patrick and Moretti, Enrico "Local Economic Development, Agglomeration Economies and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority" 2014, Quarterly Journal of Economics

40. Melo, Patricia, Graham, Daniel and Noland, Robert , " A meta-analysis of estimates of urban agglomeration economies" 2009, Regional Science and Urban Economics, vol. 39, issue 3, pages 332-342

41. Moretti, Enrico "Human Capital Externalities in Cities" 2004, Handbook of Regional and Urban Economics, North Holland-Elsevier, 4: 2243.2291

42. Moretti, Enrico "The New Geography of Jobs" 2012, Houghton Mifflin Harcourt

43. Moretti, Enrico "Local Labor Markets" 2011, In Vol. 4 of Handbook of Labor Economics, Chapter 14, 1237.1313. Elsevier

44. Puga, Diego, De la Roca, Jorge and Ottaviano, Gianmarco “City of dreams”, 2015, Working Paper

45. Roback, Jennifer “Wages, Rents, and the Quality of Life” 1982, Journal of Political Economy, 90(6): pp. 1257.1278

46. Rossi-Hansberg, Esteban and Wright, Mark “Establishment size dynamics in the aggregate economy” 2007, American Economic Review, 97 (5), 1639

47. Rosen, Sherwin “Wages-based Indexes of Urban Quality of Life” 1982, In Current Issues in Urban Economics

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

49. Seegert, Nathan, “Land Regulation and the Optimal Distribution of Cities ” 2016, Working paper

50. Serrato, Juan Carlos and Owen Zidar, “Who benefits from state corporate tax cuts? A local labor markets approach with heterogeneous firms” 2014, Technical Report, Manuscript

51. Shapiro, J.M. “Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital” 2006, The Review of Economics and Statistics, 88(2): 324.335

52. Van Nieuwerburgh, Stijn and Weill, Pierre-Olivier “Why Has House Price Dispersion Gone Up?” 2010, Review of Economic Studies, Oxford University Press, vol. 77(4), pages 1567-1606

53. Zubizarreta, J. R., "Stable Weights that Balance Covariates for Estimation with Incomplete Outcome Data," 2015, Journal of the American Statistical Association, 110, 910–922

Appendix: For Online Publication

A Two-Cities model

Given that for the general system of cities we can only characterize an approximation, we will study the case with two cities, in which the approximation coincides with the exact solution of the model. As we are able to provide analytical results for this specific case, these will be informative of the general properties of the model. In particular, we want to understand the role of housing restrictions in the model, and how these translate into different allocations of talent and consequent wage differentials between cities.

Proposition A.1. *There exist a unique equilibrium for each competitive equilibrium in a city.*

Proof. Since there is free mobility, the utility of a worker of ability s must be equal across space. Workers must be indifferent between the two cities. Then, for every worker s, it must be the case that

$$\mathbb{E}U^1(s) = \mathbb{E}U^2(s) \quad \forall s \in [\underline{s}, \bar{s}] \ .$$

The second condition is that we have a fixed amount of talent within the country (closed economy), which means that

$$V(s) = V^1(s) + V^2(s) \quad \forall s \in [\underline{s}, \bar{s}] \ .$$

Thus we can re-write the above condition as:

$$\frac{1-V^1(s)}{v^1(s)}e^{-\theta^1(1-V^1(s))} = \frac{1-V(s)+V^1(s)}{v(s)-v^1(s)}e^{-\theta^2(1-V(s)+V^1(s))} \;\; \forall s \in [\underline{s}, \bar{s}] \; .$$

There are two important observations to make here: This is a difficult problem, since we must search for a function $V^1()$ that can hold this condition for every s. Now, it is crucial to note that, since this must hold for every point and that $v^1 = \frac{\partial V^1}{\partial s}$, this is equivalent to solving an ordinary differential equation (ODE) for every point s. Also notice that $\langle V(s), v(s), \theta^1, \theta^2 \rangle$ are exogenous parameters.

