Internal Migration and Relocation Aversion

Axel Watanabe*

May 15, 2025

Abstract

Urban economics has long relied on the assumption of perfect mobility, where workers relocate in direct response to differences in local economic conditions. However, empirical evidence on internal migration reveals persistent frictions and a strong distance gradient. This paper develops a general equilibrium model of migration in which workers differ in their aversion to relocation. We focus on the role of distance, modeling it as a continuous variable that workers perceive logarithmically. The model predicts that large cities disproportionately attract workers with low relocation aversion and high tolerance for long-distance migration. It also shows that, when distance-related frictions are eliminated, smaller cities experience a greater reduction in agglomeration intensity than larger ones.

Keywords: geographic mobility, internal migration, agglomeration
JEL classification: J61, R23

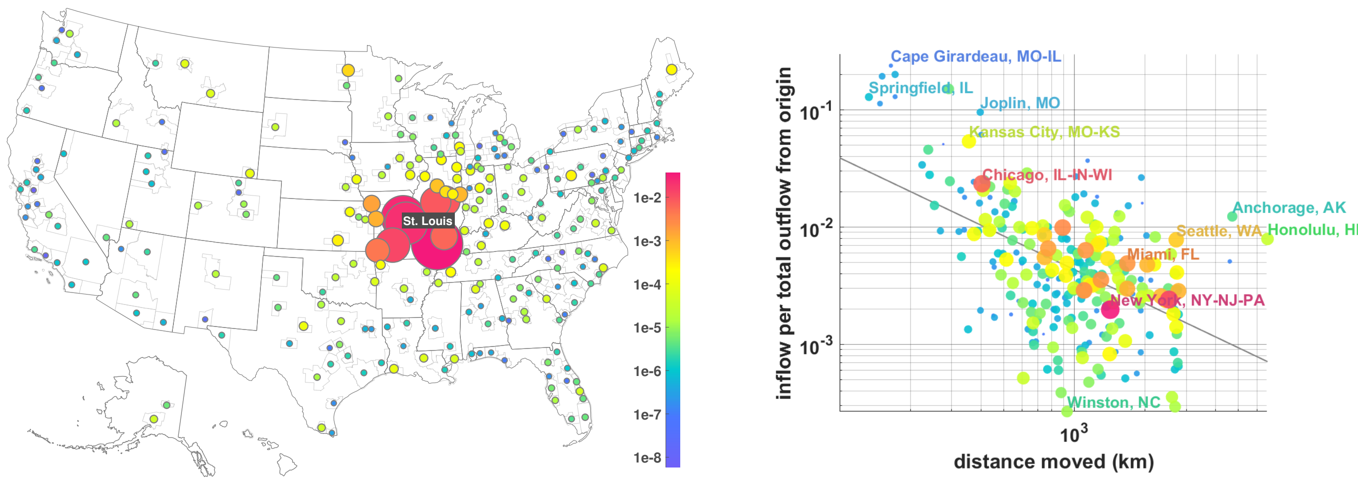
1 Introduction

1.1 Domestic Migratory Patterns

Internal migration in the U.S. exhibits a pronounced geographic gradient: the number of movers sharply declines with distance. Most in-migrants to cities originate from nearby areas, and long-distance relocation remain rare. Figure 1(a) illustrates this pattern for the St. Louis metropolitan area, where most new residents are from Missouri and Illinois despite national mobility being unrestricted. Figure 1(b) confirms this exponential decay in inflows with distance. The sharp distance decay in migration flows suggests that geographic proximity plays a central role in relocation decisions.[1]

*Department of Economics, Concordia University, and CIREQ. Address: 1455, boul. De Maisonneuve Ouest, Montréal (Québec) H3G 1M8. E-mail: axel.watanabe@concordia.ca. I am grateful for valuable inputs from Prosper Dovonon, Xintong Han, Amrutha Manjunath, and Tomoya Mori. This work was supported by Soutien à la recherche pour la relève professorale, 2021-NP-282155, of Le Fonds de recherche du Québec – Société et culture. The usual disclaimer applies.

[1]This observation echoes Tobler's First Law of Geography: "everything is related to everything else, but near things are more related than distant things." (Tobler, [Tob70]).



(a) Origins of incoming residents. A dot size represents the inflow from each location, normalized by the total outflow from it.

(b) Incoming residents by distance moved. Dots are proportionate to origin sizes.

Figure 1. In-migrants to St. Louis Metropolitan Statistical Area (MSA). *Data source*: US Census Bureau, ACS 2009-2013.

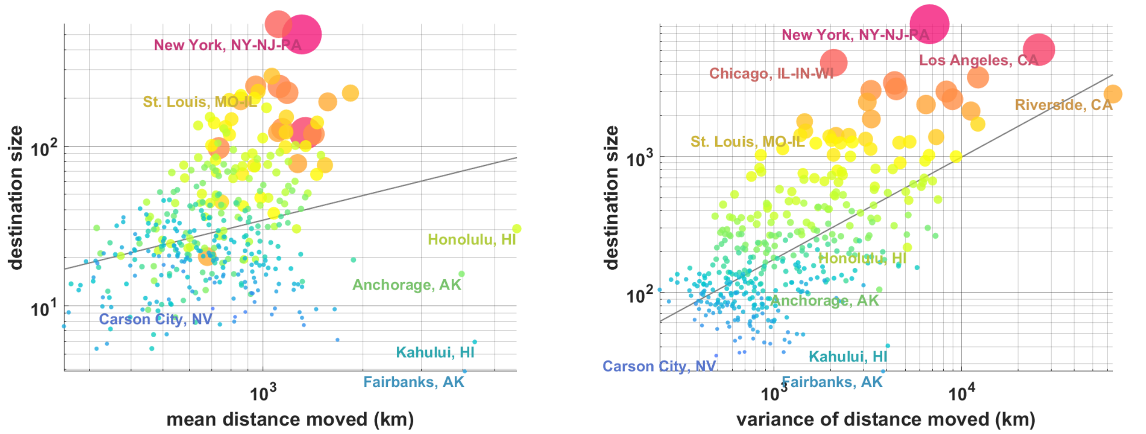
Moreover, the nature of migration varies with city size. Larger cities draw migrants from a broader set of origins and over greater average distances, as shown in figure 2. In addition to the variation in the origins of migrants and the distances they travel, cities also differ in the overall volume of domestic migration. Figure 3 shows the annual inflows and outflows by city, with dot sizes scaled to city population. Some cities experience churn rates as high as 12% annually, while others see much lower levels of turnover.

