Imperfect Mobility

Zhengyu Cai
Southwestern University of Finance and Economics
Institute of Development Studies
Chengdu, Sichuan 611130, China
and
Global Labor Organization (GLO)
Email: czy@swufe.edu.cn

Oct. 23, 2019

Abstract

This study investigates why the strong form of the spatial equilibrium is weakly supported in the literature. Using a discrete choice model, it shows that the strong form of the spatial equilibrium is rarely observed because workers are imperfectly mobile from the perspective of researchers. Incorporating the discrete choice model, a Markov chain is used to model the spatial dynamics of the population distribution. For a given location choice set, the population distribution is shown to converge to a unique spatial steady state. Microdata from the American Community Survey show that the model assumption is reasonable and support the model predictions.

Keywords: imperfect mobility, heterogeneity, spatial steady state, discrete choice model, Markov chain analysis
JEL Classification: J61, R10, C25, C15, C44

The author is grateful to Mian Huang, Dan S. Rickman, Huiyao Wang, John V. Winters, Houwen Wu, Shu Xu, and Jipeng Zhang for their inspiring and helpful suggestions on the early version of this paper. The author would also like to thank participants and discussants in several seminars and conferences for their insightful comments and suggestions. Financial support from the Fundamental Research Funds for the Central Universities (JBK1901007) is gratefully acknowledged. The usual disclaimer applies.

No two leaves are ever exactly alike.

—Gottfried Wilhelm Leibniz[1]

1. Introduction

The spatial equilibrium model, also referred to as the Rosen–Roback model, assumes workers are perfectly mobile (Roback 1982). At the spatial equilibrium, utility is equalized for workers everywhere.[2] Local attribute shocks/differentials are capitalized into factor prices. Migration is thus not expected at the spatial equilibrium.

A substantial number of empirical studies test the spatial equilibrium (Partridge et al. 2015; Goetzke and Islam 2017). Most such studies use U.S. data as labor in the United States is highly geographically mobile (Coen-Pirani 2010; Boman 2011; Saks and Wozniak 2011; Partridge et al. 2015). However, some only support the weak form of the spatial equilibrium (see Partridge et al. [2015] for a thorough survey), which has drawn increasing attention from researchers as many applications and policy considerations are built on the assumption of a strong-form spatial equilibrium, such as quality of life indices (Gabriel and Rosenthal 2004; Reynolds and Rohlin 2014), urban amenity capitalization (Cortés and Iturra 2019), urban planning (Acolin and Green 2017; Leibowicz 2017; Yang et al. 2019), and place-based policies (Glaeser and Gottlieb 2008; Kline and Moretti 2013). If the spatial equilibrium does not hold, quality of life indices are biased (Bayer, Keohane, and Timmins 2009; Krupka and Donaldson 2013) and place-based policies are not considered to be ineffective (Hall and Romer 2008; Partridge et al. 2015; Goetzke and Islam 2017).

This study uses a simple theoretical framework to explain why the strong form of the spatial equilibrium is rarely empirically observed. Few studies have thus far directly and systematically answered this question. Although mobility of current people is very

[1] This is not an exact quote by Leibniz. Yet, the connotation of this famous sentence originates from his philosophy.

[2] In addition, no interregional cost or profit differential exists for firms. This study only focuses on workers, assuming that the behaviors of firms are the same as in a conventional spatial equilibrium model.

high (Merino and Prats 2020), González-Val (2019) points out that the central city of each regional subsystem in the United States is far away from others, which does not facilitate long-distance migrations. Winters (2009) finds indirect evidence of the existence of a spatial equilibrium, stating that whether workers are fully compensated is an empirical exercise depending on which econometric methods and price measures are used. Greenwood et al. (1991) and Clark et al. (2003) argue that the spatial equilibrium is a long-run rather than a short-run phenomenon. Partridge et al. (2012) investigate whether recent decreasing net internal migration is a sign of a spatial equilibrium. Yet, they find that it is more likely to represent a structural shift in the local labor market.

As utility is unobservable, Winters and Li (2017) review the literature on unequal local subjective well-being (Oswald and Wu 2010, 2011; Glaeser, Gottlieb, and Ziv 2014, 2016). The first strand of the literature attributes subjective well-being differentials to the trade-off between subjective well-being and other factors not affecting it (Glaeser, Gottlieb, and Ziv 2014, 2016). The other strand links subjective well-being differentials to the barriers to migration that deter spatial arbitrage. For example, Krupka (2009) finds that mobility is deterred by location-specific human capital investment and Krupka and Donaldson (2013) argue that heterogeneous moving costs and land supply constraints bias conventional quality of life estimates.

Although few studies directly and systematically examine why the spatial equilibrium is rarely empirically supported, the findings in the literature imply that most frictions deterring the spatial equilibrium can be categorized into heterogeneity and moving costs. Roback (1988) argues that the utility levels of different types of workers are not necessarily equalized across locations at the spatial equilibrium if heterogeneous preferences exist. Equivalently, the spatial equilibrium condition only applies to marginal movers in a heterogeneity setting (Gyourko, Kahn, and Tracy 1999; Kahn 2006; Krupka and Donaldson 2013). However, as discussed by Krupka and Donaldson (2013, 847), “[h]eterogeneity has generally been dealt with in a footnote noting that in its presence, the results hold for the marginal migrant." Most empirical studies assume a spatial equilibrium under homogeneity without indicating whether their object of study is the marginal mover or how they identify marginal movers. Such studies also fail to explain whether and how the results based on marginal movers could be generalized to other individuals.

Some recent studies in urban economics and new economic geography explicitly consider heterogeneity. For example, Bayer et al. (2009) use a discrete choice model to integrate heterogeneity and moving costs, finding that moving costs bias the hedonic value estimates of local amenities. Behrens et al. (2017) propose a spatial equilibrium condition based on a discrete choice model considering heterogeneous preferences. The condition indicates that at the spatial equilibrium, the average probability of choosing each location is the same as its current population share. The Behrens equilibrium condition characterizes a collective and dynamic equilibrium rather than a static equilibrium at the individual level, meaning that it allows for unobserved heterogeneity and the idiosyncratic migration of individual workers at that equilibrium. More importantly, the Behrens equilibrium condition does not require the location choice probability to be equalized everywhere, implying that revealed utility is not necessarily identical across locations at the equilibrium. Although the Behrens equilibrium does not require the existence and uniqueness of a conventional spatial equilibrium, Behrens et al. (2017) do not prove the existence of their proposed spatial equilibrium.

Although discrete choice models are widely used to incorporate heterogeneity, few studies use them to explicitly discuss the relationship between heterogeneity and the conventional spatial equilibrium condition. Based on the foregoing, this study uses a discrete choice model to bridge the unobserved heterogeneity and imperfect mobility observed by researchers to explain why the strong form of the spatial equilibrium is rarely observed. Specifically, it shows that unobserved heterogeneity generates "immobility" and "over-mobility" from the perspective of researchers. It is intuitive that some people move or stay because of unknown idiosyncratic reasons. Besides, one advantage of discrete choice models is that moving costs can be directly incorporated (see Bayer et al. (2009) as an example). Yet, the existing literature does not formally distinguish psychological moving costs and pecuniary moving costs, which affect mobility differently. This study therefore argues that travel preference, as a source of unobserved heterogeneity, is a part of psychological moving costs, whereas pecuniary moving costs solely depend on moving distance. It is further shown that unobserved psychological moving costs generate the imperfect mobility observed by researchers and that pecuniary moving costs only shift imperfect mobility.

Region size is implicit in the Rosen–Roback model. Glaeser and Tobio (2008) solve for population endogenously in a homogeneous spatial equilibrium model. Krupka and Donaldson (2013) discuss population in a heterogeneous spatial model. More recently, Behrens et al. (2017) link the location choice probability to the population distribution; however, their model is static. By contrast, this study incorporates a discrete choice model with a finite time-homogeneous Markov chain to analyze the spatial dynamics of the population distribution. It is demonstrated that in a heterogeneous setting and under a reasonable assumption, a location choice set converges to a unique spatial steady state. The local population share is explicitly determined at this spatial steady state. The spatial equilibrium condition proposed by Behrens et al. (2017) is then shown to result from the spatial steady state derived from a special case of the Markov chain model. The model assumption implies that interregional revealed utility differentials (spatial disequilibrium) are persistent. The data also show positive evidence supporting the validity of the assumption.

The remainder of the paper is organized as follows. Section 2 demonstrates the duality and completeness between perfect mobility and the spatial equilibrium condition in a homogeneous setting as a benchmark. Section 3 introduces unobserved heterogeneity into the framework using a discrete choice model, showing the existence of imperfect mobility by comparing it with the benchmark presented in Section 2. Section 4 integrates pecuniary moving costs into the model and discusses its effect on mobility. Section 5 incorporates the discrete choice model with a finite Markov chain, demonstrating the existence of a spatial steady state. Section 6 provides six numerical examples in different scenarios to illustrate how the models work. Section 7 evaluates the suitability of the model assumption. The last section concludes.

2. Perfect Mobility and The Spatial Equilibrium Condition

In a standard Rosen–Roback-type model, a representative worker solves the utility maximization problem at a given location. The spatial equilibrium condition for workers can be stated as

(1) $$P = 0 \Leftrightarrow V(w, r|A_j) = v, \forall j$$

where P is the probability of any worker choosing a better location. w is wages and r is rents. A_j is amenities at Location $j = 1, \ldots, J$. v is a constant. Equation (1) indicates that at the spatial equilibrium, workers have no incentive to move, which requires that wages and rents adjust for the given amenities so that maximized utility is equalized everywhere. Otherwise, some workers move to locations in which they can be better off.

The conventional literature attributes Equation (1) to the assumption that workers are perfectly mobile; however, this is not sufficiently accurate. Define perfect mobility as follows.

Definition 1 (perfect mobility): Workers are perfectly mobile if

(2) $$V(w, r|A_i) \neq V\left(w, r\middle|A_j\right), \forall j \neq i \;\Leftrightarrow P = 1$$

This definition of perfect mobility indicates that some workers surely choose a better location when a utility differential exists. Indeed, perfect mobility is equivalent to the equalized utility condition rather than an assumption of the latter.

Proposition 1 (duality and completeness): In the conventional spatial equilibrium model setting, perfect mobility is equivalent to the spatial equilibrium (equalized utility) condition. They constitute a complete set of spatial equilibrium conditions.

Proof. Equation (1) and Equation (2) are equivalent as they are contrapositive for each other. They are complete as their probabilities are summed to one. ■

Proposition 1 indicates that perfect mobility per se is an aspect of a complete set of spatial equilibrium conditions rather than an assumption of the equalized utility condition. This complete set of spatial equilibrium conditions reflects the duality and completeness between perfect mobility and the equalized utility condition. Thus, equivalently, the equalized utility condition is an aspect of perfect mobility. This implies that the spatial equilibrium is not strictly supported in the literature because workers are not perfectly mobile.

However, people are inherently perfectly mobile because we change our behaviors immediately if and only if our *perceived* utility levels alter. Then, why is imperfect mobility defined and observed? An intuitive reason is that unobserved heterogeneity exists, meaning that *perceived* utility is not fully known to researchers. Researchers only know *revealed* utility at best. If a person does not move when his or her *revealed* utility at the current location is found to be lower than that in other places, he or she will be defined as imperfectly mobile; however, his or her stay could be due to an unobserved factor that keeps his or her *perceived* utility level higher than or equal to the maintained level. Thus, imperfect mobility is only defined from the perspective of researchers, while it is not a characteristic of an individual.

We do not distinguish *revealed* utility from *perceived* utility in Equations (1) and (2) as the Rosen–Roback model assumes people are homogeneous. In this homogeneous setting, Equations (1) and (2) imply that if the maintained utility level of the spatial equilibrium is v, the probability of a perfectly mobile worker choosing Location i is

(3) $$P_i = \begin{cases} 0, & if\ V(w, r|A_i) \leq v \\ 1, & if\ V(w, r|A_i) > v \end{cases}$$

As shown in Figure 1, Equation (3) indicates that a perfectly mobile worker would not choose a location offering the same or lower utility level; he or she would surely choose a better location. However, researchers rarely face homogeneity. In the next section, heterogeneity is thus introduced into the framework.

