The Size Distribution of Cities with Distance-Bound Households

Watanabe, Axel

Concordia University and CIREQ

17 April 2020

The Size Distribution of Cities with Distance-Bound Households

Axel Watanabe*

October 8, 2020

Abstract

There has been a long tradition of presumed perfect mobility in urban economics. Workers switch their locations in direct response to differences in local economic performance. Recent empirical observations prove otherwise. The number of movers rapidly declines with distance while there is a positive correlation between distance moved and city size. I build a general equilibrium model of a system of cities to explain the city-size distribution as a result of imperfect mobility. Consumers' logarithmic perception of distance makes the city-size distribution heavy tailed. I also extrapolate how tolerant residents are to distance in each US city from the data on city size and interurban migration.

Keywords: geographic mobility, internal migration, city-size distribution
JEL classification: J61, R12

1 Introduction

1.1 Consumers Are Not Footloose

Labor mobility exhibits distinct geographic patterns. There is a log linear relationship between the number of incoming residents and their distance moved as

*Department of Economics, Concordia University, and CIREQ. Address: 1455 boul. de Maisonneuve O, Montréal (Québec) H3G 1M8. E-mail: axel.watanabe@concordia.ca. I am grateful for valuable inputs from Editor Jan Brueckner, two anonymous referees, Xintong Han and Prosper Dovonon. 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.

can be seen in figure 1. An exceeding share of domestic migration occurs within a close proximity and there are only a few who move coast to coast. Take St. Louis for example. The vast majority of incoming residents are from Missouri and Illinois when in fact workers are free to move anywhere in the country. The inflow drops at an exponential rate with distance. When the distance increases by 1%, the inflow from that area drops by 1%.

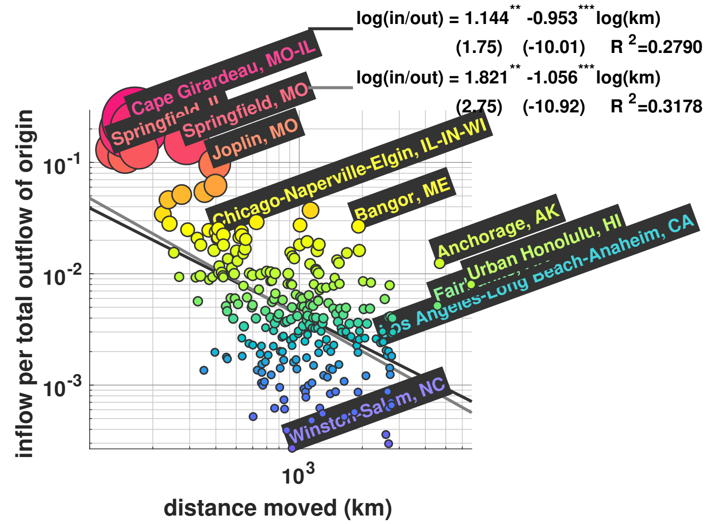
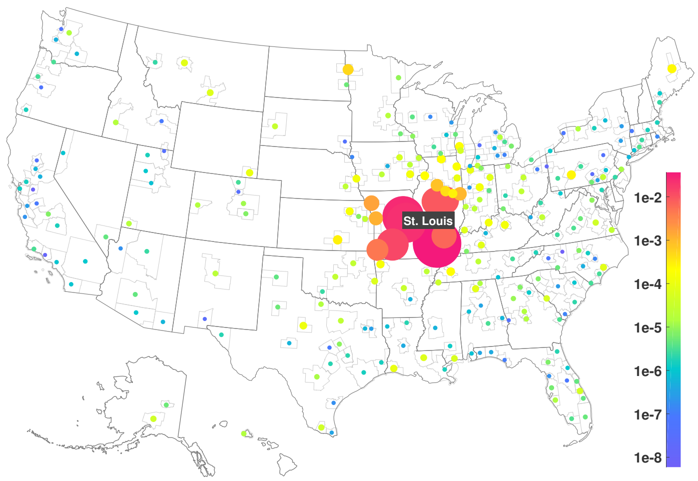


Figure 1. Number of in-migrants to St. Louis by distance. Colors and dots are size proportionate. Two lines represent ordinary least squares regression. Black line includes Alaska and Hawaii. Gray line does not. *** and ** denote coefficient significant at 1% and 5% respectively.