These differences matter. Migration is not merely a demographic backdrop—it plays an active role in shaping the urban economy. Since agglomeration economies scale with population, variation in migration intensity alters the strength of those economies. Inflows expand the labor market, introduce new skills, and reinforce scale effects; outflows can diminish them. Understanding migration behavior—and the frictions that modulate it—is thus central to explaining the formation, persistence, and divergence of cities.

These patterns highlight the importance of understanding how variation in geographic mobility influences population distribution and urban agglomeration. We develop a parsimonious model in which individuals differ in their relocation aversion—a latent trait that governs the cost they associate with moving. We focus especially on the role of distance as a constraint, modeling its influence on migration decisions and, in turn, on equilibrium city sizes.

1.2 Related Literature

The idea that workers relocate in response to spatial differences in economic conditions has a long tradition. Seminal work by Sjaastad [Sja62] and Harris and Todaro [HT70]



(a) Mean distance moved, controlled for variance.

(b) Variance of distance moved, controlled for mean.

Figure 2. Mean and standard deviation of distance moved among in-migrants to each destination. Dot sizes are proportionate to destination sizes. The distance moved exhibits high variance. Among those who made intercity migrations, the average distance moved was 981 km—roughly the distance between New York City and Indianapolis—with a standard deviation of 1,186 km. The geometric mean was 478 km, with a standard deviation of 427 km. *Data source*: US Census Bureau, ACS 2009-2013.

modeled migration as an investment in future earnings, focusing on expected wage differentials. McFadden's [McF74] discrete choice framework further formalized location decisions in terms of observable and unobservable characteristics of places and individuals.

Subsequent research extended these models to incorporate heterogeneous preferences and market imperfections. Borjas et al. [BBT92] and Kennan and Walker [KW11] modeled skill-dependent sorting, while Diamond [Dia16] shows how cities endogenously generate amenities that attract high-skilled workers, reinforcing spatial inequality.

In parallel, much of urban economics has traditionally assumed perfect mobility—that is, workers face no barriers in relocating to utility-maximizing locations (Starrett [Sta78]; Boyd and Conley [BC97]). This simplifying assumption enabled key theoretical insights in general equilibrium modeling but contrasts with empirical evidence that highlights persistent frictions in relocation.

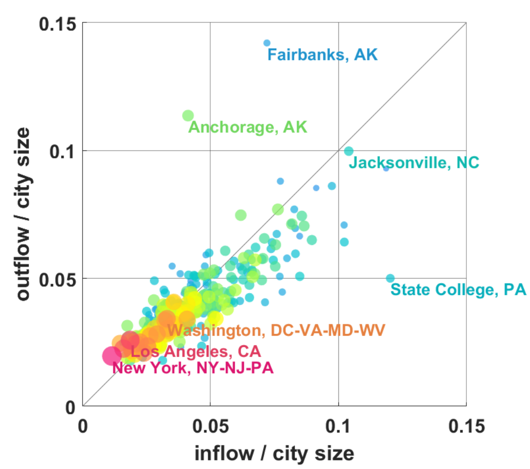


Figure 3. Annual inflows and outflows by city, as a percentage of population. Dot sizes reflect city population. Migration intensity varies widely across metropolitan areas, with significant implications for urban growth and agglomeration dynamics. *Data source*: U.S. Census Bureau, ACS 2009–2013.

Cultural, familial, and institutional factors often inhibit mobility. Falck et al. [FHLS12] document that Germans are reluctant to leave regions sharing their dialect. Helliwell [Hel97] and Woodard [Woo11] find similar effects in Canada and the U.S., where social or linguistic boundaries constrain movement. Moretti [Mor12] emphasizes monetary and psychological relocation costs. Green [Gre24] shows that U.S. Navy veterans randomly assigned to WWII ships were more likely to co-locate later in life, demonstrating the geographic persistence of peer ties.

Despite such explanations, a puzzling empirical fact remains: relocation rates are persistently lower than what standard models predict. Jia et al. [JMSW23] note that the relocation costs needed to rationalize observed migration flows often exceed plausible financial or logistical barriers. This has motivated research into the deeper, often unobservable sources of mobility frictions—whether psychological, cultural, or network-based.

While much of the spatial sorting literature emphasizes observable traits like skill, the present paper focuses on how unobserved heterogeneity in relocation aversion shapes migration flows and agglomeration outcomes. Even if two individuals face identical wage gaps across cities, their migration decisions may diverge depending on how they perceive and tolerate the cost of distance.

1.3 Outline

We build a static general equilibrium model of city formation in which individuals draw their birthplace and relocation tolerance from an exogenous distribution. Migration decisions depend on three forces: the productivity advantages of cities (via agglomeration), relocation costs that increase with distance, and individual heterogeneity in relocation aversion. The model predicts that larger cities are disproportionately populated by workers more tolerant of relocation and more willing to move long distances. It also rationalizes the commonly observed logarithmic decay in migration rates with distance.

Empirically, we estimate cutoff levels of relocation tolerance by destination size and distance moved, using U.S. Census migration data across 381 MSAs. We show that ignoring variation in distance tolerance underestimates the extent to which spatial frictions shape urban agglomeration. In a counterfactual scenario with perfect mobility, the variance in city sizes drops significantly, particularly for smaller cities that rely on short-distance inflows.

Our model highlights how imperfect mobility, mediated by distance aversion, shapes migration flows and influences the degree of agglomeration across cities of different sizes. In this article, **perfect mobility** refers to homogeneity in individuals' attitudes toward relocation. While consumers may incur relocation costs, if these costs are identical across individuals, they do not generate spatial variation and are thus consistent with the assumption of perfect mobility. This paper focuses instead on **imperfect mobility**, where individuals differ in their aversion to moving. The analysis explores how such heterogeneity influences migration flows and, in turn, the strength of agglomeration forces.

The rest of the paper is organized as follows: In the upcoming section we lay out the model and describe the relationship among relocation, distance moved, inflow and agglomeration. We will empirically validate our theoretical predictions in section 3. Section 4 concludes our analysis.

2 Model

2.1 Landscape

Consider a closed production economy situated on a circle with unit circumference. There are $I \in \mathbb{N}$ cities along the circumference, indexed by i.

Each city is endowed with a unit mass of potential consumers who may choose to relocate there. This framework assumes that each city hosts a unique set of potential in-migrants, referred to as type-i households. To simplify notation, we focus on a representative city and suppress the index i unless necessary.