3. Unobserved Heterogeneity

While people are heterogeneous, a representative economic agent is often used for modeling as we are thought to behave alike in some certain circumstances. This belief implies that different individuals have a "common core." Thus, all individuals are assumed to be the same after their finite heterogeneity is peeled off. Although this insight provides us with a simplified approach for theoretical modeling, it clearly indicates that individual heterogeneity should be thoroughly controlled for if representative-based theoretical models are empirically tested regardless of whether they are in a reduced form or a structural form. Unfortunately, it is difficult to eliminate heterogeneity, even if it is finite, as unobserved factors always exist.

Amenities vary across locations. Although a careful researcher could control for a large set of amenities, such a set is never exhaustive. Additionally, amenities are valuable only if people can perceive them. Otherwise, they do not affect the behaviors of people. Thus, amenities are not only different by themselves, but also perceived in various ways. The interaction between worker heterogeneity and amenity heterogeneity implies that if we allow heterogeneity to enter the model, the utility of worker n at Location j is

$$(4) \qquad V_n(w, r|A_j) = V\left(w, r\middle|A_j, X_n\right) + \varepsilon_{nj}, \quad j = 1, \dots, J; n = 1, \dots, N$$

where n indexes individual workers. X_n denotes observed individual heterogeneity such as measurable demographic attributes. ε_{nj} denotes unobserved utility, which depends on unobserved heterogeneity. The two-dimensional error term can be interpreted as the idiosyncratic preference of worker n for Location j. $V_n(\cdot)$ is the perceived utility of worker n, which is only known to him- or herself but not to researchers. Equation (4) indicates that perceived utility is decomposed into two parts. The first part, $V(\cdot)$, is revealed or representative utility, which is observable by researchers. Unobserved utility is assumed to be random. For simplicity, $V(w, r|A_j, X_n)$ is assumed to be well specified, and each ε_{nj} is an independently and identically distributed type I extreme value.[3] In the random utility setting in Equation (4), a discrete choice model shows that imperfect mobility is observed by researchers when unobserved heterogeneity exists.

Proposition 2 (imperfect mobility): A worker with unobserved heterogeneity is imperfectly mobile from the perspective of researchers.

Proof. Following Equation (4), the observed probability of worker n choosing Location i is

$$P_{ni} = Prob\big(V_n(w, r|A_i) > V_n(w, r|A_j), \forall j \neq i\big) \tag{5}$$

$$= Prob\big(V(w, r|A_i, X_n) + \varepsilon_{ni} > V(w, r|A_j, X_n) + \varepsilon_{nj}, \forall j \neq i\big)$$

$$= \frac{e^{V(w, r|A_i, X_n)}}{\sum_J e^{V(w, r|A_j, X_n)}}$$

Assume that the maintained utility level of the spatial equilibrium is v; then,

$$P_{ni} \in \begin{cases} \left(0, \frac{1}{J}\right], & if\ V(w, r|A_i, X_n) \leq v \\ \left(\frac{1}{J}, 1\right), & if\ V(w, r|A_i, X_n) > v \end{cases} \tag{6}$$

where J is the number of the locations in a given choice set. P_{ni} never takes the values of zero or one as it requires $V(w, r|A_i, X_n)$ to approach an infinitely small or large value, which is not possible. Therefore, compared with Equation (3), worker n is not perfectly mobile from the perspective of researchers. ■

Figure 2 illustrates Equations (5) and (6). As a benchmark, the perfect mobility case

[3] To have a closed-form result, we assume that this is a distributed type I extreme value, which is not empirically different from the normal distribution (Train 2009). Generally, one can specify any other distributions for ε_{nj}, which would not change the results qualitatively.

shown in Figure 1 is also integrated. Figure 2 shows that P_{ni} takes a logit form, which is bounded between zero and one. The observed probability of choosing Location i continuously grows with the increase in revealed utility at Location i (the solid curve), which deviates from the perfect mobility case (the dashed lines) in that there exists a jump at the maintained utility level of the spatial equilibrium. Thus, considering unobserved heterogeneity, the location decision observed by researchers is no longer a yes-or-no question under homogeneity. Now, imperfect mobility can be defined by researchers. Figure 2 shows that two types of imperfect mobility exist from the perspective of researchers: "immobility" and "over-mobility."

Definition 2 ("immobility"): Given that any revealed utility at Location i is larger than the maintained utility level of the spatial equilibrium ($V_i > v$), "immobility" is defined as $1 - P_{ni}(V_i)$.

When the revealed utility at Location i is above the maintained level, researchers expect the probability of choosing Location i to be one. A probability of choosing the more desirable location below one from the perspective of researchers generates "immobility," which is presented in the dotted area bounded between the dashed line and solid curve in Figure 2: for example, if $V(w, r|A_i, X_n) = V_1$, then "immobility" regarding Location i is $1 - P_{ni}(V_1)$. Proposition 2 demonstrates that "immobility" results from unobserved heterogeneity. Because researchers can only observe revealed utility, they expect workers to choose the location offering higher revealed utility; however, workers make decisions based on perceived utility rather than revealed utility. Hence, some amenities at the location offering lower revealed utility may only be perceived by workers, but not observed by researchers. If the utility generated by these amenities plus revealed utility is larger than or equal to the maintained utility level, workers are willing to choose this location even though it offers lower revealed utility. Consider the example of a location choice set with two options: Location a and

Location b. If researchers find that Location a offers higher utility than Location b, they would expect that people will surely choose Location a. However, researchers observe that fewer than the expected number of people choose Location a as some choose Location b because of an unobserved idiosyncratic preference for amenities that can only be or more easily accessed at Location b such as family ties (Costa and Kahn 2000; Compton and Pollak 2007; Mulder and Malmberg 2014), local social networks (Spilimbergo and Ubeda 2004; David, Janiak, and Wasmer 2010), and local food.[4] Thus, researchers would believe that workers are "immobile" regarding Location a.

Definition 3 ("over-mobility"): Given that any revealed utility at Location i is less than or equal to the maintained utility level of the spatial equilibrium ($V_i \leq v$), "over-mobility" is defined as $P_{ni}(V_i)$.

When the revealed utility at Location i is less than or equal to the maintained level, researchers expect the probability of choosing Location i to be zero. A positive probability of choosing the undesired or indifferent location from the perspective of researchers defines "over-mobility," which is represented by the striped area bounded between the solid curve and dashed line in Figure 2. For example, if $V(w, r|A_i, X_n) = V_2$, then "over-mobility" regarding Location i is $P_{ni}(V_2)$. Similarly, "over-mobility" also results from unobserved heterogeneity. Researchers expect that workers never choose a less desired or indifferent location based on revealed utility. Nonetheless, some workers choose a less desired location because their perceived utility (i.e., their unobserved idiosyncratic preference for the location plus revealed utility) is larger than the maintained utility level. In the two-location example above, researchers would expect workers to surely choose Location a and no one to choose Location b. However, as mentioned, they observe fewer than the expected number of people

[4] Even if these can be observed and potentially measurable, incomplete fungibility also generates idiosyncratic preferences.

choosing Location a and more than the expected number choosing Location b owing to their unobserved preference for amenities that can only be or more easily accessed at Location b. In this case, researchers find that workers are "over-mobile" regarding Location b.

Unlike the prediction of a conventional spatial equilibrium model, our model shows that positive incentives to choose a specific location exist at the spatial equilibrium if unobserved heterogeneity is considered. At the spatial equilibrium, representative utility is equalized everywhere, that is,

$$V(w, r | A_j, X_n) = v, \forall j \tag{7}$$

Substituting Equation (7) into Equation (5), the probability of choosing Location i becomes

$$P_{ni} = \frac{1}{J} \tag{8}$$

where J is the total number of the locations in a given choice set. Equation (8) shows that there still exists a positive probability of choosing a specific location at the spatial equilibrium, which contrasts with the homogeneous case shown in Equation (3) that the probability of choosing a specific location is zero at the spatial equilibrium. Thus, researchers will observe that workers are "over-mobile" even though the spatial equilibrium is thought to have been achieved.

4. Moving Costs

Moving costs are often treated as another friction deterring the spatial equilibrium (Bayer, Keohane, and Timmins 2009; Krupka and Donaldson 2013; Diamond 2016). To more precisely discuss the effects of moving costs on the spatial equilibrium from the perspective of researchers, the definition of moving costs should be specified.

Definition 4 (moving costs): Moving costs are the full set of factors deterring the moving decision of a worker. They consist of psychological moving costs and pecuniary moving costs. Psychological moving costs come from heterogeneity. Pecuniary moving costs equal moving distance.

This is a generalized definition of moving costs. Table 1 shows why and how we have this definition.[5] Psychological moving costs are related to the preferences of workers and themselves. As discussed in the previous section, workers are heterogeneous. Observable heterogeneity affects the moving decisions of workers related to such factors as age, health status, and education (Vigdor 2002; Machin, Salvanes, and Pelkonen 2012). Thus, observable heterogeneity is a part of psychological moving costs. As these factors can be directly controlled for, they are not further discussed in this study. As mentioned in the previous section, unobserved heterogeneity also influences the moving decisions of workers related to such factors as family ties, social networks, and local amenities. Thus, unobserved heterogeneity is another part of psychological moving costs.[6]

Now, the remaining factors are all about moving or traveling per se, which are usually regarded as moving costs in the narrow sense. These factors can be decomposed into travel preferences and moving distance. Travel preferences such as the perception of distance, vehicle preference, airline preference if traveling by air, and choice of route if traveling by automobile are mostly unobserved by researchers. Thus, travel preferences are sources of unobserved heterogeneity, with the other unobserved heterogeneity mentioned above generating imperfect mobility from the perspective of researchers, as discussed in the previous section.

If travel preferences and other heterogeneity, namely psychological moving costs,

[5] Carrington, Detragiache, and Vishwanath (1996) discuss endogenous moving costs, which decrease with the number of migrants already settled in the destination. For simplicity, endogeneity is assumed to be ignorable. As discussed later, moving costs do not qualitatively influence the existence of imperfect mobility. Endogenous moving costs are less likely to qualitatively alter the results. Further, as shown later, the number of migrants at each location in the choice set are stable at the spatial steady state. Therefore, at least, such endogeneity in moving costs does not have temporal variation.

[6] Unobserved psychological moving benefits may exist as some unobserved heterogeneity causes "over-mobility," as discussed in the previous section.

can be completely controlled for and travel cost per unit distance is normalized to one, the only remaining variation in moving costs is moving distance. Thus, moving distance is equivalent to pecuniary moving costs.[7] The remainder of this section discusses how pecuniary moving costs affect mobility from the perspective of researchers.

Define τ_{ki} as the moving distance from Location k to Location i. Hence,

$$\tau_{ki} = \begin{cases} \tau_{ik}, \forall i \neq k \\ 0 \quad, \forall i = k \end{cases} \tag{9}$$

Now, the observed probability of worker n choosing Location i is

$$P_{ni|k} = \frac{e^{V(w,r|A_i,X_n)+T(\tau_{ki})}}{\sum_J e^{V(w,r|A_j,X_n)+T(\tau_{kj})}} \tag{10}$$

where k is the initial location. $T(\cdot)$ is a monotonic disutility function on moving distance: $T(0) = 0$, and $T'(\cdot) < 0$.

Proposition 3 (biasness): If pecuniary moving costs are not accurately counted by researchers, imperfect mobility is observed with bias.