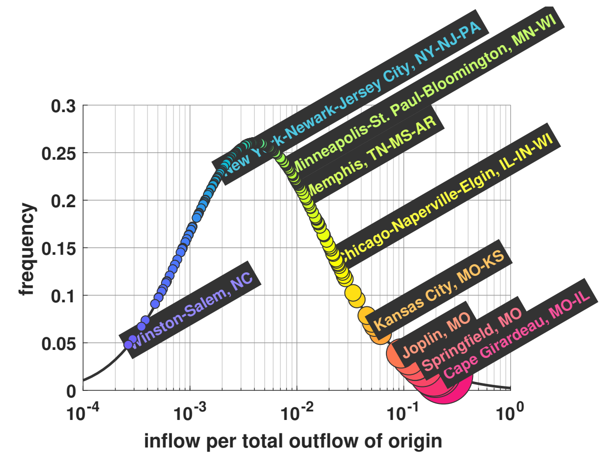
The city-size distribution is a result of household relocation. Any city size is the sum of the inflows into, less the outflows out of, the city over time. It is then logical to speculate that mobility weighs in on its determination. It is known that the city-size distribution has a heavy tail (cf. figure 2(c), Gabaix and Ioannides [GI04], and Duranton [Dur07]). Intercity migration itself has a heavy tail as well, from which the city-size distribution is derived (compare figure 2(b) to figure 2(d)).

Models of the city-size distribution traditionally assume perfect mobility. Workers move to another city in direct response to local economic conditions regardless of how far their destination is. The resulting city-size distribution is independent of where workers were in the period before.

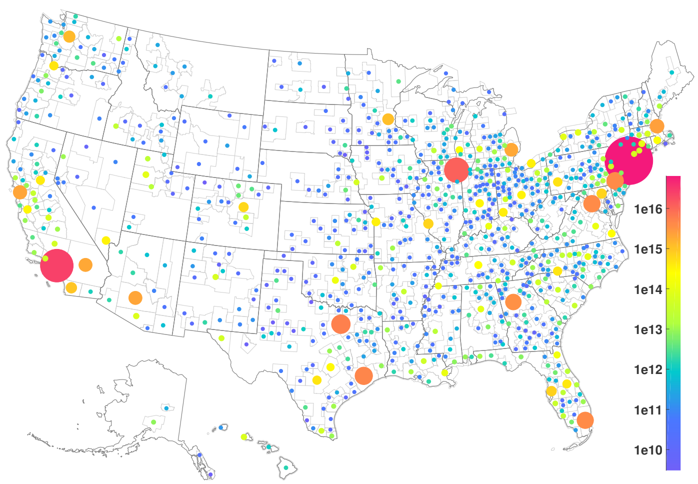
This paper aims to explain the city-size distribution as a result of the observed intercity migration patterns above. In a general equilibrium model of a system of cities, workers draw their type and tolerance level for relocation. They make their location choices based on their type and how far they were born from their destination city. The equilibrium city size depends on their attitude towards relocation. Empirical estimations reveal that workers perceive distance on a logarithmic scale. Consequently, the majority of cities source their residents only from within its immediate vicinity and thus remain small. However, there are certain types of workers who do not mind moving far in favor of earning opportunities in a large city. These types are rare but most of them **will** come to the city because distance



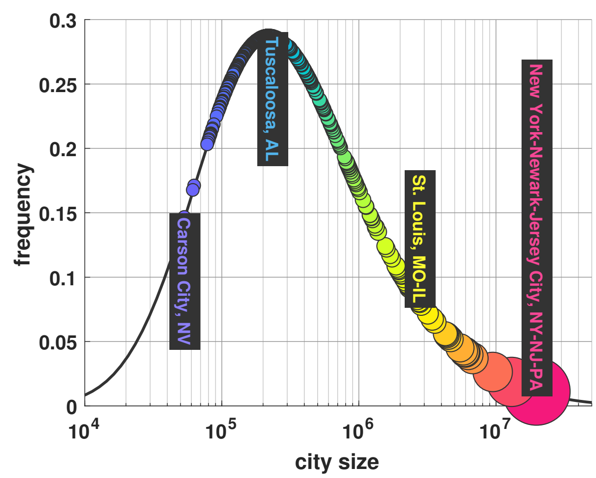
(a) Inflow into St. Louis metropolitan statistical area (MSA) from other MSA's normalized by the total outflow of origin.



(b) Probability density function of figure 2(a).



(c) Size distribution of core-based statistical areas.



(d) Probability density function of figure 2(c).

Figure 2. *Data source*: US Census Bureau, 2009-2013 American Community Survey. Colors and dots are size proportionate.

plays only a limited part in their location decision. Consequently, the city gathers workers from across the country and becomes gravitationally large. The resultant city-size distribution features a large number of small cities made up of nearby in-migrants, paired with a small number of exceptionally large cities filled with globally oriented workers.