Each potential migrant draws a pair of attributes $(x, y) \in \left[0, \frac{1}{2}\right] \times \mathbb{R}$ that jointly determine her migration behavior. Here, x denotes the distance from her birthplace to the city in question, and y denotes her individual tolerance for relocation. We refer to $f_X(x)$ and $f_Y(y)$ as the marginal probability density functions (PDFs) for distance and relocation tolerance, respectively, with $F_X(x)$ and $F_Y(y)$ denoting the corresponding cumulative distribution functions (CDFs).

If a worker is born in the destination city, $x = 0$; if she is born at the furthest point on the circle, $x = \frac{1}{2}$. To distinguish the two locations conceptually, we refer to the destination city as "urban" and the birthplace x as "rural", regardless of the local density $f_X(x)$.

As discussed in section 1, many factors shape a worker's willingness to relocate. In this model, we isolate the role of distance by decomposing relocation costs into two parts: (i) geographic distance x between birthplace and destination, and (ii) all other idiosyncratic factors, summarized by y.[2] This follows the spirit of Borjas et al. [BBT92], who treat unobserved mobility determinants via a reduced-form index.

We assume the model is static: each worker simultaneously and independently decides whether to relocate. A type-i individual chooses between remaining at her birthplace x or moving to city i. We assume that y is independently drawn from $f_Y(y)$ and uncorrelated with x. This assumption reflects the view that relocation tolerance is not inherently tied to birthplace.[3]

[2]The tolerance variable y absorbs all non-distance determinants of relocation behavior—such as income, education, family structure, and social ties—without modeling them separately. This reduced-form approach isolates the role of geographic distance while accommodating realistic heterogeneity in individual mobility.

[3]While y and x may be correlated in practice—for example, if certain regions foster greater openness to

Furthermore, we assume y is drawn from a common distribution $f_Y(y)$ across all types. This reflects the view that relocation tolerance is not determined by acquired traits such as education or earnings potential. Instead, heterogeneity in migration outcomes arises endogenously from individual decisions, not from type-specific assumptions about mobility.

Each worker's migration choice depends on three variables: distance x, relocation tolerance y, and city size s. City size s affects urban productivity and, in turn, wage levels. A lower x, higher y, or larger s increases the likelihood of relocation.

All types are ex ante identical in distribution: a type-i worker draws (x, y) from the same joint distribution as a type-j worker. Equilibrium variation in city sizes arises ex post from differences in the realized draws of (x, y).

2.2 Moving Workers

Each worker consumes a single homogeneous good produced using labor and experiences disutility from relocation. Following the spatial equilibrium tradition, a relocating worker's utility depends on both the consumption level in the destination and the cost of moving from her origin.[4]

A consumer is endowed with a unit of time and an equal stock of capital. She divides her time into leisure l and labor $1-l$, from which she produces numéraire composite consumption good c.[5] Along with labor and leisure, she also chooses her residence, either her birthplace x or city i. Her preferences are represented by $u(c, l) = c + (1-l)^{\beta}$, where $\beta \in (0, 1)$.

With l units of labor, she produces output z and earns labor income at the rate of w per unit of labor input. In addition, each consumer receives dividend π. From her income $wl + \pi$, **the cost of relocation** $r = r(x, y)$ (≥ 0) is deducted. Her budget constraint is then $wl + \pi - r(x, y) \geq c$. We will describe the cost of relocation next.

Relocation from birthplace x to city i unfolds two counteracting effects on earnings. If a worker moves to city i, her wage rate w will be different from a rural rate at her birthplace as she taps into the elevated productivity in the city. Agglomerative advantages of the city are passed onto her through w alone. On the other hand, its cost is represented by $r(x, y)$. It captures the relocation costs other than the urban-rural wage differential above. The opportunity cost of staying put is thus $w - r(x, y)$. We characterize the relocation cost as below:

migration—we treat them as independent to focus on distance itself as a friction.

[4]This setup parallels the structure in Allen and Arkolakis [AA14], who model utility as a function of wages and bilateral migration costs. Our framework simplifies their treatment by focusing on behavioral aversion to distance rather than trade or congestion frictions.

[5]We do not model land or congestion as counterweights to agglomeration. In this setup, geographic distance plays that role, curbing concentration through relocation frictions rather than land scarcity or housing costs.

Assumption 2.1 Moving Distance, Relocation Tolerance and Costs

Relocation cost $r(x, y)$ is continuously differentiable and satisfies the following:

$$\frac{\partial r(x, y)}{\partial x} > 0, \tag{1}$$

$$\frac{\partial^2 r(x, y)}{\partial x^2} < 0, \quad \textit{and} \tag{2}$$

$$\frac{\partial r(x, y)}{\partial y} < 0, \tag{3}$$

for $(x, y) \in \left(0, \frac{1}{2}\right) \times \mathbb{R}$.

These properties imply: 1) Relocation cost increases with distance; 2) individuals perceive distance non-linearly (concave in x); and 3) Greater tolerance y reduces relocation cost.

Turning to production, a worker who moved to city i produces non-tradable composite goods in a perfectly competitive environment according to $c = s^{\alpha} k^{1-\beta} l^{\beta}$. City size $s \in [0, 1]$ is endogenous. Exponent $\alpha (> 0)$ reflects agglomeration-driven productivity. Each firm employs labor l along with fixed capital k. Exponent $\beta \in (0, 1)$ is the labor share. The citywide total productivity s^{α} itself is increasing in size as documented in Rosenthal and Strange [RS04]. Furthermore, the equilibrium real wage $w = \beta s^{\alpha} \left(s^{\frac{\alpha}{\beta-1}} + k\right)^{-\beta+1}$ increases with s for sufficiently large s and/or k.[6]

Putting this together, the indirect utility function of a relocating worker is $V(x, y, s) = v(s) - r(x, y)$, where $v(s) := s^{\alpha} \left(s^{\frac{\alpha}{\beta-1}} + k\right)^{-\beta+1}$. This parallels McFadden's framework [McF74], where the utility of a location depends on observable characteristics (here, distance x and city size s) and unobserved individual traits (relocation tolerance y), summarized through $r(x, y)$.

2.3 Non-Moving Workers

Outside the city, the economy is autarkic. Non-moving workers incur no relocation costs, but they also forgo the productivity benefits associated with urban agglomeration. Let $\bar{V}$ denote the utility level that a non-moving worker achieves. We assume $\bar{V}$ is constant across all birth locations. This simplifying assumption enables closed-form solutions in the analysis that follows.