Proof. For $J > 2$, given the moving distance to the alternative locations, the partial derivative of Equation (10) with respect to τ_{ki} is, $\forall i \neq k$,

$$\frac{\partial P_{ni|k}}{\partial \tau_{ki}} = T'(\tau_{ki}) e^{V(w,r|A_i,X_n)+T(\tau_{ki})} \frac{\sum_{J \neq i} e^{V(w,r|A_j,X_n)+T(\tau_{kj})}}{\left(\sum_J e^{V(w,r|A_j,X_n)+T(\tau_{kj})}\right)^2} < 0 \tag{11}$$

Thus, if the moving distance to Location i is ignored or undercounted by researchers, "over-mobility" is overstated, while "immobility" is understated by researchers. Similarly, for $J > 2$, given the moving distance to Location i, according to Equation (10), $\forall j \neq k$,

$$\frac{\partial P_{ni|k}}{\partial \tau_{kj}} = -T'(\tau_{kj}) \frac{e^{V(w,r|A_i,X_n)+T(\tau_{ki})+V(w,r|A_j,X_n)+T(\tau_{kj})}}{\left(\sum_J e^{V(w,r|A_j,X_n)+T(\tau_{kj})}\right)^2} > 0 \tag{12}$$

Thus, if any moving distances to the alternative locations are ignored or undercounted

[7] These two terms are used interchangeably hereafter.

by researchers, “over-mobility” is understated, while “immobility” is overstated.

For $J = 2$, the partial derivative in Equation (12) is undefined as the distance between two locations is constant regardless of the moving direction. Equation (11) still works in this case, implying that “over-mobility” is overstated, while “immobility” is understated if the moving distance between two locations is ignored or undercounted.

■

Inequation (11) shows that given any revealed utility level at Location i, the longer the moving distance to Location i, the lower the probability of choosing Location i, which is shown in Figure 3(a). Figure 3(a) illustrates that the height of the “over-mobility” area shrinks, while the height of the “immobility” area increases if pecuniary moving costs are considered. Thus, if the moving distance to a target location is ignored or undercounted by researchers, “over-mobility” is overstated and “immobility” is understated.

Inequation (12) indicates that for any revealed utility level at Location i, the larger the moving distance to any alternative location, the higher the probability of choosing Location i. Figure 3(b) illustrates this and implies that “over-mobility” is understated, while “immobility” is overstated if the moving distance to any alternative location is ignored or undercounted by researchers.

Proposition 3 also implies that imperfect mobility always exists from the perspective of researchers no matter whether pecuniary moving costs are accurately counted. As moving distance is easier to directly control for than unobserved heterogeneity, unobserved heterogeneity rather than pecuniary moving costs is more likely to deter researchers from observing the spatial equilibrium.

5. Spatial Steady State

Propositions 2 and 3 indicate that it is difficult for researchers to observe the spatial equilibrium because of unobserved heterogeneity. Idiosyncratic migration can be observed even if revealed utility is equalized across locations. Yet, if revealed utility differentials exist, will they disappear over time in a heterogeneous setting? In a homogeneous spatial equilibrium model setting as in Proposition 1, local revealed utility converges through spatial arbitrage, implying that interregional revealed utility differentials are capitalized into factor prices pushed by interregional migration, which affects region size. However, Rosen–Roback models do not explicitly discuss equilibrium region size.

In this section, considering unobserved heterogeneity, it is shown that if location choice probabilities are assumed to be constant over time (i.e., interregional revealed utility differentials, if any, are persistent), region sizes converge to a unique distribution (i.e., a spatial steady state exists for a given location choice set).

Definition 5 (spatial steady state): Given a location choice set, in a spatial steady state, the population distribution is stationary.

Consider a finite location choice set with N locations. The distribution vector of the population at time t is $\boldsymbol{s}_t = (s_{1t} \quad \cdots \quad s_{Nt})$, where s_{it} is the local population share.[8] The average probability of choosing Location i from Location k at time t is given by an averaged version of Equation (10):

$$P_{i|k}^{t} = \frac{e^{V\left(w(S_{it}), r(S_{it}) \middle| A_i, \bar{X}_n\right) + T(\tau_{ki})}}{\sum_{j=1}^{N} e^{V\left(w\left(S_{jt}\right), r\left(S_{jt}\right) \middle| A_j, \bar{X}_n\right) + T(\tau_{kj})}} \tag{13}$$

where $i = 1, \ldots, N$ and $k = 1, \ldots, N$. S_{it} is the local population. It is convenient to regard the evolution of the population distribution, $\boldsymbol{s}_t$, as a finite Markov chain process. Then, Equation (13) is the expression of the entries of the transition matrix. Equation (13) indicates that the location choice probability is affected by the local population through its effect on factor prices and hence, revealed utility. The literature on urban economics describes agglomeration externalities as a determinant of region size (Abdel-Rahman and Anas 2004). Agglomeration brings about benefits, while negative

[8] Natural population growth is not considered in this study.

externalities such as pollution, congestion, and crime lower revealed utility. However, because the non-linearity of agglomeration lacks consensus in the literature (Glaeser and Gottlieb 2008), it is difficult to determine the functional forms of $w(S_{it})$ and $r(S_{it})$. Further, Equation (13) indicates that the finite Markov chain is time-dependent, which loses some well-behaved properties compared with a time-homogeneous one. To simplify the analysis and draw analytical solutions, the following assumption is therefore made.

Assumption 1: The population elasticity of revealed utility is sufficiently small to be ignored, such that Equation (13) becomes time-independent:

$$P_{i|k} = \frac{e^{V\left(w,r\middle|A_i,\bar{X}_n\right)+T(\tau_{ki})}}{\sum_{j=1}^{N} e^{V\left(w,r\middle|A_j,\bar{X}_n\right)+T(\tau_{kj})}} \tag{14}$$

Assumption 1 is a strong assumption, ignoring the effect of the population on revealed utility. Ideally, the marginal benefit of making this assumption would be larger than the marginal cost, so that the assumption simplifies the model without substantially affecting its performance. First, this assumption implies that the spatial disequilibrium is persistent, which is supported by the literature (Spilimbergo and Ubeda 2004; Partridge et al. 2015; Goetzke and Islam 2017). Second, it avoids the functional-form concern of $w(S_{it})$ and $r(S_{it})$. Third, it is evaluated using real data in Section 7, where the results show that the assumption is reasonable. Fourth, it only matters during convergence. Revealed utility is constant after the process has converged (in a spatial steady state) as the population is constant. This study focuses on what happens in the spatial steady state rather than convergence. Fifth, it is possible for a time-varying version of the Markov chain above to converge to a stationary distribution using speculated functional forms of $w(S_{it})$ and $r(S_{it})$ under a weaker version of

Assumption 1 that the population elasticity of revealed utility is small but not ignored.[9] Finally, although stationarity is not guaranteed for a time-varying Markov chain without Assumption 1, it does not provide any evidence to show that local revealed utility converges to a single level.

Proposition 4 (convergence): Under Assumption 1 and if there is no subsequent shock, a given location choice set with unobserved heterogeneity converges to a unique spatial steady state.

Proof. The transition matrix is given by

$$\overline{\boldsymbol{P}} = \begin{bmatrix} P_{1|1} & \cdots & P_{N|1} \\ \vdots & \ddots & \vdots \\ P_{1|N} & \cdots & P_{N|N} \end{bmatrix} \tag{15}$$

Each transition matrix entry is given by Equation (14) under Assumption 1. As $0 < P_{i|k} < 1$ for all $i, k \in \{1, \dots, N\}$, the process is ergodic. Therefore,

$$\lim_{t \to \infty} \boldsymbol{s}_t = \boldsymbol{s} \tag{16}$$

where $\boldsymbol{s}$ is a unique stationary population distribution vector such that $\boldsymbol{s}\overline{\boldsymbol{P}} = \boldsymbol{s}$. ■

As the Markov chain is ergodic, the steady-state population distribution does not depend on the initial distribution. For simplicity, take the discrete time case with a two-location choice set (Locations a and b) as an example. Given the initial population of Location a is a_0 and of Location b is b_0, the total population is $S = a_0 + b_0$. The probabilities of choosing Locations a and b from Location a are $P_{a|a}$ and $P_{b|a}$, respectively. The probabilities of choosing Locations a and b from Location b are $P_{a|b}$ and $P_{b|b}$, respectively. Thus, the relationships between the current and future populations of Locations a and b, respectively are

$$a_{t+1} = P_{a|a} a_t + P_{a|b} b_t \tag{17}$$

$$b_{t+1} = P_{b|a} a_t + P_{b|b} b_t \tag{18}$$

[9] See Haag and Weidlich (1984) for a formal discussion in a reduced-form setting.

Solve for a_t and b_t:

$$(19) \qquad a_t = \frac{P_{b|a}\left(P_{a|a} + P_{b|b} - 1\right)^t + P_{a|b}}{P_{a|b} + P_{b|a}} a_0 + \frac{-P_{a|b}\left(P_{a|a} + P_{b|b} - 1\right)^t + P_{a|b}}{P_{a|b} + P_{b|a}} b_0$$

$$(20) \qquad b_t = \frac{-P_{b|a}\left(P_{a|a} + P_{b|b} - 1\right)^t + P_{b|a}}{P_{a|b} + P_{b|a}} a_0 + \frac{P_{a|b}\left(P_{a|a} + P_{b|b} - 1\right)^t + P_{b|a}}{P_{a|b} + P_{b|a}} b_0$$

As $\left|P_{a|a} + P_{b|b} - 1\right| < 1$,

$$(21) \qquad \lim_{t\to\infty} a_t = \frac{P_{a|b}}{P_{a|b} + P_{b|a}} S$$

$$(22) \qquad \lim_{t\to\infty} b_t = \frac{P_{b|a}}{P_{a|b} + P_{b|a}} S$$

According to Equations (21) and (22), the population of Locations a and b both converge to a constant that does not depend on the initial local population; rather, it only depends on the total population and a steady state share. The steady state share for Location a is $\frac{P_{a|b}}{P_{a|b}+P_{b|a}}$, which is intuitive as the population at Location a is fixed when the population choosing Location a from b is the same as that choosing Location b from a. The same intuition applies to Location b.

The further the distance from one location to the other locations, the closer to one is the probability of staying at the location. This is close to a Markov chain with an absorbing state at which the steady-state distribution depends on the initial distribution (multiple spatial steady states). However, this is less likely to happen in the setting used in this study. If a location is completely isolated, it is not reasonable to include it in a location choice set. Therefore, in a given location choice set, if some locations are remote from the other locations, the speed of convergence to a unique spatial steady state just becomes slow rather than having multiple spatial steady states. On the contrary, it is easy to see that the Markov chain converges immediately if pecuniary moving costs are ignored, as each row entry of the transition matrix is identical to that in any other row. Thus, the transition matrix is identical to any power of itself, implying that the lower pecuniary moving costs, the faster the speed of convergence.

Corollary (existence of the Behrens equilibrium condition): Without considering pecuniary moving costs, the steady-state population share of each location is the same as the probability of choosing the location.

Proof. Ignoring pecuniary moving costs, the transition matrix becomes

$$\overline{\boldsymbol{P}} = \begin{bmatrix} P_1 & \cdots & P_N \\ \vdots & \ddots & \vdots \\ P_1 & \cdots & P_N \end{bmatrix} \tag{23}$$

Suppose the spatial steady-state population distribution is

$$\boldsymbol{s} = (s_1 \quad \cdots \quad s_N) \tag{24}$$

Then,

$$\begin{aligned} \boldsymbol{s}\overline{\boldsymbol{P}} &= (s_1 \quad \cdots \quad s_N) \begin{bmatrix} P_1 & \cdots & P_N \\ \vdots & \ddots & \vdots \\ P_1 & \cdots & P_N \end{bmatrix} \\ &= (P_1 \quad \cdots \quad P_N) \end{aligned} \tag{25}$$

According to Proposition 4, in the spatial steady state,

$$\boldsymbol{s}\overline{\boldsymbol{P}} = \boldsymbol{s} \tag{26}$$

Therefore,

$$(P_1 \quad \cdots \quad P_N) = (s_1 \quad \cdots \quad s_N) \tag{27}$$

Namely,

$$P_i = s_i,\ \forall i \tag{28}$$ ■

Equation (28) is simply the spatial equilibrium condition given by Behrens et al. (2017).[10] Thus, the Behrens equilibrium condition is proven to exist as a result of the spatial steady state derived from a special case of our model ignoring pecuniary moving costs.