1.2 Related Literature

Geographic mobility literature theorises about the probability of internal migration as a function of associated net gains (e.g., Sjaastad [Sja62] or Harris and

Todaro [HT70]). Various factors involved in relocation choice have been studied (see Molloy et al. [MSW11]). Davis and Dingel [DD12] and Rauch [Rau13] speculate on the heterogeneous search behavior by workers of different skill levels as a possible cause. Falck et al. [FHLS12] turn to regional and cultural factors. They document Germans' reluctance to move outside of their shared area of regional dialect. Woodard [Woo11] suggests similar cultural divides in the US.

On the empirical side, the primary focus is on whether a consumer moves or not, but not on by how far he moves. In Bowles [Bow70], there are two distances: whether a consumer is in the South or not. In Ladinsky [Lad67], there are four. I measure the distance between every pair of 381 metropolitan statistical areas (MSA), totalling 72,390 distinct distances, which enables me to examine the exact role that distance plays. Haag et al. [HMP+92] feature the city-to-city distance in France as well. Their work differs from the current paper as they work with a reduced form. Moreover, the literature concerns about the identification of relevant causes for relocation, but not necessarily about the resulting city-size distribution.

By contrast, the city-size distribution has been one of the primary areas of focus in urban economics.

Urban economics usually assumes no relocation cost (cf. Starrett [Sta78], and Boyd and Conley [BC97]), and in labor economics, relocation cost is usually a fixed cost that does not depend on the distance moved. For instance, in Manning [Man10] and Hirsch et al. [HJO16], market imperfections lead to reduced mobility in terms of type-matched industry, but not geographic mismatch.

The present paper is based on the works of Behrens et al. [BDRN14] and Eeckhout et al. [EPS14]. Behrens et al. show that workers sort into a city and select their occupation according to their skill level. Along with skill levels, location-variant serendipity determines the productivity and in turn the degree of agglomeration in each city. Eeckhout et al. find evidence in support of extreme skill complementarity where the co-presence of workers from top- and bottom-tier skill levels does not undermine but rather enhances their productivity. Migration in the present paper is also motivated by heterogeneous skills. However, urban productivity is simplified in the interest of incorporating distance-dependent relocation costs.

A myriad of socioeconomic and political factors are involved in the determination of city size. Considering that the city-size distribution is the upshot of these individual cities that are already convoluted in and of themselves, it is unlikely that one factor can single-handedly explain it all. Urban economists pick one factor of interest and examine its explanatory power such as random growth (Eeckhout [Eec04]), transportation cost (Berliant and Watanabe [BW18]) or collective decision (Duranton and Puga [DP19]) while keeping other factors such as imperfect mobility turned off. Here, I will tune in to imperfect mobility and tune out other factors. The ultimate model of the city-size distribution is likely an amalgam of all these models. I intend to add a model with a yet unexplored perspective to the pool of existing models that, when combined, yield a comprehensive description of the subject.

1.3 Reasons for Imperfect Mobility

There are many reasons for imperfect geographic mobility. For example, the geographic extent of job search expands with the level of skill. Ph.D students on the job market fly everywhere for interviews, whereas it is not likely to see high school graduates doing the same. According to the US Census, 35% of degree holders moved for employment reasons whereas that of non high school graduates is only 13%.[1] The same goes for the receiving end (cities) as well. 30% of the residents in the 100 largest core-based statistical areas hold BA or Ph.D, whereas that of the smallest 100 is 19%.

Along with the heterogeneous skill levels, uncertainty aversion may deter long-distance relocation. It is usually difficult to know the quality of life in a new city in advance. Furthermore, even if a worker herself may be mobile per se, it is prohibitively costly to move the entire network of people she meets in her daily life. The pecuniary cost of relocation is a one-time expense, but out-of-towners may incur implicit costs as above over a long period of time, be it personal, social, cultural or economic.

The rest of the paper is organized as follows: In the upcoming section I will lay out the model and uncover the relationship among distance, inflow and city size. I will empirically validate my theoretical predictions in sections 3 and 4, interpret them in section 5, and summarize them in section 6.

[1] Data source: https://www.census.gov/data/tables/2015/demo/geographic-mobility/cps-2015.html.

2 Model

2.1 Landscape

Consider a closed production economy. I take the country to be a circle with its perimeter normalized to 1 in order to remove the border that may otherwise produce an unwanted asymmetric result. Cities line up along the perimeter X. There are I of them indexed by i.