In reality, local economic conditions may vary across birthplaces. We assume such variation is partly captured by the distribution $f_X(x)$, that is, by the likelihood of being

[6] This is consistent with empirical findings on agglomeration economies; see Combes et al. [CDGR12]. Although many urban models use a reduced-form specification such as $w(s) = As^{\gamma}$, our model derives wages endogenously from a structural form based on production, linking agglomeration and labor market conditions to wage levels; see Moretti [Mor12] and Allen and Arkolakis [AA14] for examples of reduced-form log-linear wage functions. Our structural specification incorporates agglomeration via s^{α} with additional modulation from labor market size s.

born at a given distance x from the city. As discussed later, the migration decision implied by this structure resembles a gravity model.

2.4 Migration Decision

Let $\varphi(x, y, s; \bar{V})$ denote the utility difference between migrating to the city and staying in one's birthplace:

$$\varphi\left(x, y, s; \bar{V}\right) := V(x, y, s) - \bar{V} = v(s) - r(x, y) - \bar{V}.$$

<u>Assumption 2.2 Urban-Rural Utility Differential</u>

Given destination size s, the marginal worker at the threshold of relocating satisfies:

$$\varphi\left(x, y, s; \bar{V}\right) = 0. \tag{4}$$

Remark. For simplicity, we suppress $\bar{V}$ and write $\varphi(x, y, s)$ in what follows.

Along the locus $\varphi(x, y, s) = 0$, we analyze how the marginal indifference condition responds to changes in distance, relocation tolerance, and city size. The marginal utility gain from increasing s is $\partial v/\partial s$, which must be offset by a corresponding increase in relocation cost $\partial r/\partial s$. That is, workers who are indifferent at higher values of s must face higher effective relocation costs—either because they are farther from the city or less tolerant of relocation.

We specify the impact of moving distance and relocation tolerance on agglomeration as follows:

<u>Proposition 2.1 Distance Moved, Relocation Tolerance and Agglomeration</u>

Along $\varphi(x, y, s) = 0$,

$$\frac{\partial x}{\partial s} = \frac{-\varphi_s}{\varphi_x} = \frac{v_s}{r_x} > 0, \tag{5}$$

$$\frac{\partial y}{\partial s} = \frac{-\varphi_s}{\varphi_y} = \frac{v_s}{r_y} < 0, \quad \textit{and} \tag{6}$$

$$\frac{\partial y}{\partial x} = \frac{-\varphi_x}{\varphi_y} = \frac{-r_x}{r_y} > 0. \tag{7}$$

Proof. These follow from the implicit function theorem applied to (5), using the monotonicity conditions in assumption 2.1. □

Remark. Equation (6) implies that, for a fixed y, larger cities draw migrants from farther away: higher urban productivity offsets greater distance. Equation (7) shows that, for a fixed distance, higher relocation tolerance is needed to move to smaller cities. Equation (8) reveals that among marginal workers born at different distances, those farther away must exhibit higher tolerance to be indifferent.

The responsiveness of these trade-offs depends on the curvature of the relocation cost function. If $r(x, y)$ is highly sensitive to x or y, then even small changes in distance or tolerance are associated with large changes in s. Thus, the shape of $r(x, y)$ plays a central role in determining the degree of agglomeration.

Finally, note that in the absence of any centrifugal forces, all workers will move to the city and thus $s_i = 1$ for any i. In standard urban models, land scarcity or commuting costs are often introduced to prevent such a distribution. Here, relocation costs serve that purpose.

2.5 Competitive Equilibrium

We define equilibrium as follows:

Definition 2.3 Competitive Equilibrium

An equilibrium is a set of feasible allocations $\big(c(x, y),\ l(x, y)\big)$, a city size s, and a wage w such that:

1. *For each consumer with (x, y), the allocation $\big(c(x, y),\ l(x, y)\big)$ maximizes utility given w and s, and*
2. *The population size s satisfies the fixed-point condition:*

$$s - \iint_{M(s)} f_X(x) f_Y(y) dx dy = 0, \tag{8}$$

where $M(s) = \{(x,\ y) :\ \varphi(x, y, s) \geq 0\}$ is the measure of consumers for whom migrating to the city yields higher utility than staying.

This equation ensures that the realized city size equals the measure of individuals who choose to migrate there. While it defines equilibrium size implicitly, it does not generally admit a closed-form solution. We adopt a heuristic approach instead to study how imperfect mobility shapes agglomeration.

Suppose a representative consumer drew (x, y). If $\varphi(x, y, s) > 0$, she chooses to migrate, implying upward pressure on s; if $\varphi(x, y, s) < 0$, she opts not to, putting downward pressure on s. In equilibrium, these forces balance, and $\varphi(x, y, s) = 0$.

Under this logic, the probability that city size takes a particular value s corresponds to the probability that a drawn (x, y) satisfies the indifference condition $\varphi(x, y, s) = 0$. For instance, the probability that $s \geq \bar{s}$ equals the probability that $\varphi(x, y, \bar{s}) \geq 0$.

We formalize this mapping by treating s as a function of two random variables x and y distributed as $f_X(x)$ and $f_Y(y)$. Let (x, y) satisfy the condition $\varphi\big(x, y, s(x, y)\big) = 0$, and define the change of variable:

$$t = x, \quad s = s(x, y), \tag{9}$$

with inverse

$$x = x(t, s), \quad y = y(t, s). \tag{10}$$

Since $\varphi(\cdot)$ is strictly monotonic in x, y and s, the mapping is bijective.

Proposition 2.2 City-Size Distribution

Under transformation (10) and (11), the distribution of equilibrium city sizes $f_S(s)$ is given by:

$$f_S(s) = \int_0^{\frac{1}{2}} f_X(t) f_Y\big(y(t,s)\big)\, \frac{v_s(s)}{-r_y\big(t,\ y(t,s)\big)} dt \tag{11}$$

in equilibrium.

Proof. Joint pdf of s and t is

$$f_{TS}(t,s) = f_{XY}(x,y)|\det J| = f_X\big(x(t,s)\big)\; f_Y\big(y(t,s)\big) \left| \frac{-\partial y(t,s)}{\partial s} \right| \tag{12}$$

by a change of variables, where J denotes Jacobian $\frac{\partial(x,y)}{\partial(t,s)}$ of (10). Apply the implicit function theorem to (5) to replace $|-\partial y(t,s)/\partial s|$ in (13) with $|-v_s(s)/r_y(x,y)|$. It is equal to $-v_s(s)/r_y(x,y)$ for (7). Finally, marginalize t out to obtain (12). $\square$

Remark. Appendix A.1 derives an alternative form of (12).

Equation (12) relates agglomeration to mobility:

Proposition 2.3 Pass-Through

The variance of city size increases as the relocation cost function $r(x,y)$ becomes more responsive to x or y. This amplification is especially pronounced for low values of x (nearby individuals), where migration responses are more elastic.