To sum up, Proposition 4 indicates that the population shares of a given location

[10] See Equations (29) and (30) in Behrens et al. (2017).

choice set converge to a unique distribution in a spatial steady state and are not necessarily the same across locations, as the probability of choosing each location is positive, persistent, and not necessarily equalized. Migrations must be observed if the location choice probability is positive. Consequently, researchers may observe a spatial steady state, while the spatial equilibrium is not observed.

6. Numerical Example

Numerical examples in six scenarios are provided in this section to illustrate the theoretical framework. For simplicity, consider a choice set containing Locations a and b.[11] Table 2 reports the numerical results.

Scenario 1 (benchmark): At the spatial equilibrium without unobserved heterogeneity.

Column (1) of Table 2 shows the results of a spatial equilibrium in the Rosen–Roback setting. Assuming revealed utility is equalized in both locations without pecuniary moving costs, workers have no incentive to shift locations. The probability of moving from Location b to a ($P_{a|b}$) is zero and this is the same for the reverse direction ($P_{b|a}$). In this case, workers are perfectly mobile regarding either location.

Scenario 2: Equalized revealed utility with unobserved heterogeneity.

To show the difference between the models with and without unobserved heterogeneity, Column (2) reports the results. Although revealed utility is still equalized in both locations as in Scenario 1, considering unobserved heterogeneity, $P_{a|b}$ and $P_{b|a}$ calculated by Equation (5) are both 0.5 rather than zero. Hence, interregional migration still exists even if revealed utility is equalized everywhere. Specifically, half of the population in each location swap their locations. From the perspective of researchers who cannot observe idiosyncratic heterogeneity, workers are "over-mobile" regarding either location.

[11] This can be generalized to a choice set with multiple options; however, the results are similar.

Scenario 3: Equalized revealed utility with unobserved heterogeneity and pecuniary moving costs.

Assume that the disutility of pecuniary moving costs is -0.5 for either direction of moving, as the distance between Locations a and b is the same regardless of the direction. Column (3) shows that the probability of moving from Location b to a ($P_{a|b}$) is 0.378 and this is the same for the reverse direction ($P_{b|a}$), as calculated by Equation (10). As revealed utility is still identical for each location, researchers who do not observe idiosyncratic heterogeneity would not expect interregional migration. However, unobserved heterogeneity makes workers choose either location with a positive probability. According to Definition 3, workers are "over-mobile" regarding either location from the perspective of researchers. Further, compared with Scenario 2, this case also considers the role of pecuniary moving costs. $P_{a|b}$ and $P_{b|a}$ are smaller than their corresponding values in Scenario 2, as pecuniary moving costs offset the idiosyncratic utility gain from moving. This case also illustrates Proposition 3 that "over-mobility" is overstated if pecuniary moving costs are not considered.

Scenario 4 (benchmark): Positive revealed utility shock (differential) at Location a without unobserved heterogeneity.

In this scenario, assume that the revealed utility offered at Location a is higher than that at Location b. Column (4) reports the benchmark results in a homogeneous setting. All homogeneous workers choose Location a as they can obtain higher utility[12] there, while none choose Location b. Workers are perfectly mobile regarding either location.

Scenario 5: Positive revealed utility shock (differential) at Location a with unobserved heterogeneity.

Owing to unobserved heterogeneity, the probability of moving from Location b to a is not one. Column (5) shows that this probability is 0.731, indicating that 26.9% of workers at Location a choose Location b even though it offers lower revealed utility.

[12] This is either revealed or perceived utility as they are identical for homogeneous workers.

Thus, compared with the benchmark in Scenario 4, workers are "immobile" regarding Location a and "over-mobile" regarding Location b from the perspective of researchers. According to Definitions 2 and 3, "immobility" regarding Location a and "over-mobility" regarding Location b are both 0.269 because with zero pecuniary moving costs, regardless of the initial location, the probability of workers choosing Location b is 0.269. Hence, 26.9% of workers move from Location a to Location b, resulting in "over-mobility" regarding Location b; moreover, 26.9% of workers from Location b choose to stay at Location b, generating "immobility" regarding Location a from the perspective of researchers.

Scenario 6: Positive revealed utility shock (differential) at Location a with unobserved heterogeneity and pecuniary moving costs.

Column (6) shows that the probabilities of moving from Location b to a and from Location a to b are 0.622 and 0.182, respectively. As discussed in Scenario 5, unobserved heterogeneity makes researchers observe "immobility" and "over-mobility" regarding Locations a and b, respectively. Such "immobility" and "over-mobility" are 0.378 and 0.182 according to Definitions 2 and 3, respectively. Compared with Scenario 5, "immobility" is larger and "over-mobility" is smaller in this scenario because pecuniary moving costs serve as a friction that reinforces "immobility" and offsets "over-mobility," consistent with the prediction in Proposition 3 that "immobility" and "over-mobility" are respectively understated and overstated when ignoring pecuniary moving costs in a two-location case.

These six hypothetical scenarios illustrate that unobserved heterogeneity generates imperfect mobility from the perspective of researchers, resulting in some difficulty detecting the spatial equilibrium. Although ignoring pecuniary moving costs both understates and overstates "immobility" and "over-mobility," it does not deny the existence of imperfect mobility from the perspective of researchers. To show the evolution of the six scenarios over time, the corresponding Markov chain processes are simulated.

As discussed in Section 5, as the Markov chain for the population distribution with unobserved heterogeneity is ergodic, the steady-state distribution is unique and does not depend on its initial distribution. Thus, the initial distribution choice could be arbitrary. In the two-location example, assume that the initial populations at Locations a and b are equally distributed.[13] Figure 4 reports the corresponding simulation results for Table 2.

Figure 4(1) shows the results for Column (1) of Table 2. This is a benchmark scenario in which the spatial equilibrium is achieved in a homogeneous setting. As revealed utility is equalized everywhere, homogeneous workers have no incentive to shift their locations. Thus, no interregional migration is observed. The population distribution is persistent (i.e., in a spatial steady state). Compared with Figure 4(1), unobserved heterogeneity is considered in Figure 4(2). The difference is clear: although revealed utility is still equalized, migration is always observed and persistent because of idiosyncratic preferences. Pecuniary moving costs are considered in Figure 4(3). This figure shows a similar pattern except that migrants in both directions and total migrants are less than those in Figure 4(2). In general, these figures illustrate the possible reason for the difficulty of observing the spatial equilibrium: the persistent existence of migrants would lead researchers to observe "over-mobility" and reject the strong form of the spatial equilibrium even if revealed utility is equalized across locations.

Figures 4(4)–4(6) illustrate the circumstances that Location a provides higher revealed utility than Location b according to Columns (4)–(6) in Table 2. Figure 4(4) shows the benchmark case in which workers are homogeneous. In this setting, all homogeneous workers move to Location a at time 1, reaching a new spatial equilibrium. As there are no subsequent shocks, after migration at time 1, no migration is observed, as in Scenario 1. However, considering unobserved heterogeneity, Figure 4(5) shows that not all workers move to Location a even though they can obtain higher revealed utility there. Some workers at Location a even move to Location b for

[13] Figure A1 in the Appendix shows an example of an unequal initial population distribution.

idiosyncratic reasons. These unexpected migrations observed by researchers represent imperfect mobility. As there are no pecuniary moving costs, the revealed utility difference can be realized on the population distribution immediately at time 1. The population distribution then converges to a stationary distribution. Moreover, another difference between Figures 4(5) and 4(4) is that migration still exists after the population distribution becomes steady because of unobserved heterogeneity. Considering pecuniary moving costs, the evolution pattern of the population distribution in Figure 4(6) is similar to that in Figure 4(5) except that migration in either direction and total migration are less because of the friction and the speed of convergence becomes slower, as discussed in the previous section. Figures 4(5) and 4(6) are consistent with the empirical findings of Clark et al. (2003) and Goetzke and Islam (2017) that workers move to locations in which they are better off. However, unlike that claimed by Clark et al. (2003) and Goetzke and Islam (2017), the figures show that such migration does not tend toward the spatial equilibrium. Instead, Goetzke and Islam (2017) actually provide evidence supporting Figures 4(5) and 4(6) that workers move to better locations, but that a disequilibrium is persistent. Spilimbergo and Ubeda (2004) also show that labor immobility is persistent.

The homogeneous cases (Figures 4(1) and 4(4)) are not ergodic Markov chain processes as the transition probabilities are either zero or one. Therefore, the population distributions in these cases either depend on the initial distribution (Figure 4(1)) or vanish (Figure 4(4)), as observed by comparing Figure 4 with Figure A1 in the Appendix. Moreover, the results in Figure 4 are based on Assumption 1 that population change does not affect revealed utility. This is a somewhat strong assumption. In the next section, data are used to evaluate its validity.

7. Validity of The Model Assumption

Assumption 1 suggests that the location choice probability is rarely affected by revealed utility through population change, meaning that the transition probabilities are time-independent. Under this assumption, it is proven that a given location choice set converges to a unique spatial steady state. This section evaluates the suitability of Assumption 1. As the transition probabilities are unobserved, they must be estimated. However, theoretically, the population shares are equal to the estimated location choice probabilities using discrete choice models, subject to rounding errors (Train 2009). Thus, we can evaluate Assumption 1 by directly checking the stability of the population shares without estimating the location choice probabilities. Simultaneously, this allows us to directly examine the model predictions (i.e., the spatial steady state).

We use microdata from the 2005–2017 American Community Survey retrieved from IPUMS (Ruggles et al. 2019) to evaluate the model assumption for several reasons. First, our model works for a given location choice set. The past location information in the dataset allows us to exclude those individuals not previously in a given location choice set. Second, we use employment information to restrict the sample to full-time workers who work at least 35 hours per week. Third, the personal survey weights combined with residential information can be used to recover the population shares for different location choice sets. Specifically, we restrict the sample to full-time workers living in the 50 states and District of Columbia during the survey years that did not live abroad one year before. We use this sample to recover the population shares for location choice sets at three levels: public use microdata area (PUMA), metropolitan statistical area (MSA), and state.[14]

Figures 5–7 illustrate the population shares for the different location choice sets. First, the location choice set at the state level is considered (Figure 5). The population shares for the 50 states and District of Columbia over 2005–2017 are mostly stable, implying that the location choice probabilities at the state level are relatively time-independent, which supports Assumption 1. This also provides direct evidence that the population

[14] The PUMA boundaries were redrawn in 2012. To be consistent, we thus use consistent PUMA (CPUMA) to identify the PUMAs over the study years. The 2013 boundaries for MSAs from the U.S. Office of Management and Budget are used to identify the MSAs. The non-MSA areas of each state are also included.

distribution at the state level is in a spatial steady state, which is consistent with our model predictions. The dashed line in the figure panels is equal to 0.0196, as calculated by Equation (8) as a reference, indicating the location choice probability when revealed utility is equalized in all states. However, few states sit exactly on this line, and there is no trend converging to this line. The population shares of some states such as California, Florida, New York, and Texas are always above the reference line, indicating they are persistently preferred, whereas workers are immobile regarding these states from the perspective of researchers. The population shares of other states such as Alaska, Idaho, Maine, and Wyoming are continuously below the reference line, showing "over-mobility." According to the one-to-one relationship between the location choice probability and local revealed utility discussed in Section 3, Figure 5 therefore implies that revealed utility differentials always exist (i.e., the spatial disequilibrium is persistent).

The pattern is robust to using different location choice sets at different geographic scales. There are too many CPUMAs in the United States to display (i.e., $N=1078$). Thus, we define individual states as the location choice sets with CPUMAs as the location options and choose certain states to report.[15] Figure 6 shows that the population shares of CPUMAs are highly persistent in all these states, indicating a spatial steady state. The population shares in Figure 6 can thus be interpreted as the conditional location choice probabilities of CPUMAs given a state. It can then be inferred that the unconditional location choice probabilities of CPUMAs are also steady, as the unconditional choice probability of a CPUMA is derived by multiplying the choice probability of the state in which the CPUMA is located (Figure 5) by the conditional choice probability of the CPUMA given the state (Figure 6). Figure 7 shows the population shares for selected MSAs.[16] This figure shows a similar pattern to that

[15] The subsamples are restricted to workers who always stay in the given states. The patterns are similar for unreported states. The figures for unreported states are available upon request.