Thus by rearranging the terms, we arrive at the following expression for the equilibrium ODE:

$$\frac{\partial V^1(s)}{\partial s} = v(s)\frac{[1-V^1(s)]e^{-\theta^1(1-V^1(s))}}{[[1-V^1(s)]e^{-\theta^1(1-V^1(s))} + [1-V(s)+V^1(s)]e^{-\theta^2(1-V(s)+V^1(s))}]} \qquad \forall s \in [\underline{s}, \bar{s}]$$

Given that we have a first-order differential equation with an initial condition, we must show that$f(s, V1(s))$ is Lipschitz–continuous in V^1 and continuous in s. Lipschitz continuity (i.e., bounded variation) is easy to observe. Notice that $f(s, V1(s))$ is bounded by one (since it is a weighted average), thus it is Lipschitz. Second, since V is continuous by definition, the composition of continuous functions is also continuous, thus $f(s, V1(s))$ is continuous in s. Note that $\langle j^1, j^2 \rangle$ are endogenous parameters, since they depend on the number of workers that come to produce to the city. So, it is necessary to find a fixed point. We fix j^i, solve the system of differential equations, and compute a new tightness parameter. We do this until j^i converges to a stable parameter. We can now recall the Picard–Lindeloff Theorem, which states that there exists a unique solution to this contraction. □

The next property captures the idea that City 1 (i.e., the city with a more inelastic housing supply) have relatively more high-skill workers than City 2. We formalize this result in the next proposition.

Proposition A.2. (Monotone likelihood ratio property) *For two cities that have the same level of amenities, whenever $\theta^1 > \theta^2$, the monotone likelihood property for the distribution of talents in City 1 versus City 2 holds:*

$$v^1(s')v^2(s) \geq v^2(s')v^1(s), \qquad \forall s' > s \ .$$

Proof. Given that our analytical results are comparisons of different levels of housing tightness, as this is an endogenous parameter, we would first like to know whether there is a monotonic relation between supply and tightness (since we will be observing supply and comparing different housing supplies). We must show that the elasticity of new buyers to new houses is smaller than one. The condition we must look for is

$$\frac{B_1 + \triangle B}{S_1 + \triangle S} < \frac{B_2 - \triangle B}{S_2}.$$

This always holds as long as $\frac{\triangle B}{\triangle S} < 1/2$.

To prove that this is true, notice that

$$B_1 = \int \left(v(s) \frac{[1 - V^1(s)]e^{-\frac{B_1}{S_1}(1-V^1(s))}}{\left[[1 - V^1(s)]e^{-\frac{B_1}{S_1}(1-V^1(s))} + [1 - V(s) + V^1(s)]e^{-\frac{(1-B_1)}{S_2}(1-V(s)+V^1(s))}\right]} \right) ds$$

After some algebraic manipulation, we can show that

$$\partial B_1 / \partial S_1 < 1.$$

This is always lower than 1, Thus, tightness is monotonic in housing supply. What about relative tightness? Given that extra (or less) supply has an effect on the tightness of both cities, the last thing we must check is whether the change in tightness in City 1 is larger than the change in tightness in City 2:

$$\frac{B_1}{S_1} - \frac{B_1 + \triangle B}{S_1 + \triangle S} > \frac{B_2}{S_2} - \frac{B_2 - \triangle B}{S_2} \Longleftrightarrow \frac{B_1 \triangle S - S_1 \triangle B}{S_1(S_1 + \triangle S)} > \frac{\triangle B}{S_2}.$$

This condition depends on the level of the original elasticities. For example, if City 2 originally had a extremely restricted housing supply, then the number of people that leave the city creates a very strong change in its local housing market. In general, the following condition must hold:

$$\frac{\triangle B}{\triangle S} < \frac{1}{2}.$$

Given this expression, this elasticity becomes (except for extreme values of S_2)

$$\vartheta \equiv \theta_1/\theta_2 = \frac{S_2}{S_1}\frac{B_1}{B_2} = \frac{S_2}{S_1}\int\left(\frac{[1-V^1(s)]e^{-\frac{B_1}{S_1}(1-V^1(s))}}{\left[[1-V(s)+V^1(s)]e^{-\frac{(1-B_1)}{S_2}(1-V(s)+V^1(s))}\right]}\right)ds$$
$$= \frac{S_2}{S_1}\int\left(\frac{[1-V^1(s)]e^{\frac{1}{\vartheta}\frac{(1-V^1(s))}{[1-V(s)+V^1(s)]}}}{[[1-V(s)+V^1(s)]]}\right)ds.$$