Proof. In (13), $|\det J|$ becomes smaller when $r_y(x,y)$ becomes larger in magnitude, i.e., when $r(x,y)$ becomes more sensitive to y. Then a region in $X \times Y$ maps to a larger region in $S \times T$ in (10), rendering the distribution of s more spread out. See appendix A.1 for the proof in terms of x. $\square$

Imperfect mobility introduces dispersion in the city size distribution via two channels:

1. Heterogeneity in (x,y): If relocation tolerance and distance are homogeneous, $f_S(s)$ collapses to a point mass.
2. Relocation friction function $r(x,y)$: A more responsive cost function amplifies the impact of heterogeneity in (x,y).

Thus, $r(x,y)$ governs how much of the joint distribution $f_{XY}(x,y)$ passes through to the equilibrium distribution of city sizes.

2.6 Tolerance Cutoff

To assess the impact of imperfect mobility on agglomeration, we would ideally observe x, y, and s, estimate the function $r(x, y)$, and compute $|\det J|$ to measure the extent to which heterogeneity in mobility shapes locational variation in agglomeration. In practice, while x and s are observable, y is not, making direct empirical validation infeasible. Against this backdrop, we focus on the threshold value of y derived in assumption 2.2. Specifically, we express $y(x, s)$ using the inverse function of $r(x, y)$ as follows:

DEFINITION 2.4 CUTOFF LEVEL OF TOLERANCE

*Define the **cutoff level of tolerance** by*

$$y(x, s) := r^{-1}\left(x, s^{\alpha}\left(s^{\frac{\alpha}{\beta-1}} + k\right)^{-\beta+1} - \bar{V}\right). \tag{13}$$

If $y \lesseqqgtr y(x, s)$, then $\varphi(x, y, s) \lesseqqgtr 0$. This threshold characterizes the minimum tolerance level required for a worker born at x to choose the city over her birthplace. Thus, from observed (x, s) pairs, we can infer the implied threshold $y(x, s)$ that separates movers from non-movers.

We restate proposition 2.3 in terms of the cut off tolerance:

PROPOSITION 2.4 CUTOFF LEVEL AND AGGLOMERATION

If $|\partial y(x, s)/\partial s|$ declines, the variance of s increases.

Proof. Immediate from (13). ◻

Intuitively, a low value of $|\partial y(x, s)/\partial s|$ implies that a small change in tolerance is associated with a large change in city size. As shown in proposition 2.3, this flattens the distribution of s, increasing its variance.

While $r_y(x, y)$ is unobservable, $\partial y(x, s)/\partial s$ can be computed from data. A low empirical value of $|\partial y(x, s)/\partial s|$ suggests a high sensitivity of relocation costs to y, amplifying the influence of heterogeneous mobility on agglomeration outcomes.[7]

Figure 4 illustrates the relationship between the cutoff tolerance and migration behavior. The blue curve represents urban utility at x_N, while the gray line represents rural utility $\bar{V}$. The cutoff tolerance $y(x, s)$ is located where these two lines intersect. Workers drawing y above this threshold relocate to the city, while those below remain.

Lower cutoff levels increase the likelihood of migration. As evident in the figure, the fraction of movers increases as the cutoff $y(x, s)$ falls.

[7]Analogously, one could define a moving distance cutoff $x(y, s)$ and relate $\partial x(y, s)/\partial s$ to the influence of distance on agglomeration (see appendix A.1). However, empirical implementation is difficult because y is not directly observed. Consequently, we proceed by estimating $y(x, s)$ rather than $x(y, s)$ in the empirical analysis to follow.

Notably, spatial sorting occurs along both x and y. An individual with a given y_0 would relocate if born at x_N but not if born at $x_F > x_N$, reflecting the joint dependence of migration decisions on birthplace and relocation tolerance.

Turning to estimation, we model the migration flow from x to a city of size s as

$$m(x, s) = f_X(x) \int_{y(x,s)}^{\infty} dF_Y(y) = f_X(x) G_Y\big(y(x, s)\big), \quad (14)$$

where $G_Y(y)$ denotes the survival function of $F_Y(y)$. Since $G_Y\big(y(x, s)\big)$ declines with x for (8), the normalized migration flow $m(x, s)/f_X(x)$ also declines with x.

In equilibrium, city size satisfies $s - \int_0^{\frac{1}{2}} m(x, s) dx = 0$ as defined in definition 2.3. Although no closed-form solution exists, the cutoff tolerance level $y(x, s)$ can be extrapolated from observed migration flows by rearranging (15):

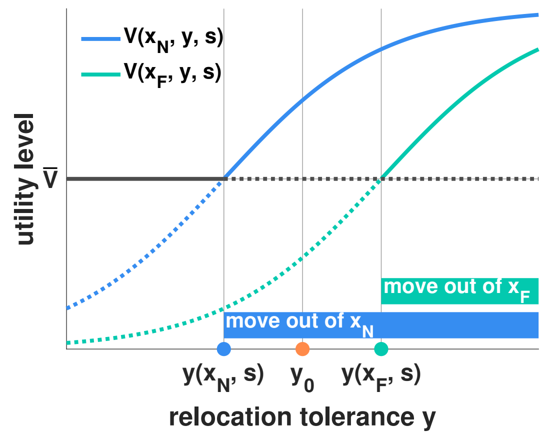


Figure 4. Urban and rural utility levels at x_N near the city and $x_F (> x_N)$ farther out. The solid line traces utility profiles across y. Workers with $y \geq y(x, s)$ relocate to the city. Note that $V(\cdot, s)$ shifts upwards as s increases.

$$y(x, s) = G_Y^{-1}\left(\frac{m(x, s)}{f_X(x)}\right). \quad (15)$$

To implement (16), the form of $G_Y(y)$ must be specified. Given that y likely reflects a sum of multiple individual factors (e.g., job stability, risk tolerance, housing tenure), we appeal to the Central Limit Theorem and assume y follows a standard normal distribution. Nonetheless, (16) holds under any strictly increasing $F_Y(y)$.

Homogeneous mobility emerges as a special case where $y(x, s) \to -\infty$, meaning all workers relocate regardless of tolerance. In this limit, $\partial y(x, s)/\partial s \to 0$, leading to degenerate $f_S(s)$ and uniform city sizes.

The present framework attributes variation in agglomeration not to external differences across cities, but to internal frictions arising from heterogeneous mobility among workers. In the next section, we compute empirical cutoff values and quantify their impact on agglomerative patterns, leveraging the theoretical foundation developed here.