[16] Non-MSA areas by state are also included as location options. The unreported figures of MSAs are available upon request. The results are robust to excluding non-MSA areas.

in the figures for states and CPUMAs: the probability of choosing an MSA is mostly stable and the population distribution at the MSA level is also in a steady state.

Figure 8 presents the estimated kernel densities of the population shares at the CPUMA and MSA levels. This figure shows that the variation in population shares is small at both levels. We also estimate a year fixed effect model showing that the annual change in the population shares at different geo-unit levels is small and mostly not significantly different from zero.[17] This evidence indicates that Assumption 1 is reasonable and that spatial steady-state theory applies to different location choice sets.

As discussed previously, the literature finds persistence in the spatial disequilibrium (Spilimbergo and Ubeda 2004; Partridge et al. 2015; Goetzke and Islam 2017). The evidence in this section thus confirms the persistence at different geo-unit levels. To sum up, the strong form of the spatial equilibrium is rarely observed because (i) researchers observe imperfect mobility that results from unobserved heterogeneity and (ii) interregional revealed utility differentials (spatial disequilibrium) are highly persistent. Thus, place-based policies may be effective as discussed in the literature (Hall and Romer 2008; Partridge et al. 2015; Goetzke and Islam 2017).

8. Conclusions

This study explains why the strong form of the spatial equilibrium is rarely observed. A simple discrete choice model is used to illustrate that researchers observe imperfect mobility because of the unobserved heterogeneity of workers. Thus, the strong form of the spatial equilibrium is rarely detected as spatial equilibrium conditions are not observed. Although imperfect mobility is more accurately observed if pecuniary moving costs are considered, this does not alter the main conclusion. Assuming that the population elasticity of revealed utility is ignorable, it is shown that the population distribution of a given location choice set converges to a unique stationary distribution, which is defined as a spatial steady state, implying that the spatial disequilibrium is

[17] The results are suppressed to conserve space.

persistent. The higher the pecuniary moving costs, the slower the population distribution converges to a spatial steady state. The assumption of the Markov chain is evaluated using American Community Survey 2005–2017 pooled microdata. These data show that the assumption is reasonable and that spatial steady states for different location choice sets are observed.

The models used in this study, namely the discrete choice model (Carlton 1983; Anas and Chu 1984; Haynes and Fotheringham 1990; Guimarães, Figueirdo, and Woodward 2003, 2004; Bayer, Keohane, and Timmins 2009; Behrens et al. 2017) and Markov chain process (Tarver and Gurley 1965; Brown 1970; Brown and Longbrake 1970; Spilerman 1972; Ginsberg 1973; Singer and Spilerman 1973; Haag and Weidlich 1984; Constant and Zimmermann 2012), have been widely used in the literature. This study thus contributes to the body of knowledge on this topic by combining them to explain the imperfect mobility and persistence of the spatial disequilibrium observed by researchers.

Such a combination has advantages over classical models. First, our framework avoids the discussion of marginal movers by integrating unobserved heterogeneity. Second, it analyzes the causes of the spatial disequilibrium and its persistence from both the static and the dynamic perspectives of researchers. Third, the model assumption is supported by the data and the model prediction fits the data well.

Yet, caveats remain. Agglomeration is ignored in the model for simplicity. Indeed, although our model fits the data well, more profound spatial dynamics should exist during the convergence toward a spatial steady state due to agglomeration. Hence, before any progress can be made, agglomeration per se should be further studied. Nonetheless, this study implies that it is one of the criteria for an ideal functional form of agglomeration that it should ensure the convergence of the Markov chain with the transition probabilities given by Equation (13). Further, this study assumes that workers face the same location choice set with the same information about their options. Cases with incomplete information sets are left to future studies.

References

Abdel-Rahman, Hesham M., and Alex Anas. 2004. “Theories of Systems of Cities.” In *Handbook of Regional and Urban Economics,* edited by J. Vernon Henderson and Jacques-François Thisse, 4:2293–2339. Elsevier. https://doi.org/10.1016/S1574-0080(04)80009-9.

Acolin, Arthur, and Richard K. Green. 2017. “Measuring Housing Affordability in São Paulo Metropolitan Region: Incorporating Location.” *Cities* 62 (February): 41–49. https://doi.org/10.1016/j.cities.2016.12.003.

Anas, Alex, and Chaushie Chu. 1984. “Discrete Choice Models and the Housing Price and Travel to Work Elasticities of Location Demand.” *Journal of Urban Economics* 15 (1): 107–23. https://doi.org/10.1016/0094-1190(84)90025-1.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins. 2009. “Migration and Hedonic Valuation: The Case of Air Quality.” *Journal of Environmental Economics and Management* 58 (1): 1–14. https://doi.org/10.1016/j.jeem.2008.08.004.

Behrens, Kristian, Giordano Mion, Yasusada Murata, and Jens Suedekum. 2017. “Spatial Frictions.” *Journal of Urban Economics* 97: 40–70. https://doi.org/10.1016/j.jue.2016.11.003.

Boman, Anders. 2011. “The Mobility of Immigrants and Natives: Evidence from Internal Migration Following Job Displacement.” *Regional Studies* 45 (3): 283–97. https://doi.org/10.1080/00343400903431003.

Brown, Lawrence A. 1970. “On the Use of Markov Chains in Movement Research.” *Economic Geography* 46 (Sup 1): 393–403. https://doi.org/10.2307/143152.

Brown, Lawrence A., and David B. Longbrake. 1970. “Migration Flows in Intraurban Space: Place Utility Considerations.” *Annals of the Association of American Geographers* 60 (2): 368–84. https://doi.org/10.1111/j.1467-8306.1970.tb00726.x.

Carlton, Dennis W. 1983. “The Location and Employment Choices of New Firms: An

Econometric Model with Discrete and Continuous Endogenous Variables." *The Review of Economics and Statistics* 65 (3): 440–49. https://doi.org/10.2307/1924189.

Carrington, William J., Enrica Detragiache, and Tara Vishwanath. 1996. "Migration with Endogenous Moving Costs." *American Economic Review* 86 (4): 909–30.

Clark, David E, William E Herrin, Thomas a Knapp, and Nancy E White. 2003. "Migration and Implicit Amenity Markets: Does Incomplete Compensation Matter?" *Journal of Economic Geography* 3 (3): 289–307. https://doi.org/10.1093/jeg/3.3.289.

Coen-Pirani, Daniele. 2010. "Understanding Gross Worker Flows across U.S. States." *Journal of Monetary Economics* 57 (7): 769–84. https://doi.org/10.1016/j.jmoneco.2010.08.001.

Compton, Janice, and Robert A. Pollak. 2007. "Why Are Power Couples Increasingly Concentrated in Large Metropolitan Areas?" *Journal of Labor Economics* 25 (3): 475–512. https://doi.org/10.1086/512706.

Constant, Amelie F., and Klaus F. Zimmermann. 2012. "The Dynamics of Repeat Migration: A Markov Chain Analysis." *International Migration Review* 46 (2): 362–88. https://doi.org/10.1111/j.1747-7379.2012.00890.x.

Cortés, Yasna, and Victor Iturra. 2019. "Market versus Public Provision of Local Goods: An Analysis of Amenity Capitalization within the Metropolitan Region of Santiago de Chile." *Cities* 89 (June): 92–104. https://doi.org/10.1016/j.cities.2019.01.015.

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. https://doi.org/10.1162/003355300555079.

David, Quentin, Alexandre Janiak, and Etienne Wasmer. 2010. "Local Social Capital and Geographical Mobility." *Journal of Urban Economics* 68 (2): 191–204. https://doi.org/10.1016/j.jue.2010.04.003.

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. https://doi.org/10.1257/aer.20131706.

Gabriel, Stuart a., and Stuart S. Rosenthal. 2004. "Quality of the Business Environment Versus Quality of Life: Do Firms and Households Like the Same Cities?" *Review of Economics and Statistics* 86 (1): 438–44. https://doi.org/10.1162/003465304774201879.

Ginsberg, R B. 1973. "Stochastic Models of Residential and Geographic Mobility for Heterogeneous Populations." *Environment and Planning A: Economy and Space* 5 (1): 113–24. https://doi.org/10.1068/a050113.

Glaeser, Edward L., and Joshua D Gottlieb. 2008. "The Economics of Place-Making Policies." *Brookings Papers on Economic Activity* 2008: 155–239. https://doi.org/10.1353/eca.0.0005.

Glaeser, Edward L., Joshua D Gottlieb, and Oren Ziv. 2016. "Unhappy Cities." *Journal of Labor Economics* 34 (2): S129–82. https://doi.org/10.3386/w20291.

Glaeser, Edward L., Joshua Gottlieb, and Oren Ziv. 2014. "Maximising Happiness Does Not Maximise Welfare." *VoxEU.Org*. http://voxeu.org/article/maximising-happiness-does-not-maximise-welfare.

Glaeser, Edward L., and Kristina Tobio. 2008. "The Rise of the Sunbelt." *Southern Economic Journal* 74 (3): 609–43. https://doi.org/10.2307/20111988.

Goetzke, Frank, and Samia Islam. 2017. "Testing for Spatial Equilibrium Using Happiness Data." *Journal of Regional Science* 57 (2): 199–217. https://doi.org/10.1111/jors.12311.

González-Val, Rafael. 2019. "The Spatial Distribution of US Cities." *Cities* 91 (August): 157–64. https://doi.org/10.1016/j.cities.2018.11.015.

Greenwood, Michael J, G L Hunt, D S Rickman, and G I Treyz. 1991. "Migration, Regional Equilibrium, and the Estimation of Compensating Differentials." *American Economic Review* 81 (5): 382–90. https://doi.org/10.2307/2006927.

Guimarães, Paulo, Octávio Figueirdo, and Douglas Woodward. 2003. "A Tractable Approach to the Firm Location Decision Problem." *Review of Economics and Statistics* 85 (1): 201–4. https://doi.org/10.1162/003465303762687811.

———. 2004. "Industrial Location Modeling: Extending the Random Utility Framework." *Journal of Regional Science* 44 (1): 1–20. https://doi.org/10.1111/j.1085-9489.2004.00325.x.

Gyourko, Joseph, Matthew Kahn, and Joseph Tracy. 1999. "Quality of Life and Environmental Comparisons." In *Handbook of Regional and Urban Economics*, edited by Edwin S Mills and Paul Cheshire, 3:1413–54. Elsevier. https://doi.org/10.1016/S1574-0080(99)80006-6.

Haag, Günter, and Wolfgang Weidlich. 1984. "A Stochastic Theory of Interregional Migration." *Geographical Analysis* 16 (4): 331–57. https://doi.org/10.1111/j.1538-4632.1984.tb00820.x.

Hall, Robert E, and Paul Romer. 2008. "The Economics of Place-Making Policies. Comments and Discussion." *Brookings Papers on Economic Activity* 2008: 240–53. https://doi.org/10.2307/27561618.

Haynes, Kingsley E., and A. Stewart Fotheringham. 1990. "The Impact of Space on the Application Of Discrete Choice Models." *The Review of Regional Studies* 20 (2): 39–49.

Kahn, Matthew E. 2006. "Environmental Valuation Using Cross-City Hedonic Methods." In *Environmental Valuation: Interregional and Intraregional Perspectives*, edited by John I. Carruthers and Bill Mundy, 27–48. Ashgate Publishing, Ltd.

Kline, Patrick, and Enrico Moretti. 2013. "Place Based Policies with Unemployment." *American Economic Review* 103 (3): 238–43. https://doi.org/10.1257/aer.103.3.238.

Krupka, Douglas J. 2009. "Location-Specific Human Capital, Location Choice and Amenity Demand." *Journal of Regional Science* 49 (5): 833–54.

https://doi.org/10.1111/j.1467-9787.2009.00614.x.

Krupka, Douglas J., and Kwame N. Donaldson. 2013. “Wages, Rents, and Heterogeneous Moving Costs.” *Economic Inquiry* 51 (1): 844–64. https://doi.org/10.1111/j.1465-7295.2012.00475.x.

Leibowicz, Benjamin D. 2017. “Effects of Urban Land-Use Regulations on Greenhouse Gas Emissions.” *Cities* 70 (October): 135–52. https://doi.org/10.1016/j.cities.2017.07.016.