We now analyze how this relative tightness measure would change if, for example, the supply of housing in City 2 change. This have a mechanical effect in terms of relaxing the tightness of the housing market in City 2, but as people move from City 1 to city two, this have the same effect in City 1. The system of cities receives a “positive” shock; the question now is: In which city will the tightness condition relax more? If the monotonicity condition holds, we should expect that a higher supply of houses in City 2 makes the City 2 housing market less tight relative to City 1, thus relative tightness should increase:

$$\frac{\partial \vartheta}{\partial S_2} = \frac{1}{S_1}\int \big(\frac{[1-V^1(s)]e^{\frac{1}{\vartheta}\frac{(1-V^1(s))}{[1-V(s)+V^1(s)]}}}{[[1-V(s)+V^1(s)]]}\big)ds - \frac{S_2}{S_1}\frac{[1-V(s)+V^1(s)]}{[1-V^1(s)]}\frac{1}{\vartheta^2}\int \big(\frac{[1-V^1(s)]e^{\frac{1}{\vartheta}\frac{(1-V^1(s))}{[1-V(s)+V^1(s)]}}}{[[1-V(s)+V^1(s)]]}\big)ds\frac{\partial \vartheta}{\partial S_2}$$

$$\iff \frac{\partial \vartheta}{\partial S_2} = \frac{\vartheta}{S_2}\frac{1}{(1+\frac{S_2}{S_1}\frac{[1-V(s)+V^1(s)]}{[1-V^1(s)]}\frac{1}{\vartheta})} > 0$$

Finally, we must show that $v^1(s)/v^2(s)$ is an increasing function. Writting

$$\frac{v^1(s)}{v^2(s)} = \frac{(1-V^1(s))e^{-\theta^1(1-V^1(s))}}{(1-V^2(s))e^{-\theta^2(1-V^2(s))}} = \frac{1-V^1(s)}{1-V^2(s)}\exp\left(\frac{\theta^2}{\theta^1}\frac{(1-V^2(s))}{(1-V^1(s))}\right),$$

and defining $g(s) = \frac{1-V^1(s)}{1-V^2(s)}$, we obtain

$$v^1/v^2(s) = g(s)\exp\left(\frac{\theta^2}{\theta^1}\frac{1}{g(s)}\right).$$

Thus,

$$\frac{\partial(v^1/v^2)(s)}{\partial s} = g'(s)\exp\left(\frac{\theta^2}{\theta^1}\frac{1}{g(s)}\right)\left(1-\frac{\theta^2}{\theta^1}\frac{1}{g(s)}\right).$$

It remains to show that $[1-\frac{\theta^2}{\theta^1}\frac{1}{g(s)}] > 0 \iff \frac{\theta^2}{\theta^1}\frac{1}{g(s)} < 1 \iff \frac{\theta^2}{\theta^1} < g(s)$.

Since $\theta^1 > \theta^2$, then $\frac{\theta^2}{\theta^1} < 1$. Finally, we must show that $g(s)$ is an increasing function (since $g(0) = 1$), if $g(s)$ in an increasing function then $\frac{\theta^2}{\theta^1} < 1 \leq g(s)$ will hold for every $1 > s > 0$

Notice that

$$g'(s) = \frac{-v^1(s)(1-V^2(s)) + v^2(s)(1-V^1(s))}{(1-V^2(s))^2}$$

$$= \frac{v(s)g(s)}{[1-V^1(s)]e^{-\theta^1(1-V^1(s))}+[1-V^2(s)]e^{-\theta^2(1-V^2(s))}}[e^{-\theta^2(1-V^2(s))}-e^{-\theta^1(1-V^1(s))}]$$

Therefore, we must show again that

$$e^{-\theta^2(1-V^1(s))}-e^{-\theta^1(1-V^1(s))}>0 \Leftrightarrow \frac{\theta^2}{\theta^1}<1<\frac{(1-V^1(s))}{(1-V^2(s))}$$

This is equivalent to showing that

$$V^1(s)<V^2(s) \quad \forall s$$

If we can show that

$$v^1(s)<v^2(s) \quad \forall s$$

then the above will also hold.