3 Empirical Validation

3.1 Data Employed

To test our model predictions, we use the US Census Bureau's American Community Survey (ACS) 2009-2013.[8] In the span of five years, ACS surveyed a sample of 10.6 million households nationwide.

[8] Data available from US Census.

The questionnaire asked respondents where they lived a year prior to the survey, along with their residence at the time of survey. Migration records are tabulated by MSA, totalling 381·380=144,780 entries of inflow and outflow traced between city pairs.[9]

3.2 Correlation between Tolerance Cutoff and Agglomeration

We locate the cutoff tolerance value using specification (16). We discretize distances by origin-destination pairs and scale inflows by the originating population. To match the model's unit mass assumption, we normalize standardized inflows so that the destination with the highest total inflow attains a value of one. This normalization allows us to estimate cutoff tolerances consistently across cities.

Figure 5 plots the cutoff values across the 381 MSAs in the sample. The cutoff values

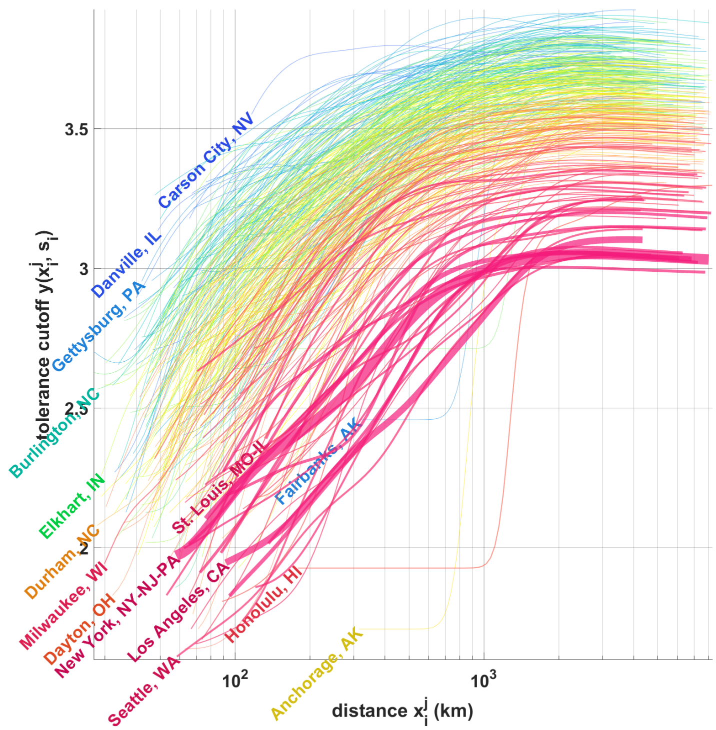


Figure 5. Tolerance cutoff (kernel smoothed).

differ systematically with destination size, consistent with proposition 2.4. The correlation coefficient between the cutoff and the log of destination size is -0.3968 with t-statistic -95.22. It is unlikely that heterogeneity in mobility and variations in agglomeration are independent.

[9] The US terrain is not a circle of unit perimeter as set up in the preceding sections. We discuss reconciliation between the model and data in appendix A.2.

That said, tolerance cutoff is a function of s **and** x. To isolate the role of distance, we approximate

$$y(x,s) \approx y_x(x_0,s_0)\,x + y_s(x_0,s_0)\,s + \frac{1}{2} y_{xx}(x_0,s_0)\,x^2 + \bar{y} \tag{16}$$

about some $(x_0, s_0) \in \left[0, \frac{1}{2}\right] \times \mathbb{R}_+$, where $\bar{y}$ is a constant.

	1	2	3
constant	4.225	0.2122	5.662
	(205.08)	(3.36)	(318.08)
$\log x$	0.2119	1.448	
	(116.72)	(78.01)	
$(\log x)^2$		-0.09267	
		(66.89)	
$\log s$	-0.1693	-0.1707	-0.1611
	(-123.38)	(-130.02)	(-120.31)
in-state			-0.6949
			(-125.97)
R^2	0.3421	0.3977	0.3651
adjusted R^2	0.3421	0.3976	0.3651

Table 1. The numbers in brackets denote t-statistics. All coefficients are significant at 1%. The log specifications help distribute data points more evenly.

Table 1 summarizes regressions of the cutoff value on distance and size. We emphasize that these regressions serve only to gauge the marginal effect of moving distance.

Under perfect mobility, the coefficient on moving distance would be zero. The data reject this: moving distance significantly affects relocation decisions.

Moreover, the signs predicted in proposition 2.1 align with our estimates: $y_x > 0$ and $y_s < 0$. Distance raises the cutoff, while destination size lowers it. In addition, $y_{xs} \approx 0$ (see figure 6), namely, the cutoff values run parallel to each other in figure 5.[10] If this value was positive, smaller cities would have a wider catchment area than larger cities. The data reject this as well.

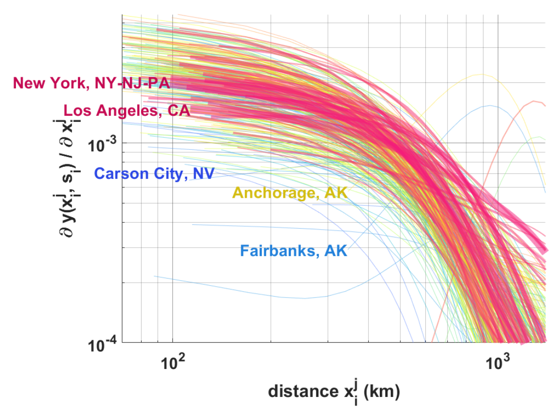


Figure 6. The slope of tolerance cutoff.

The coefficient on $(\log x)^2$ is negative, consistent with the assumption in assumption 2.1 that relocation costs grow at a decreasing rate with distance. Thus, cutoff sensitivity to distance diminishes at longer distances. This result provides theoretical justification for the common

[10]The coefficient of correlation between $\partial y(x,s)/\partial x$ and s comes to -0.01067 and it is not significant at the 5% level.

empirical practice of grouping moves by broad distance categories, such as local, intrastate, or interstate moves—as reflected in the strong predictive power of the in-state indicator in column 3 of table 1.