The model rolls out in two stages. Initially, there are $N\,(\in \mathbb{R}_+)$ consumers uniformly distributed over X. Each consumer is endowed with a pair $(t,\ y) \in \{1, \cdots, T\} \times Y\ (\subseteq \mathbb{N} \times \mathbb{R})$. The first entry is type t representing her skill. It identifies her best suited industry to work in. There are n_t of type-t consumers. Along with the type, she also draws her distance-tolerance factor y from the distribution with probability density function (pdf) $f_t(y)$ and cumulative distribution function (CDF) $F_t(y)$. A high y implies that she does not mind moving far. Note that $\int_Y f_t(y)dy = n_t$ for all t, totaling up to $\sum_t \int_Y f_t(y)dy = N$ nationwide.

Type distribution $f_t(y)$ may depend not only on t but also on birthplace. However, since consumers cannot choose a place to be born in,[2] it is safe to assume that $f_t(y)$ takes the same form regardless of the location.

In the second stage, consumers of type t make simultaneous and uncoordinated decisions on their location. A type-t consumer can either stay at her initial location or move to the city that matches her type. Following the example of Eeckhout [Eec04], each city i produces one commodity c_i. For simplicity, I assume that there are as many cities as there are types, that is, $I = T$. From here on I refer to type t by its corresponding city i and use the term "city", "type" and "industry" interchangeably where applicable.

I write $x_i (\in X)$ to mark the birthplace of a type-i consumer measured by the shorter arc length from city i. If I place this country on a compass with city i facing the east, $x_i = \frac{1}{4}$ is found on the north or south and $x_i = \frac{1}{2}$ is on the west end of the country.

2.2 Consumption and Location Choice

Consider a type-i consumer born distance x_i away from city i. Her preferences over a numéraire composite consumption good c_i and housing h_i are represented

[2] I will discuss the possibility of intergenerational dependence over space in appendix A.1.

by

$$u(c_i,\ h_i) = c_i + \eta \log h_i, \tag{1}$$

where η measures the portion of her expenditure on housing. She is endowed with a unit of time, which she converts into c_i to earn wage w_i. Her budget constraint is

$$w_i \geq c_i + p_i h_i + \rho(x_i,\ y), \tag{2}$$

where $\rho(\cdot)$ measures the lifelong opportunity cost of relocation as discussed in section 1.3. I will make following assumptions regarding $\rho(\cdot)$:

Assumption 2.1 Logarithmic Perception of Distance:
For any given $y \in Y$, the opportunity cost of relocation $\rho(\cdot)$ satisfies

$$\rho(0,\ y) = 0, \tag{3}$$

$$\frac{\partial \rho(x_i,\ y)}{\partial x_i} > 0, \quad \textit{and} \tag{4}$$

$$\frac{\partial^2 \rho(x_i,\ y)}{\partial x_i^2} < 0 \tag{5}$$

over X.

Consumers' nonlinear perception of distance gives grounds for (5). A St. Louis native finds a move from St. Louis to Chicago more draining than a move from Fairbanks to Anchorage (roughly the same distance apart). The additional cost increase wears out with distance.

Coupled with assumption 2.1, I also assume that $\partial \rho(x_i,\ y)/\partial y < 0$, i.e., the higher the distance tolerance is, the lower the relocation cost will be.

In addition to finding the optimal consumption bundle, she also needs to decide whether she will move to city i or stay put at x_i. I will denote her decision by a location choice function $g_i(x_i,\ y) : X \times Y \to \{0,\ 1\}$. If she drew $(x_i,\ y)$ and decides to move out of her birthplace, $g_i(x_i,\ y) = 1$. Otherwise, $g_i(x_i,\ y) = 0$.

2.3 Feasibility

Given the location choice function, a measure of type-i residents in city i is

$$s_i = \int_X \int_Y g_i(x_i,\ y)\ dF_i(y) dx_i. \tag{6}$$

Let H denote the housing supply in each city. Using (6), define the feasible allocation in this economy as follows:

Definition 2.2 Feasible Allocation:

An allocation is a list of functions $[c_i(x_i, y), h_i(x_i, y), g_i(x_i, y)]_{i=1}^I$ *with* $c_i : X \times Y \to \mathbb{R}_+$, $h_i : X \times Y \to \mathbb{R}_+$, *location choice* $g_i : X \times Y \to \{0, 1\}$, *and output* $(z_i)_{i=1}^I \in \mathbb{R}_+^I$. *Given type-size distribution* $(n_i)_{i=1}^I$ *and distance-tolerance distribution* $[f_i(y)]_{i=1}^I$, *an allocation is feasible if*

$$\begin{aligned} s_i z_i &= \int_X \int_Y g_i(x_i, y)\, c_i(x_i, y) dF_i(y) dx, && (7)\\ H &= \int_X \int_Y g_i(x_i, y)\, h_i(x_i, y) dF_i(y) dx, \quad \textit{and} && (8)\\ s_i &\leq n_i \end{aligned}$$

for any i, *where* s_i *is defined by* (6).