Machin, Stephen, Kjell G. Salvanes, and Panu Pelkonen. 2012. “Education and Mobility.” *Journal of the European Economic Association* 10 (2): 417–50. https://doi.org/10.1111/j.1542-4774.2011.01048.x.

Merino, Fernando, and Maria A. Prats. 2020. “Why Do Some Areas Depopulate? The Role of Economic Factors and Local Governments.” *Cities* 97 (February): 102506. https://doi.org/10.1016/j.cities.2019.102506.

Mulder, Clara H., and Gunnar Malmberg. 2014. “Local Ties and Family Migration.” *Environment and Planning A* 46 (9): 2195–2211. https://doi.org/10.1068/a130160p.

Oswald, Andrew J, and S. Wu. 2010. “Objective Confirmation of Subjective Measures of Human Well-Being: Evidence from the U.S.A.” *Science* 327 (5965): 576–79. https://doi.org/10.1126/science.1180606.

Oswald, Andrew J, and Stephen Wu. 2011. “Well-Being across America.” *Review of Economics and Statistics* 93 (4): 1118–34. https://doi.org/10.1162/REST_a_00133.

Partridge, Mark D., Dan S. Rickman, M. Rose Olfert, and Kamar Ali. 2012. “Dwindling U.S. Internal Migration: Evidence of Spatial Equilibrium or Structural Shifts in Local Labor Markets?” *Regional Science and Urban Economics* 42 (1–2): 375–88. https://doi.org/10.1016/j.regsciurbeco.2011.10.006.

Partridge, Mark D., Dan S. Rickman, M. Rose Olfert, and Ying Tan. 2015. “When

Spatial Equilibrium Fails: Is Place-Based Policy Second Best?” *Regional Studies* 49 (8): 1303–25. https://doi.org/10.1080/00343404.2013.837999.

Reynolds, C. Lockwood, and Shawn Rohlin. 2014. “Do Location-Based Tax Incentives Improve Quality of Life and Quality of Business Environment?” *Journal of Regional Science* 54 (1): 1–32. https://doi.org/10.1111/jors.12035.

Roback, Jennifer. 1982. “Wages, Rents, and the Quality of Life.” *Journal of Political Economy* 90 (6): 1257–78. https://doi.org/10.1086/261120.

———. 1988. “Wages, Rents, and Amenities: Differences Among Workers and Regions.” *Economic Inquiry* 26 (1): 23–41. https://doi.org/10.1111/j.1465-7295.1988.tb01667.x.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2019. “IPUMS USA: Version 9.0 [Dataset].” MN: IPUMS.

Saks, Raven E, and Abigail Wozniak. 2011. “Labor Reallocation over the Business Cycle: New Evidence from Internal Migration.” *Journal of Labor Economics* 29 (4): 697–739. https://doi.org/10.1086/660772.

Singer, Burton, and Seymour Spilerman. 1973. “Social Mobility Models for Heterogeneous Populations.” *Sociological Methodology* 5: 356–401. https://doi.org/10.2307/270841.

Spilerman, Seymour. 1972. “The Analysis of Mobility Processes by the Introduction of Independent Variables into a Markov Chain.” *American Sociological Review* 37 (3): 277–94. https://doi.org/10.2307/2093468.

Spilimbergo, Antonio, and Luis Ubeda. 2004. “A Model of Multiple Equilibria in Geographic Labor Mobility.” *Journal of Development Economics* 73 (1): 107–23. https://doi.org/10.1016/j.jdeveco.2002.11.001.

Tarver, James D., and William R. Gurley. 1965. “A Stochastic Analysis of Geographic Mobility and Population Projections of the Census Divisions in the United States.” *Demography* 2 (1): 134–39.

Train, Kenneth E. 2009. “Discrete Choice Methods with Simulation.” *Cambridge University Press, New York*, no. 2nd: 1148. https://doi.org/10.1016/S0898-1221(04)90100-9.

Vigdor, Jacob L. 2002. “The Pursuit of Opportunity: Explaining Selective Black Migration.” *Journal of Urban Economics* 51 (3): 391–417. https://doi.org/10.1006/juec.2001.2250.

Winters, John V. 2009. “Wages and Prices: Are Workers Fully Compensated for Cost of Living Differences?” *Regional Science and Urban Economics* 39 (5): 632–43. https://doi.org/10.1016/j.regsciurbeco.2009.05.001.

Winters, John V, and Yu Li. 2017. “Urbanisation, Natural Amenities and Subjective Well-Being: Evidence from US Counties.” *Urban Studies* 54 (8): 1956–73. https://doi.org/10.1177/0042098016631918.

Yang, Tianren, Haozhi Pan, Geoffrey Hewings, and Ying Jin. 2019. “Understanding Urban Sub-Centers with Heterogeneity in Agglomeration Economies—Where Do Emerging Commercial Establishments Locate?” *Cities* 86 (March): 25–36. https://doi.org/10.1016/j.cities.2018.12.015.

Figure 1: Location Choice Probability in Homogenous Setting

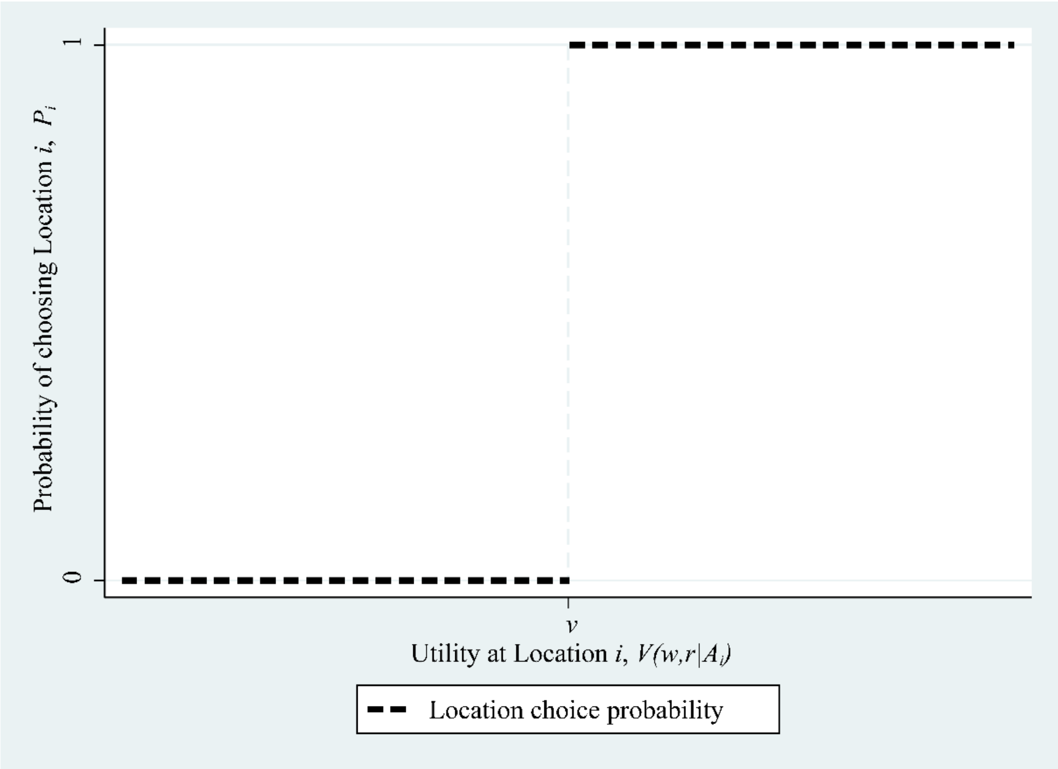


Figure 2: Location Choice Probability with Unobserved Heterogeneity

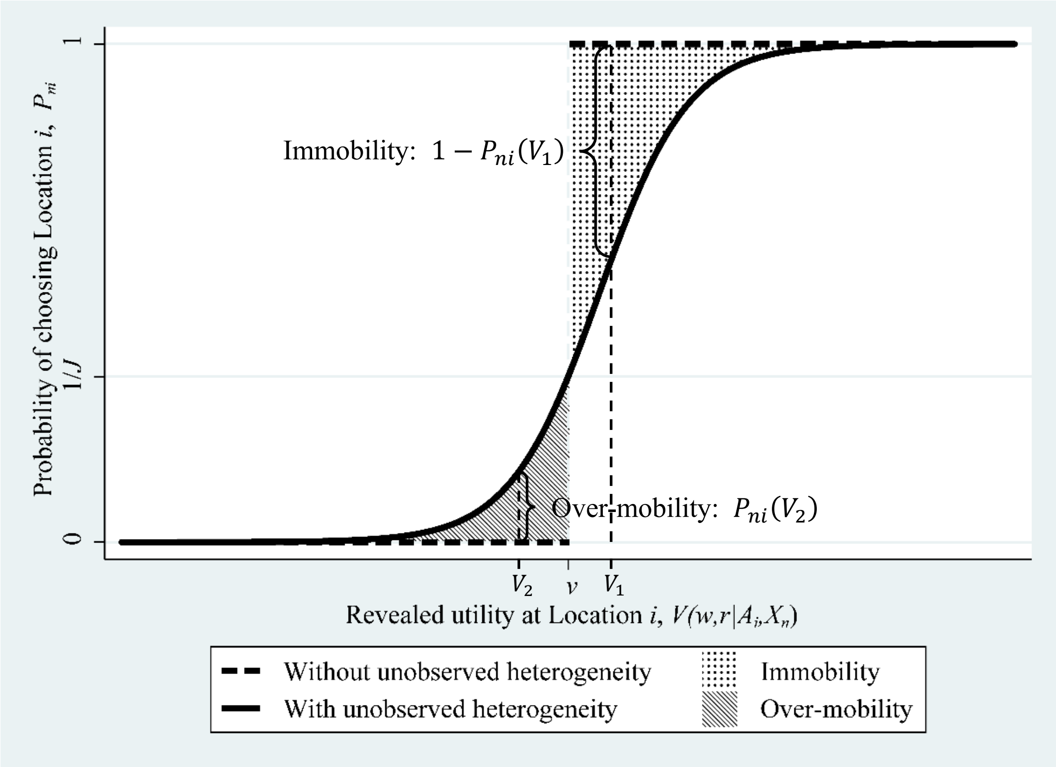
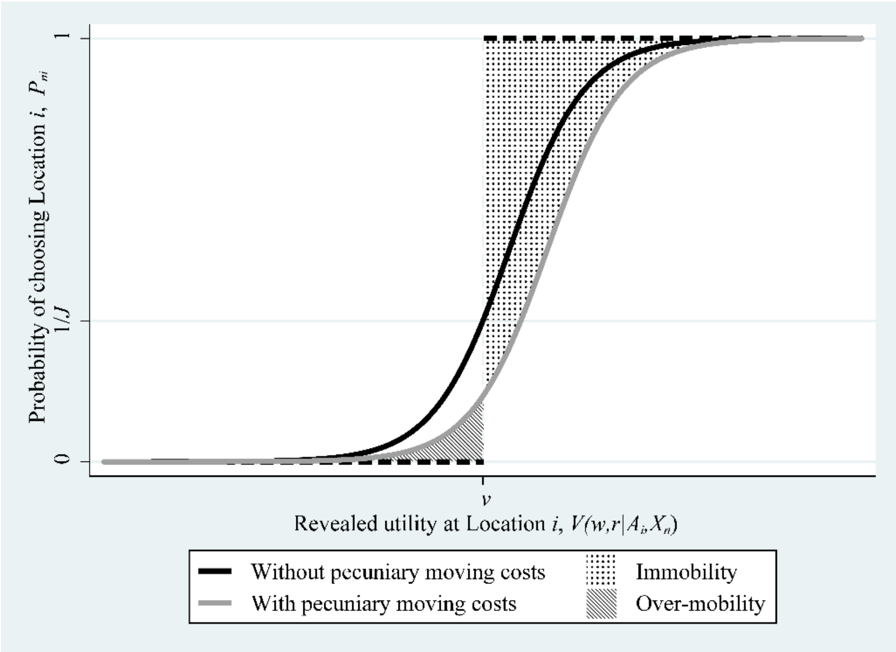