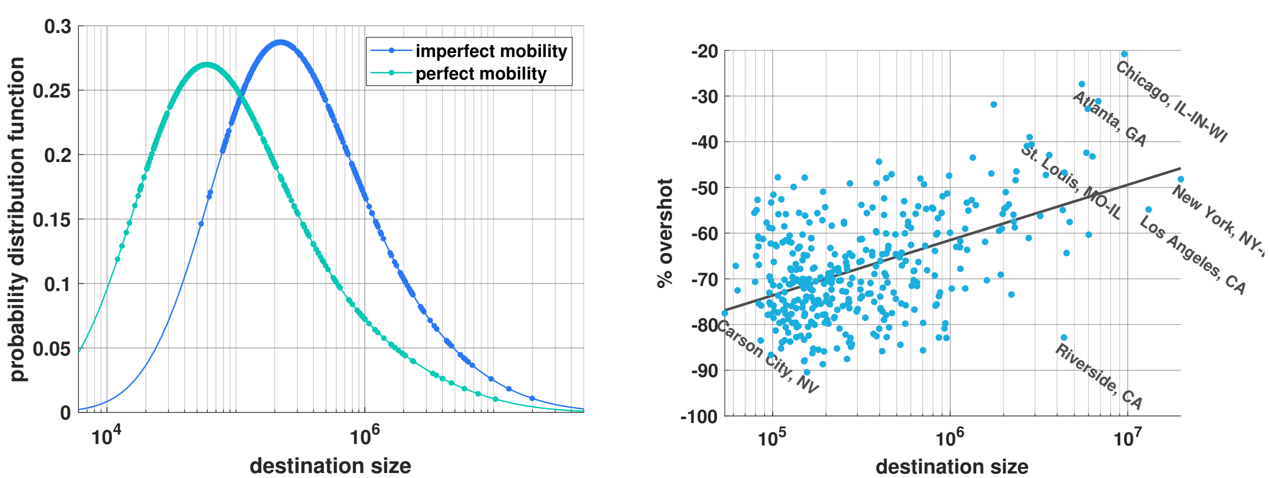
As a robustness check, we replicate the analysis using a different empirical strategy: estimating the gradient of inflow decay by city and linking it to city size. The results, presented in appendix A.3, confirm that cities drawing migrants from a broader range of distances tend to be larger, reinforcing the role of heterogeneous distance tolerance in shaping agglomeration patterns.

Overall, the empirical results validate the key comparative statics of the model: relocation frictions matter, and their effects vary systematically with both distance and agglomeration size.

3.3 Agglomeration under Homogeneous Mobility

We conduct a counterfactual exercise to isolate the role of imperfect mobility in shaping agglomeration patterns. Specifically, we consider an alternative scenario where migration decisions are independent of distance, emulating perfect mobility.

In this counterfactual, origins are randomly reassigned across locations. If distance plays no role, the inflow from any origin to a destination would be invariant to the origin's actual location. Permuting origins across the dataset effectively filters out the distance-driven component of migration, allowing us to reconstruct hypothetical destination sizes under perfect mobility.



(a) Size distribution under perfect and imperfect mobility.

(b) Percentage difference between the size under perfect mobility (counterfactual) and the size under imperfect mobility (actual).

Figure 7.

Figure 7 compares the actual and counterfactual city-size distributions. Figure 7(a) shows that variance in destination sizes declines under perfect mobility: the standard deviation falls by 59.1%, from 1.62 million to 0.896 million. Thus, imperfect mobility magnifies disparities in city sizes.

Figure 7(b) plots the relative size difference between the actual and counterfactual cases. Larger cities experience smaller relative size reductions, while smaller cities lose a larger share of their population. This pattern suggests that spatial frictions disproportionately bolster smaller destinations by limiting access to larger alternatives.

While our counterfactual focuses on physical distance, we speculate that similar patterns would emerge with respect to social or economic relocation frictions, which are not directly observable in the data. Investigating such frictions remains an avenue for future research.

4 Conclusion

This paper develops a framework linking workers' distance tolerance to the equilibrium distribution of city sizes. Each worker draws a distance tolerance from a common distribution and chooses whether to relocate to a city to access urban productivity. In equilibrium, migration decisions reflect both birthplace and distance aversion, allowing the city-size distribution to be traced back to fundamental micro-level heterogeneity.

The model generates two core predictions: first, that variation in distance tolerance underpins variation in agglomeration intensity across cities; second, that individuals perceive distance on a nonlinear, approximately logarithmic scale. This finding offers a microfoundation for the widespread empirical practice of categorizing migration into local, intrastate, and interstate moves. Both predictions are supported by U.S. migration data. Empirically, a small subset of highly mobile workers disproportionately drives the formation of large cities, while most workers remain geographically anchored near their birthplaces.

While relocation tolerance is unobservable, our framework enables inference from observed migration patterns. The cutoff value, the minimum tolerance needed to justify moving, can be recovered using equilibrium inflow data and the distribution of distances. This inverse mapping links the data to the theoretical structure and enables empirical validation of the model's key mechanisms.

Counterfactual analysis shows that removing distance aversion would significantly compress the city-size distribution, leading to smaller and more uniform cities. Imperfect mobility, therefore, acts as an amplifying force behind urban concentration and the heavy tail of the city-size distribution.

The model abstracts from several features of real-world migration dynamics. We assume each city hosts a single worker type, whereas in reality cities accommodate diverse populations. Urban productivity is not linked to migrant characteristics. Moreover, we analyze a static snapshot rather than a dynamic evolution of migration and urban growth.

Extending the framework to allow for type co-location, endogenous productivity accumulation, and dynamic feedback would offer fruitful directions for future research.

Finally, while the model assumes independence between birthplace and distance tolerance, real-world patterns such as corporate transfers or regional cultural differences may introduce systematic correlations. Relaxing this assumption would refine our understanding of how micro-level frictions aggregate into macro-level urban patterns.

A Appendix

A.1 City-Size Distribution in an Alternative Form

By substituting the transformation (10) with $z = y$ and $s = s(x, y)$; (11) with $x = x(z, s)$ and $y = y(z, s)$, we obtain an alternative representation of the city-size distribution in (12):

$$f_S(s) = \int_{\mathbb{R}} f_X\big(x(z, s)\big) f_Y(z) \frac{v_s(s)}{r_x\big(x(z, s),\, z\big)} dz. \tag{17}$$

This form is analogous to that in proposition 2.2, except that the roles of x and y are re-parameterized to facilitate analysis of how the geographic distribution of population shapes $f_S(s)$. Just as y governs heterogeneity in individual relocation tolerance, x governs spatial frictions through the relocation cost $r(x, y)$.

In addition, assumption 2.1 implies that the cost increase tapers off with distance x. Therefore, if x is high, it requires a broader difference in x to reduce $f_S(s)$ by the same amount that a low x does, to make up for the dwindling impact of distance on $r(x, y)$.