2.4 Production

Turning to production, as mentioned earlier, workers supply one unit of (perfectly inelastic) labor to produce the composite goods with a constant returns to scale technology: τ units of labor produces $z_i = A_i(s_i)\tau$ units of composite goods. Contrary to what is conventionally assumed, $A_i(s_i)$ does not vary with industry i or city size s_i (unless it is zero). In particular

$$A(s_i) = \begin{cases} 1 & \text{if} \quad s_i = 0 \\ a(>1) & \text{if} \quad s_i > 0. \end{cases} \tag{9}$$

In the current model, I do not rely on productivity differences to break the otherwise uniform distribution of workers. I shut off the channel through which productivity differences bring in variations in city sizes (as documented in Rosenthal and Strange [RS04]) in order to isolate the role that distance tolerance plays (or else I will not be able to tell how much of the size difference is the result of imperfect mobility). However, I still do need to secure some incentive for residents to clump together in one location. Absent economies of localization, no one will move to a city (cf. Glaeser et al. [GKS01]). Specification (9) is the minimally invasive way to do so without introducing added complications from type-dependent productivity.

Firms are a price taker and earn zero profit in equilibrium. Thus, each worker earns

$$w_i = a. \tag{10}$$

2.5 Rural Residents

Let us turn to a resident who stays put. He becomes a Robinson Crusoe-type rural resident to lead a life under the backyard capitalism. His marginal product gets pushed back to $A(s_i = 0) = 1(< a)$ according to (9), housing consumption becomes independent of the city size, and the cost of relocation becomes zero.[3] I mark his maximized utility level by $\underline{v}_i$, which, by construction, is independent of x_i. In order to keep the model on point, assume that the land in the rural area is abundant enough and the number of people who do not move out of the birthplace will not affect the value of $\underline{v}_i$. Furthermore, assume $\underline{v}_i = \underline{v}$ for all i in order to remove arbitrariness.

2.6 Trans-Tolerance Value

Circling back to a consumer who leaves her birthplace, her indirect utility function is

$$v(p_i,\ w_i;\ x_i,\ y,\ s_i) = w_i - \rho(x_i,\ y) - \eta + \eta(\log\eta - \log p_i). \tag{11}$$

The housing market (8) clears when

$$s_i h_i = H, \tag{12}$$

from which I obtain the equilibrium rent

$$p_i = \frac{\eta s_i}{H}, \tag{13}$$

i.e, the more crowded the city becomes, the more expensive the rent per unit will be.[4] Firm's first-order condition (10) and housing market clearance (12) further simplify her indirect utility function (11) to

$$v(a,\ x_i,\ y,\ s_i) = a - \rho(x_i,\ y) + \eta(-\log s_i + \log H - 1). \tag{14}$$

The farther she is from, the lower her utility level will be, holding everything else constant. Notice the trade-off among the economies of agglomeration a, diseconomies of agglomeration $-\eta\log s_i$ and distance tolerance y. Holding the value of

[3] Location x_i is identified by the distance from city i, which itself is located at $x_i = 0$. For a non-mover, the distance moved is $x_i - x_i$ rather than mover's $x_i - 0$ so that $\rho(x_i - x_i,\ y) = 0$ following (3).

[4] Note that the expenditure on housing is always $p_i h_i = \eta$ regardless of the city size. A city resident copes with an increasing city size by reducing her lot size without changing her expenditure share of housing.

$v_i(\cdot)$ constant, if the destination city becomes crowded or the productivity boost a gets smaller, the only residents with high enough tolerance y would move to the city.