Now, let us show that $v^1(0)<v^2(0) \Leftrightarrow v^1(0)-v^2(0)<0 \Leftrightarrow e^{-\theta^1}<e^{-\theta^2}$ and since $\theta^1>\theta^2$ this will be true.

Second, we must also show that

$$\partial v^1/\partial v^2 \leq 1$$

(The intuition is that if v^1 does not grow faster than v^2, then it will always be the case that $v^1(s)<v^2(s) \quad \forall s$)

Now, after some algebraic manipulation, we can show that $\partial v^1/\partial v^2=1$ Thus the result holds. □

Proof of Proposition 2.1. The proof follows immediately from lemma 3 of Costinot and Vogel

(2010). The argument is that, given that we have a system of ODE, for any pair of functions $A(\cdot)$ and $B(\cdot)$ that are continuous and of bounded variation, we can write $dM = F^M(M,w,\gamma)$ and $dw = F^w(M,w,\gamma)$. Given that F^M and F^w are continuous and of bounded variation, there exists a unique equilibrium for the system of ordinary differential equations. $\square$

B Housing Prices

We obtain the equilibrium bids which are representative of the equilibrium housing prices. We follow [Albrecht, Gautier and Vroman 2013] closely.

First, we write the expected bid as

$$\mathbb{E}[b(w(s)] = \frac{\int_r^{w(s)} b(w(s))h(w(s))dw(s)}{\int_r^{w(s)} h(w(s))dw(s)},$$

where $b(w(s))$ is the optimal bid in a first price auction and $h(w(s))$ is the density of the highest valuation drawn by the buyers visiting a particular seller, conditional on the seller having at least one visitor.

Assume that the reservation value will be the lowest wage available $w(\underline{s})$. Thus, we write

$$\mathbb{E}[b(w(s)] = \frac{\int_{w(\underline{s})}^{w(s)} b(w(s))h(w(s))dw(s)}{\int_{w(\underline{s})}^{w(s)} h(w(s))dw(s)} = \frac{\int_{\underline{s}}^{s} b(w(s))h(s)ds}{\int_{\underline{s}}^{s} h(s)ds}.$$

In addition, Bayes theorem implies that

$$h(s) = f(s|H) = \frac{P(H|s)v^i(s)}{P(H)}.$$

Given the properties of Poisson distributions, the probability that a buyer who has $w(s)$ has the highest valuation is

$$P(H|s) = e^{-\theta(1-V^i(s))},$$

and the unconditional probability that any buyer has the highest valuation becomes

$$P(H) = \int e^{-\theta(1-V^i(s)} v^i(s) ds = \frac{1-e^{-\theta}}{\theta}.$$

Thus,

$$h(s) = \frac{\theta e^{-\theta(1-V^i(s))} v^i(s)}{1-e^{-\theta}}.$$

Given the optimal bidding in a first price auction, it follows that

$$b(w(s)) = w(s) - \frac{V(s)}{v(s)}.$$

Basically, each bidder, who faces an uncertain number of buyers, offers her expected value minus its virtual surplus. Evaluating the expected bid, we have

$$\mathbb{E}[b(w(s)] = \frac{\int_{\underline{s}}^{s} b(w(s))h(s)ds}{\int_{\underline{s}}^{s} h(s)ds} = w(s) - \frac{\int_{\underline{s}}^{s} \frac{(V(s))}{v(s)} h(s)ds}{\int_{\underline{s}}^{s} h(s)ds},$$

which concludes the proof.

Finally, from Proposition 2, we have shown that the skill distribution in City 1 is more skill-abundant than in City 2 (i.e., the monotone likelihood ratio property holds). The rest of the proof for Proposition 3 directly follows from Costinot Voguel 2009 Lemma 3.

C Indirect Inference Estimation

We rely on the indirect inference estimation to recover the parameters for workers' taste for amenities and the technology shifter. The objective is to generate moments of the skill distribution and wage schedule that are as close as possible to the data.