This representation is useful for analyzing how sensitive agglomeration patterns are to the initial spatial distribution of the population $f_X(x)$. For example, if we hypothetically relocate a large inland city (e.g., Denver or Phoenix) to a denser coastal region, the mass of $f_X(x)$ would shift, potentially amplifying the size of the city due to higher surrounding population density. Equation (18) allows us to evaluate such geographic counterfactuals explicitly.

A.2 Adjustments for the Shape of Country

The theoretical model assumes a circular country where all cities are equidistantly reachable with a maximum distance of $1/2$. In contrast, the actual U.S. landmass is irregularly shaped, and the maximum distance to each MSA varies. For example, Carson City, NV, has a maximum range of 4,187 km (from Bangor, ME), while the greatest possible distance—between Honolulu and Bangor—reaches 8,293 km.

This geographic asymmetry could bias estimates of distance tolerance: workers located beyond a city's feasible maximum range cannot relocate there, even if their tolerance would otherwise allow it. As a result, cities like Carson City might appear smaller than predicted, while cities like Honolulu might appear larger.

In practice, we find no statistically significant relationship between agglomeration and placement of cities.[11] This suggests that geographic bounds do not meaningfully distort city sizes, particularly since distance is perceived logarithmically as discussed in (2), compressing the perceived difference between long and very long distances.

A.3 Gradient of Cutoff Tolerance

To validate the cutoff estimation, we examine whether cities that draw migrants from greater distances tend to be larger. For each MSA, we regress inflow on $\log x$ to estimate two parameters: α_i (intercept) and β_i (distance elasticity). See figure 1(b). We then estimate the following relationship: $\log s_i = \gamma_0 + \gamma_1 \alpha_i + \gamma_2 \beta_i$. The estimated γ_2 is significantly positive, indicating that cities with flatter inflow decay curves, i.e., those that retain inflows over longer distances, are systematically larger. This aligns with the model's prediction that variation in relocation tolerance contributes to differences in city size.

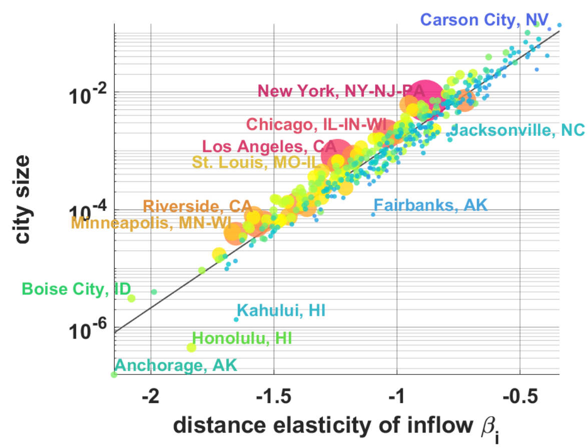


Figure 8. City size (controlled for α_i) over distance elasticity of inflow β_i. Dot size and color are proportional to city size. *Data source*: U.S. Census Bureau, ACS 2009–2013.

To reinforce this point, we also regress city size on the mean and standard deviation of inflow distances. Both are positively associated with population size: large cities not only draw from farther away but also from a more varied range of locations.

These results are visualized in figures 2 and 8 and are inconsistent with perfect mobility, which would imply invariant inflow patterns across cities.

Observed deviations in remote MSAs (e.g., in Alaska and Hawaii) appear to stem from geographic constraints rather than economic mechanisms. Including maximum feasible distance as a control confirms that our main results are not driven by boundary effects.

[11] To test whether geographic bounds bias city size estimates, we regressed log city size on the maximum distance from each city to any origin. The result,

$$\text{size} = \underset{(5.17)}{12.61} + \underset{(.02)}{.006}\ \text{max range}, \quad (R^2 = .0000)$$

(t-statistics in the parentheses) shows no significant relationship.

References

[AA14] Treb Allen and Costas Arkolakis. Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3):1085–1140, 2014.

[BBT92] George J Borjas, Stephen G Bronars, and Stephen J Trejo. Self-selection and internal migration in the united states. *Journal of Urban Economics*, 32(2):159–185, 1992.

[BC97] John H Boyd, III and John P Conley. Fundamental nonconvexities in Arrovian markets and a Coasian solution to the problem of externalities. *Journal of Economic Theory*, 72(2):388–407, 1997.

[CDGR12] Pierre-Philippe Combes, Gilles Duranton, Laurent Gobillon, and Sébastien Roux. Sorting and local wage and skill distributions in france. *Regional Science and Urban Economics*, 42(6):913–930, 2012.

[Dia16] Rebecca Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *American Economic Review*, 106(3):479–524, 2016.

[FHLS12] Oliver Falck, Stephan Heblich, Alfred Lameli, and Jens Südekum. Dialects, cultural identity, and economic exchange. *Journal of Urban Economics*, 72(2-3):225–239, 2012.

[Gre24] Allison Green. Networks and geographic mobility: Evidence from World War II navy ships. Working Paper, 2024.

[Hel97] John F Helliwell. National borders, trade and migration. *Pacific Economic Review*, 2(3):165–185, 1997.

[HT70] John R Harris and Michael P Todaro. Migration, unemployment and development: a two-sector analysis. *The American Economic Review*, 60(1):126–142, 1970.

[JMSW23] Ning Jia, Raven Molloy, Christopher Smith, and Abigail Wozniak. The economics of internal migration: Advances and policy questions. *Journal of Economic Literature*, 61(1):144–180, 2023.

[KW11] John Kennan and James R Walker. The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251, 2011.

[McF74] Daniel McFadden. Conditional logit analysis of qualitative choice behavior. In Paul Zarembka, editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, 1974.

[Mor12] Enrico Moretti. *The new geography of jobs*. Houghton Mifflin Harcourt, 2012.

[RS04] Stuart S Rosenthal and William C Strange. Evidence on the nature and sources of agglomeration economies. In *Handbook of regional and urban economics*, volume 4, pages 2119–2171. Elsevier, 2004.

[Sja62] Larry A Sjaastad. The costs and returns of human migration. *Journal of Political Economy*, 70(5, Part 2):80–93, 1962.

[Sta78] David Starrett. Market allocations of location choice in a model with free mobility. *Journal of Economic Theory*, 17(1):21–37, 1978.

[Tob70] Waldo R Tobler. A computer movie simulating urban growth in the detroit region. *Economic Geography*, 46(sup1):234–240, 1970.

[Woo11] Colin Woodard. *American nations: A history of the eleven rival regional cultures of North America*. Penguin, 2011.