Let us revisit the location choice $g_i(x_i,\ y)$. A type-i consumer will move to her type-matched city if her utility level (14) is greater than the fallback value $\underline{v}$:

$$v(a,\ x_i,\ y,\ s_i) = a - \rho(x_i,\ y) + \eta(-\log s_i + \log H - 1) \geq \underline{v}. \tag{15}$$

A resident at the margin meets (15) with equality. Since $\rho(\cdot,\ y)$ is strictly monotone decreasing in y, one can solve (15) with equality for y to define a **trans-tolerance** function

$$y_i(x_i) := \rho^{-1}\left[x_i,\ a + \eta(-\log s_i + \log H - 1) - \underline{v}\right]. \tag{16}$$

This establishes the following:

<u>Proposition 2.1 Location Choice Rule</u>

A consumer who drew $(x_i,\ y)$ makes her location choices $g_i(\cdot)$ in reference to the trans-tolerance value $y_i(x_i)$ prevalent at her birthplace x_i as follows:

$$g_i(x_i,\ y) = \begin{cases} 0 & \text{if} \quad y \leq y_i(x_i) := \rho^{-1}\left[x_i,\ a + \eta(-\log s_i + \log H - 1) - \underline{v}\right] \\ 1 & \text{otherwise.} \end{cases} \tag{17}$$

A couple of observations on (16) and (17) are in order. First off, $y_i(x_i)$ determines the fraction of people moving to city i. Anyone who drew $y \geq y_i(x_i)$ moves out because she does not show much affinity to her birthplace or her opportunity cost of staying put is too high. On the contrary, anyone with $y \leq y_i(x_i)$ has a lot to lose by relocation and thus stays in. Therefore, the higher the trans-tolerance is, the higher the ratio of non-movers will be.

Second, $y_i(x_i)$ is increasing in x_i because $\rho(\cdot)$ is increasing in x_i. In the vicinity of city i, the number of non-movers is very small because it does not take much to turn residents into a city dweller. As a result, the borderline tolerance is very low. As the distance to city i increases, the cost of relocation bears down on consumers. They will not become a city resident as easily as before unless their tolerance is high, making the threshold high as well.

However, since $\rho^{-1}(\cdot)$ is concave in x_i, the effect of increasing trans-tolerance value becomes less pronounced as x_i increases:

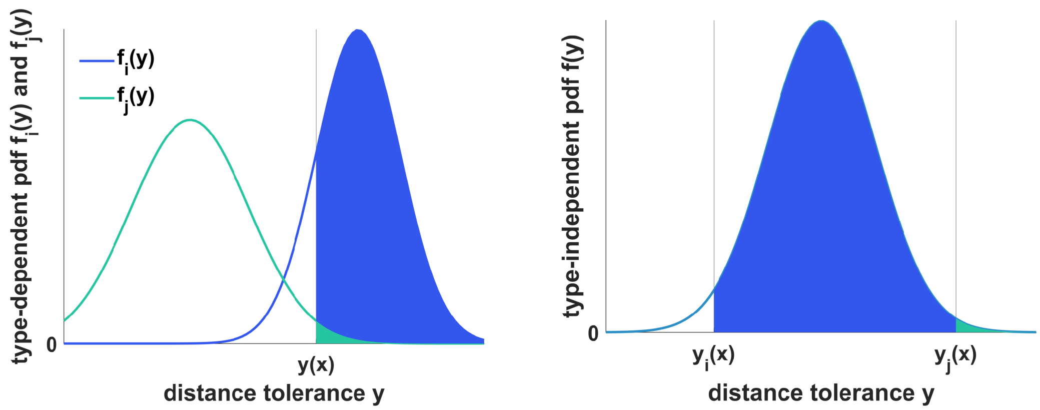
<u>Proposition 2.2 Trans-Tolerance Function Is Concave</u>

Suppose that the opportunity cost of relocation $\rho(x_i,\ y)$ is concave in x_i. Then trans-tolerance function $y_i(x_i)$ is concave in x_i as well.

Proof. Immediate from (16). ◻

The proposition hinges on assumption 2.1 that consumers perceive distance on a logarithmic scale. I will validate this supposition in section 4.

There are in fact two ways to go about the trans-tolerance value. One is to assume that the trans-tolerance function is identical over type: $y_i(x_i) = y(x_i)$ for all i. The other is to allow $y_i(x_i)$ to take different values depending on the type. I will explain the difference between them below.



(a) Type-**independent** trans-tolerance with type-**dependent** distribution of y.

(b) Type-**dependent** trans-tolerance with type-**independent** distribution of y.

Figure 3. Two possible specifications for the distribution of distance-tolerance value and the trans-tolerance function at some location $x_i = x_j = x \in X$. The shaded area represents a measure of workers who move distance x to live in their type-matched city i (in blue and green) or j (in green). The remainder represents those who stay distance x away from their type-matched city (i.e., non-movers). The outflow from location x_i is larger than that from location x_j in both scenarios.