Figure 3: Location Choice Probability with Pecuniary Moving Costs

(a)



(b)

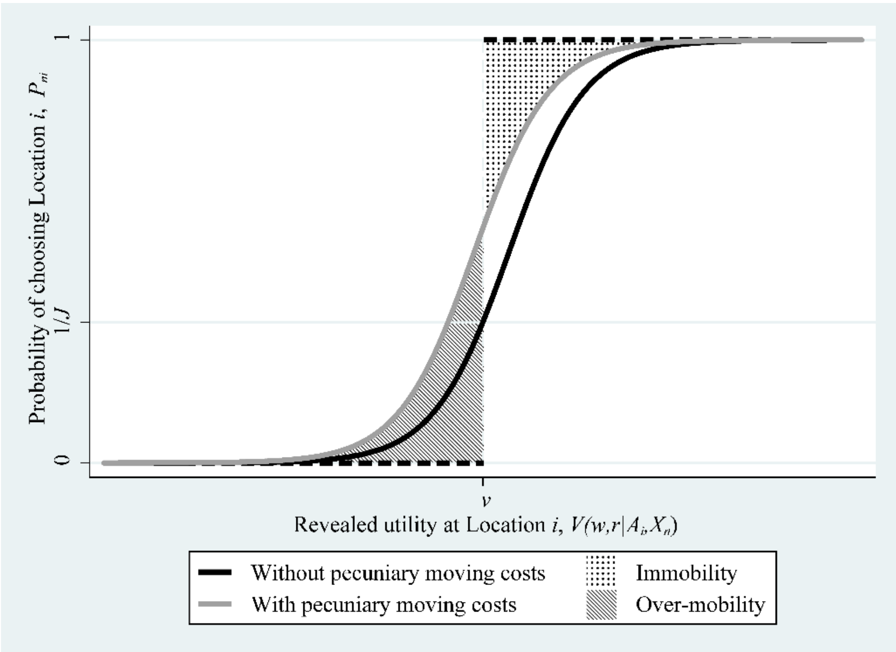
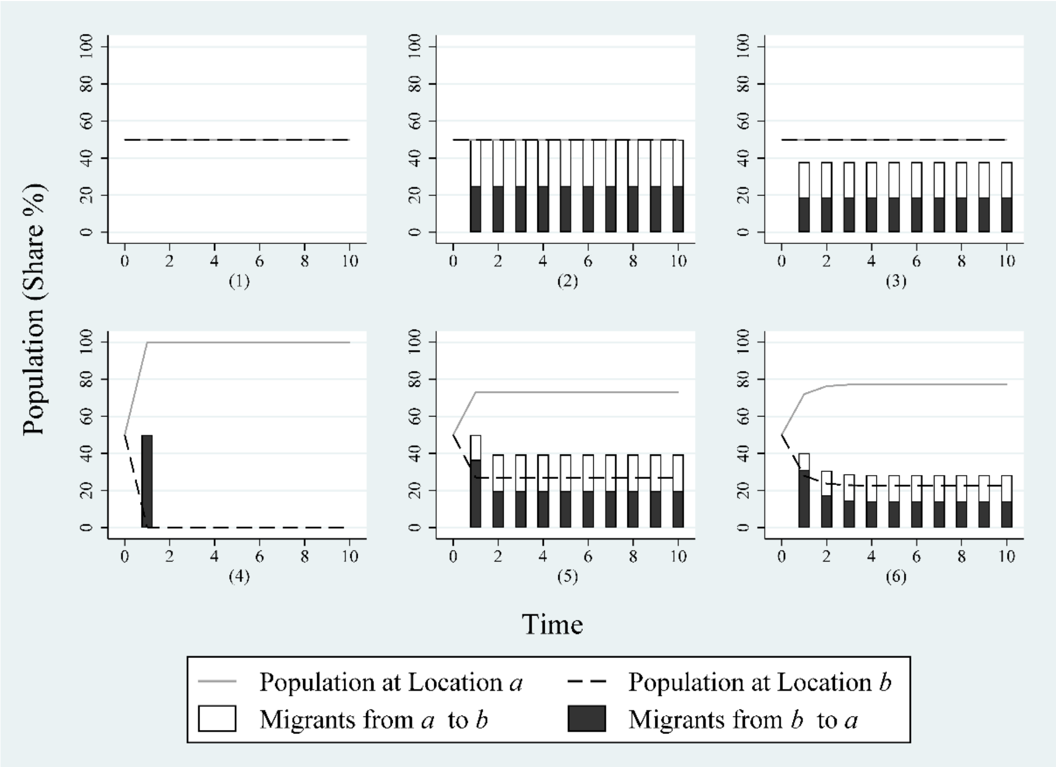
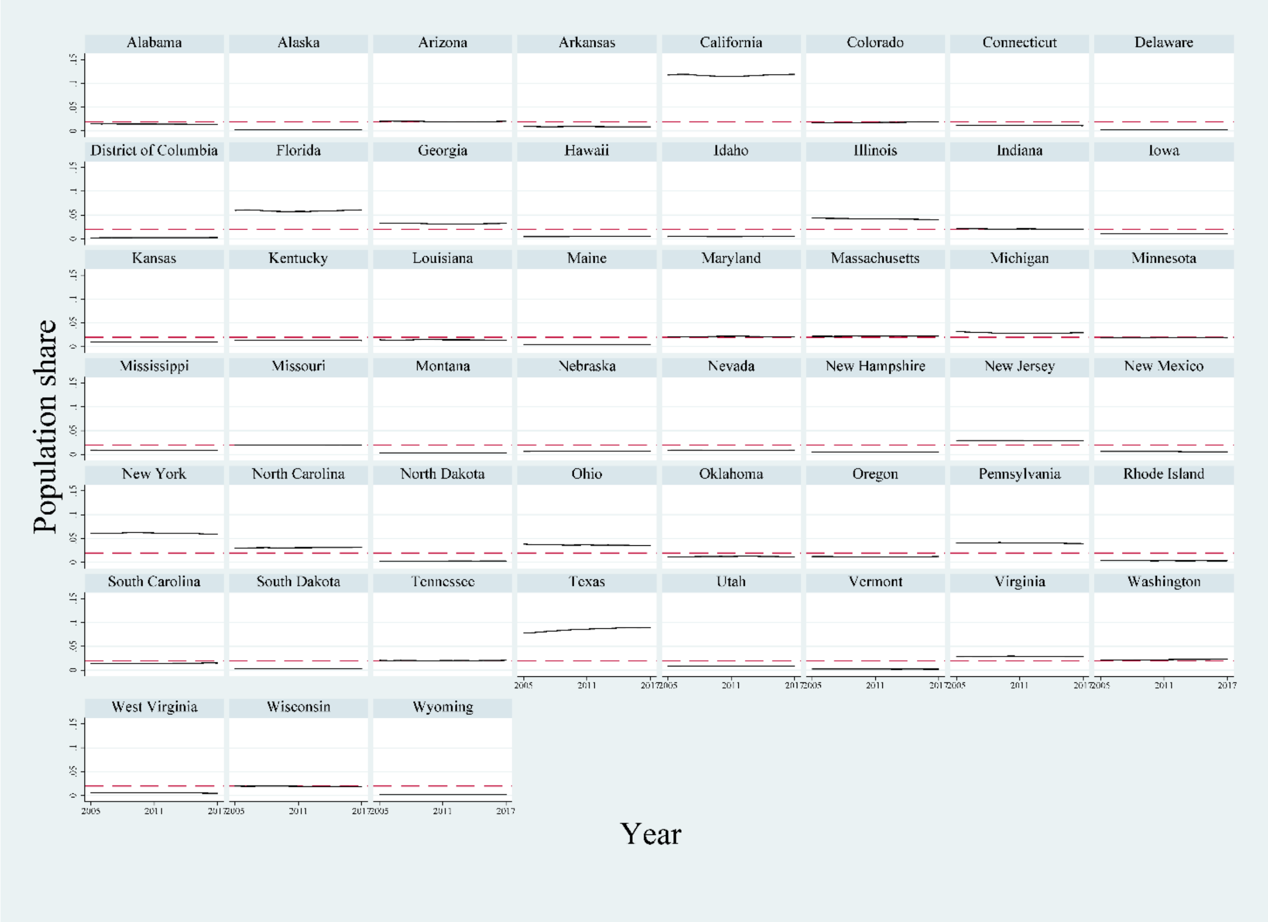


Figure 4: Spatial Dynamics of Population



Notes: The initial population of Locations a and b are assumed to be 50 and 50, respectively. The total population of the location choice set is 100, so that the vertical axis can be interpreted as either in levels or percentages.

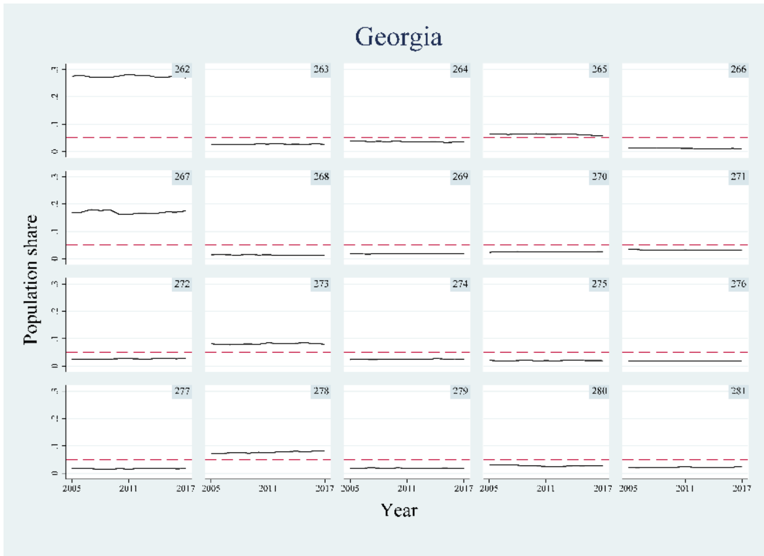
Figure 5: 2005–2017 Population Shares at State Level



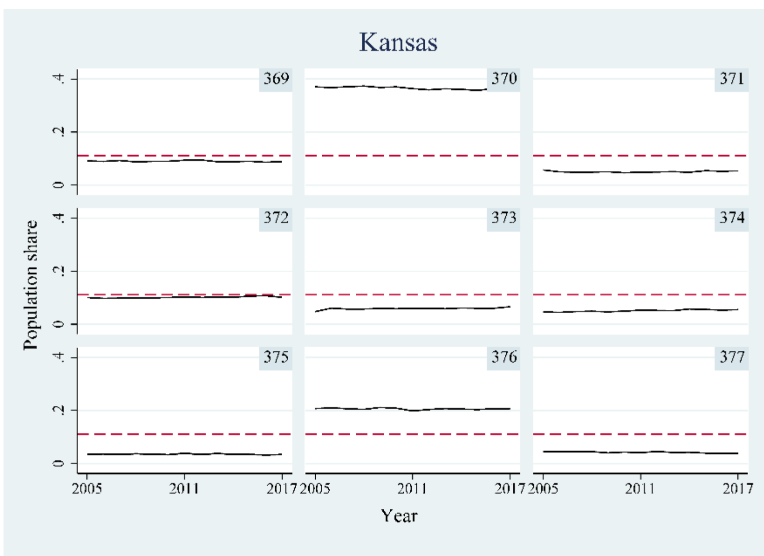
Notes: The sample is restricted to full-time workers living in the 50 states and District of Columbia during the survey years that did not live abroad one year before.