Given that we solve the theoretical model in two separate parts (i.e., the spatial sorting of workers and the competitive equilibrium in each city), we split our estimation procedure into two steps. First, we solve the spatial sorting problem and recover the parameter capturing the taste for amenities T. Once we obtain an endogenous distribution of skills for each MSA, we solve for the distribution of wages in each city (because the cities produce non-tradable services). Second, from these wage distributions, we recover the second parameter capturing skill–technology complementarity A.

Last, it is important to point out that we aggregate the distribution of skills in the economy and then "sort" between cities. We set the domain of the aggregate skill distribution to be the interval $[0, 1]$ and assume that $v(s)$ follows a truncated normal, centered in 0.5 with dispersion equal to 1.

Recovering T from talent data

The first parameter we must recover is the taste for amenities parameter T, which shifts the utility of workers by e^{Ta^i}, determining worker's preferences for places. This parameter quantifies the tradeoff between housing tightness and the quality of services provided by a city (as well as other intangibles, such as weather or natural beauty). A higher taste for amenities means that workers tend to prefer living in that region, thus they are willing to risk going into a city, even when facing a tough housing market that could leave them living in the periphery.

Although we assume that the taste for amenities is the same for all workers, the tradeoff this entails is distinct for different type of workers. In particular, high-skilled workers who know they can get a house in the city center and still have higher disposable income tend to value the amenities of a city more.

Given that this parameter fundamentally speaks to how workers of different skills sort between cities, we recover it from empirical distributions of skills. We simulate the model for different parameter values of T such that the model's cross-sectional predictions are as close as possible to the data. We define a Wald-type loss function that weights the distance between the predicted and the observed mean and variance of the talent distribution for each MSA.

The indirect inference estimator is given by,

$$\widehat{T} = \text{argmin}_T L(T) = (\rho - \widehat{\rho}(T))' W (\rho - \widehat{\rho}(T))$$

where ρ are the data moments, $\widehat{\rho}(T)$ are the simulated moments, and W is a positive definite weighting matrix, that captures relative city size.

We use a grid search procedure and divide the estimation process into four consecutive steps:

Step 1: Partition the parameter space of T.[11]

Step 2: Run the model and compute auxiliary vector $\widehat{\rho}(T)$.

Step 3: Compute the criterion function $L(T)$.

Step 4: Repeat steps 1–3 to minimize $L(T)$.

Using this process, we find that the parameter that minimizes this distance is equal to 7.[12] Notice that the taste parameter requires two different kinds of exogenous data to uniquely

[11] We let T vary from -100 to 100. We took steps of size 0.1, to ran the program for 2000 possible values of T.
[12] Which is very consistent with the 7.4 that we find using the Lumosity sample.

predict the shape and size of the skill distribution: amenities a^i and housing S^i.

Recovering the technology parameter A

One important theoretical assumption is that high-skilled workers have a comparative advantage in performing more complex tasks, and that complementarity must be log-supermodular. Thus, we use the following simple parametric characterization for empirical purposes: $A(s,\sigma) = e^{As\sigma}$, where A is a technology shifter which is recovered from the data. The parameter A captures the advantage that higher-skilled workers might have over lower-skilled workers. This parameter does not affect how workers sort between cities, but it directly affects the level and dispersion of wages. We recover this parameter from the observed distributions of wages by city.

We follow the same strategy as in the first part of the calibration. Now, however, we can solve the model for wages for each city independently. We start by choosing a parameter value for both the agglomeration parameter and the elasticity of substitution. In addition, we take the distributions of skills by city as given from the problem above. Next, we simulate the model for different parameter values of A such that the model's predictions for the distribution of wages in each city are as close as possible to the data. As before, we construct a Wald-type of loss function as a weighted distance between predicted and observed moments for the distribution of wages in each MSA, and focus the attention on the first and second moments of each distribution. The indirect inference estimator becomes

$$\widehat{A} = \text{argmin}_A L(A) = (\rho - \widehat{\rho}(A))' W (\rho - \widehat{\rho}(A)),$$

where ρ are the data moments, $\widehat{\rho}(T)$ are the simulated moments, and W is a positive definite weighting matrix that captures relative city size.

Following the same procedure as in the first part, we find that the parameter that minimizes this distance is equal to 1.36.