First suppose that $y_i(x_i) = y_j(x_j) = y(x)$ for all i, j at any $x_i = x_j = x \in X$. If $y_i(x)$ will be the same across the types, then y should have been drawn from different distributions depending on the type as in figure 3(a) (or else the city-size distribution will be uniform). In this case, if $f_i(y)$ first-order stochastically dominates $f_j(y)$, then $s_i > s_j$ (cf. proposition 2.3 below). Type i should be more distance-tolerant than type j so that at any given x, more of type i must have drawn $y \geq y(x)$ than type j. In this case, y can be interpreted as a skill level that indicates the favorable degree of concentration of workers. Industry j features low-skill labor that does not benefit from concentration of workers within the same industry. Consequently, their distance tolerance is drawn from the distribution with a low mean. By contrast, industry i involves a type of workers who capitalize on large-scale interactions among them.

Alternatively, type can be thought of as a manifestation of risk tolerance. If $f_i(y)$ first-order stochastically dominates $f_j(y)$, then type i is made up of those who are willing to take on new challenges in an unfamiliar city. Large cities are large because they attract adventurous types, who are ready to go through a long-distance relocation process and consequently, the total inflow is large.

Now suppose instead that $y_i(x)$ can differ from $y_j(x)$ at some $x_i = x_j = x \in X$ but y itself is drawn from the identical distribution $f(y)$ regardless of the type as in figure 3(b). In this case, if $y_i(x) < y_j(x)$, then $s_i > s_j$. The variation in city size arises directly from trans-tolerance (16) itself, rather than the distribution from which y is drawn. City i has a larger influx of people because the net effect of agglomeration $a - \eta \log s_i$ is large enough to make up for the fallback utility level $\underline{v}$. Once again, type i is likely to be a high-skill type whereas industry j does not call for much concentration of workers.

Empirically speaking, I cannot tell which one is at work because I do not have direct observations of $f_i(y)$ or $\rho^{-1}(\cdot)$. To cover all bases, I will consider both type-dependent and -independent trans-tolerance for the empirical analysis in sections 3 and 4. For the theoretical analysis to follow, I will take type-independent trans-tolerance (figure 3(a)) as an example but the same argument goes for figure 3(b) as well.

Most of the models of city-size distributions can be thought of as a limiting case of the present model where trans-tolerance tends to negative infinity ($y_i(x) \to -\infty$) so that everyone moves out of their place of birth no matter how far they are from their destination city. This can happen in a couple of different ways. Looking at (16), if I remove the concept of distance, that is, if the distance to the city is the same (typically zero in the literature) from anywhere, then no one bears the cost of relocation $\rho(0, \cdot) = 0$ so that for sufficiently low $\underline{v}$ everyone moves to the city and the size distribution turns uniform. Alternatively, if a becomes dominant, everyone moves to their type-concordant city. The resulting city-size distribution $(s_i)_{i=1}^I$ becomes the (exogenous) type distribution $(n_i)_{i=1}^I$ itself. Existing models endogenously derive the city-size distribution using other factors of choice than imperfect mobility to frame agglomeration (something more complex than a rudimentary urban productivity defined in (9)).

Put differently, existing models start from stage two (and usually have more steps to follow), whereas the current model focuses on the transition from stage one to two.

2.7 Competitive Equilibrium and City Size

Given the trans-tolerance function above, define

Definition 2.3 Competitive Equilibrium:

An equilibrium is a feasible allocation $[c_i(x_i, y), h_i(x_i, y), g_i(x_i, y)]_{i=1}^I$ *and* $(z_i)_{i=1}^I$*, and price system* $(p_i)_{i=1}^I \in \mathbb{R}_+^I$*, such that* $[c_i(x_i, y), h_i(x_i, y), g_i(x_i, y)]_{i=1}^I$ *maximizes the utility level and* $(z_i)_{i=1}^I$ *maximizes the profit under* $(p_i)_{i=1}^I$ *for any* $(x_i, y) \in X \times Y$ *and* $i \in \{1, \cdots, I\}$.

With the equilibrium conditions above, goods market (7) is written as

$$s_i = \frac{1}{a} \int_X \int_{y(x_i)}^{\infty} \{a - \rho(x_i, y) - \eta\} \, dF_i(y) dx_i. \tag{18}$$

Using the survival function $S_i(y) := 1 - F_i(y)$, this further simplifies to

$$s_i = n_i \int_X S_i[y(x_i)] dx_i, \tag{19}$$

from which the city-size distribution is derived.

2.8 Empirical Connection

I will reformulate theoretical predictions so far of city size and trans-tolerance in preparation for empirical testing in sections 3 and 4.