Figure 6: 2005–2017 Population Shares at CPUMA Level

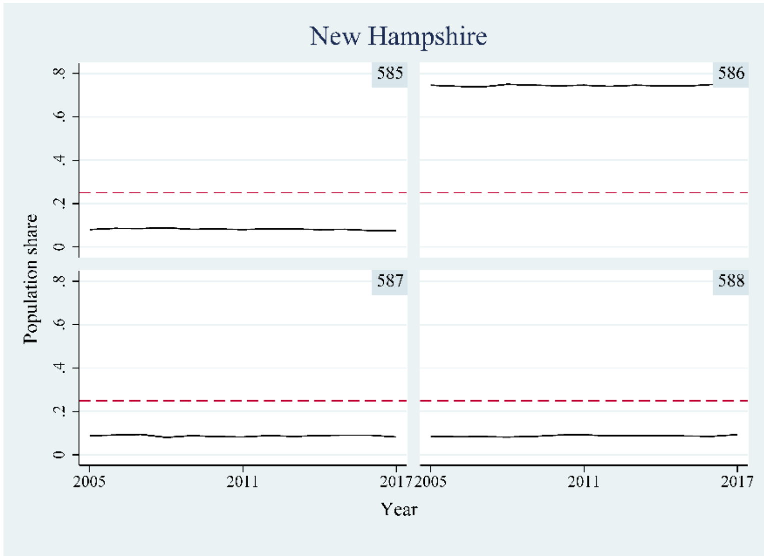
(a)



(b)

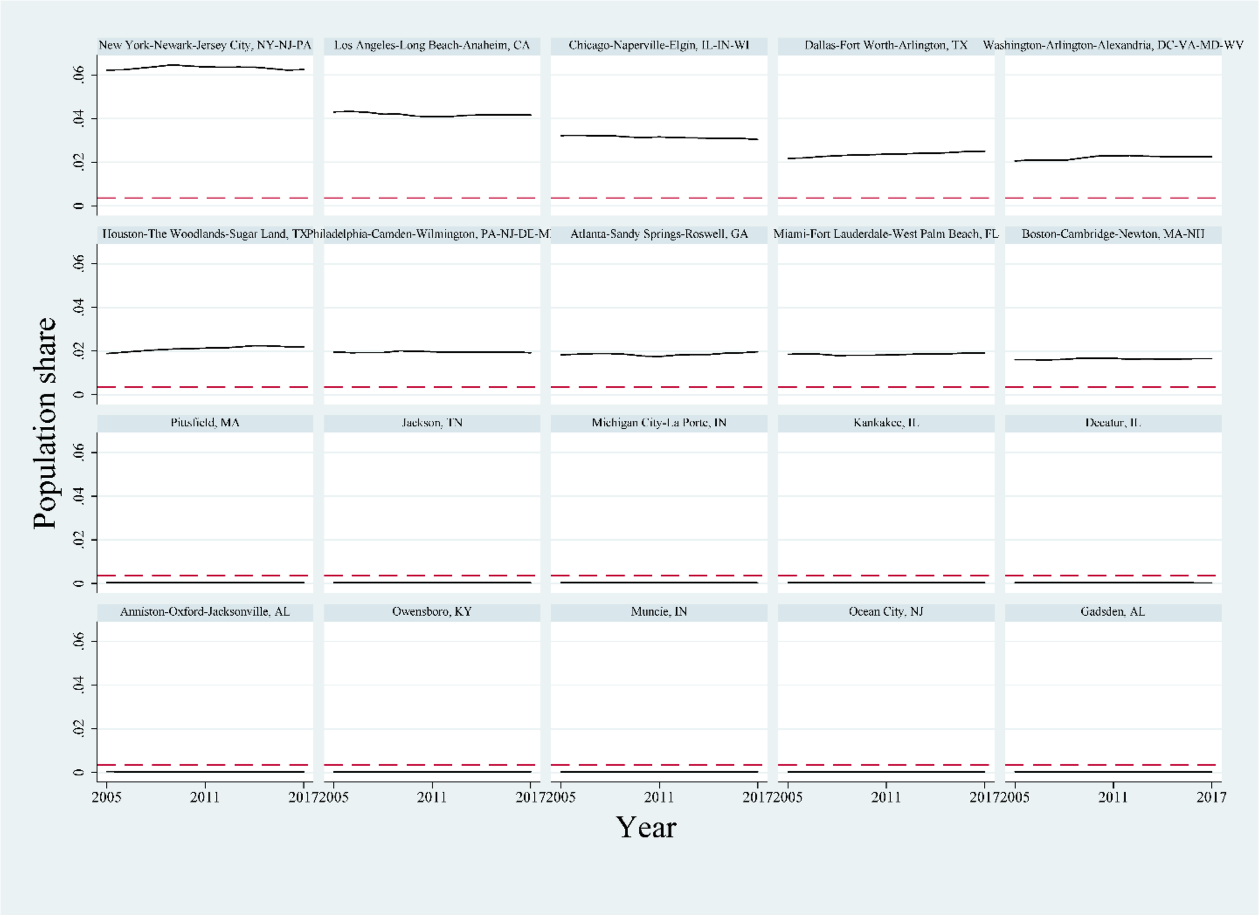


(c)



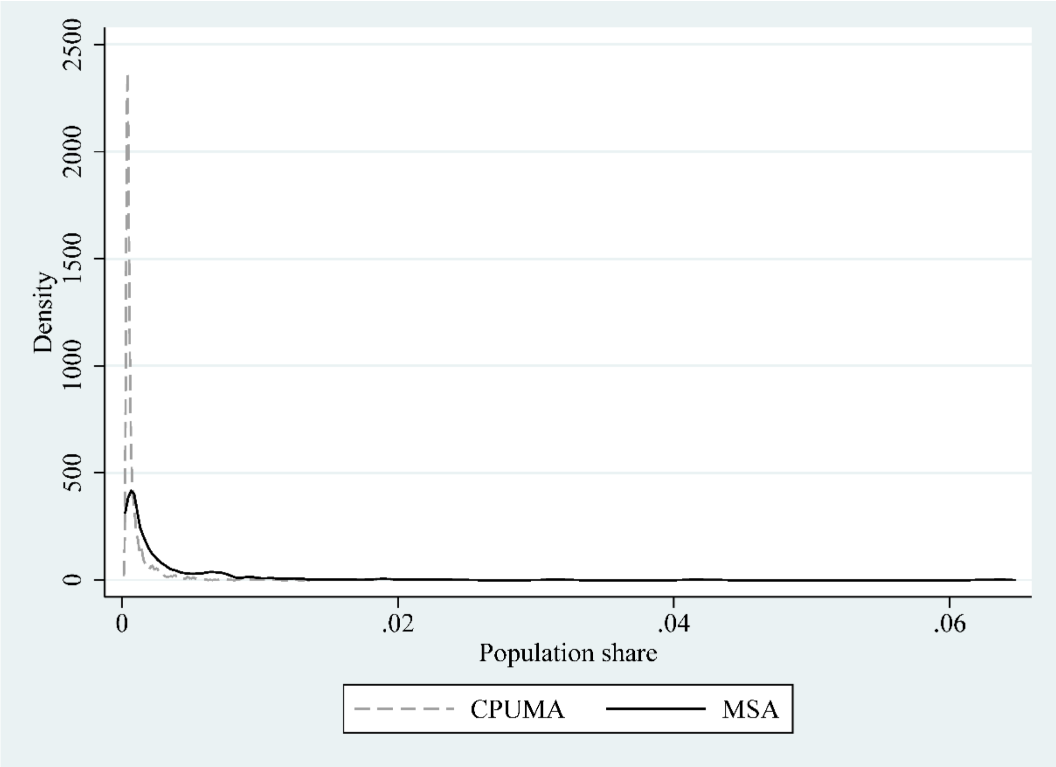
Notes: The samples are restricted to full-time workers living in the given states during the survey years.

Figure 7: 2005–2017 Population Shares of Selected MSAs



Notes: The first two rows are the top 10 MSAs with highest population shares and the last two rows are the bottom 10 MSAs with lowest population shares. The sample is restricted to full-time workers living in the United States during the survey years that did not live abroad one year before.

Figure 8: Estimated Kernel Density of Population Shares



Notes: The default kdensity command in Stata is used to estimate the kernel densities.

Table 1: Definition of Moving Costs

	Category	Source		Observed by researchers
Moving cost	Psychological moving cost	Heterogeneity	Observable: Age, health status, education, etc.	Ideally, yes. Directly control for.
			Unobservable: Travel preferences; other unobserved heterogeneity	No. Generate imperfect mobility.
	Pecuniary moving cost	Beeline distance of moving		Ideally, yes.

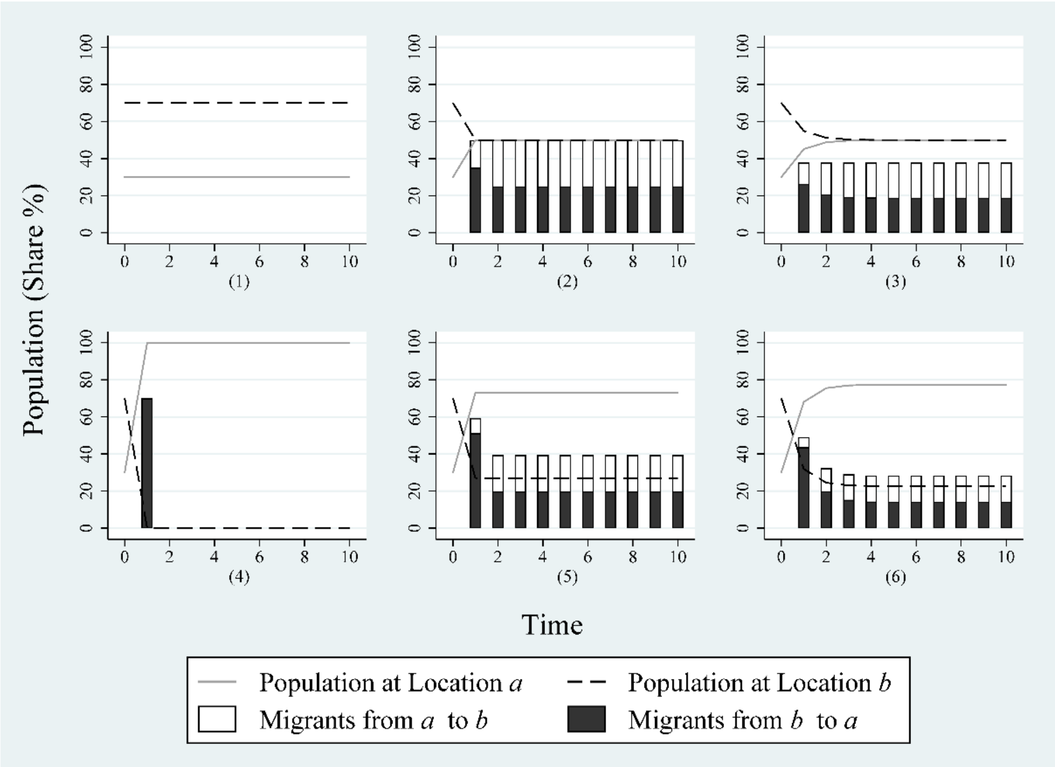
Notes: See the text for details.

Table 2: Numerical Examples in Different Scenarios

	(1)	(2)	(3)	(4)	(5)	(6)
V_a	1	1	1	2	2	2
V_b	1	1	1	1	1	1
T	0	0	-0.5	0	0	-0.5
Heterogeneity	No	Yes	Yes	No	Yes	Yes
$P_{a\|b}$	0	0.5	0.378	1	0.731	0.622
$P_{b\|a}$	0	0.5	0.378	0	0.269	0.182
Mobility regarding Location a	Perfect	"Over-mobile"	"Over-mobile"	Perfect	"Immobile"	"Immobile"
Mobility regarding Location b	Perfect	"Over-mobile"	"Over-mobile"	Perfect	"Over-mobile"	"Over-mobile"

Notes: $P_{a|b}$ and $P_{b|a}$ in Columns (1) and (4) come from Equation (3). $P_{a|b}$ and $P_{b|a}$ in Columns (2) and (5) are calculated using Equation (5). $P_{a|b}$ and $P_{b|a}$ in Columns (3) and (6) are calculated using Equation (10).

Appendix Figure A1: Spatial Dynamics of Population



Notes: The initial population of Locations a and b are assumed to be 30 and 70, respectively. The total population of the location choice set is 100, so that the vertical axis can be interpreted as either in levels or percentages.