2.8.1 City Size

I will differentiate distance-tolerance distributions by certain criteria to make testable predictions out of (18). There are various ways to rank density functions. I propose two of them below and discuss their implications for the city-size distribution.

Proposition 2.3 First-Order Stochastic Dominance and City Size

If $f_i(\cdot)$ *first-order stochastically dominates* $f_j(\cdot)$*,* $s_i \geq s_j$ *in equilibrium.*

Proof. Suppose that $f_i(\cdot)$ first-order stochastically dominates $f_j(\cdot)$. For any given location $x_i = x_j = x \in X$, $S_i[y(x)] \geq S_j[y(x)]$. Integrating both sides of the inequality over the country,

$$s_i = n_i \int_X S_i[y(x)] dx \geq n_j \int_X S_j[y(x)] dx = s_j \tag{20}$$

from (19). □

Remark. For the role of n_i in (20), see appendix A.1.

In order to prepare a testable prediction, I will further specify a component of city size (19) as follows:

$$n_i S_i\,[y(x_i)] = \alpha_i + \beta_i x_i. \tag{21}$$

The left-hand side measures the inflow from location x_i. Coefficient β_i measures the rate of decline in the inflow as a birthplace gets farther.

PROPOSITION 2.4 CHANGE IN INFLOW AND CITY SIZE

Suppose that the inflow is linear in distance as in (21)*. Then the equilibrium city size is*

$$s_i = \alpha_i + \frac{1}{4}\beta_i. \tag{22}$$

[5]

Proof. City size (19) implies $s_i = n_i \int_X S_i[y(x_i)]dx_i = \int_X (\alpha_i + \beta_i x_i)\,dx_i$, from which (22) is obtained. □

Remark. The slope β_i reflects two items that appear in (18). First, it captures trans-tolerance. If $y(x_i)$ makes a rapid ascent with x_i, then β_i will be small because the inflow will drop fast with distance. Second, it picks up the opportunity cost of relocation. If $\rho(x_i,\ y)$ does not flatten out with x_i much, then again β_i is small for the same reason.

The remaining entries in (18) do not depend on the distance and thus are picked up by α_i rather. For instance, distance-tolerance distribution $f_i(y)$ does not depend on x_i and thus it will be folded into α_i, which in turn measures type i's general distance tolerance or propensity to move out.

What applies to s_i goes for p_i as well because it is proportional to s_i via (13). Thus, if y_i has a high mean, then type-i consumers face a high rent in their city. An increased rent functions as a repellent but type i has low enough $y(x_i)$ to accept it. Conversely, if y_j has a low mean, p_j will be low, but that will not be enough to make up for their low distance tolerance and the city will be small.

2.8.2 Trans-Tolerance

Trans-tolerance is not observable but can be constructed from the recorded domestic migration patterns.[6] It is then useful to write trans-tolerance as a function

[5] The coefficients 1 and $\frac{1}{4}$ in (22) do not carry much empirical meaning as they are an artifact from having a circle of perimeter 1 for an economy.

[6] Note that (unobservable) $y(x_i)$ determines (observable) $m_i(x_i)$, not the other way around.

of the inflow by distance. Let $m_i(x_i)$ denote a measure of in-migrants to city i from location x_i.

PROPOSITION 2.5 INFERRED TRANS-TOLERANCE
Suppose $F_i(y)$ is strictly monotone increasing. In equilibrium, trans-tolerance can be extrapolated from in-migration as follows:

$$y(x_i) = S_i^{-1}\left(\frac{m_i(x_i)}{n_i}\right) \tag{23}$$

at any $x_i \in X$.

Proof. From (19), $m_i(x_i) := n_i S_i[y(x_i)]$ at any $x_i \in X$. Since $F_i(y)$ is strictly monotone increasing, its survival function $S_i(y)$ is strictly monotone decreasing in y. Thus, for any given x_i, $y(x_i) = S_i^{-1}\left(\frac{m_i(x_i)}{n_i}\right)$. □

Sections 3 and 4 investigate the empirical validity of propositions 2.4 and 2.5 respectively, namely, section 3 estimates the **size** of cities whereas section 4 estimates the **type** of cities.

3 Testing the Impact of Imperfect Mobility on City Size

3.1 Data Employed

I use the US Census Bureau's American Community Survey (ACS) 2009-2013.[7] The questionnaire asks which MSA a responder lived a year prior to the survey. A total of 381 MSA's report in- and out-migration so that there are 381×380=144,780 entries of inflow and outflow recorded between each pair of cities.

I will make several adjustments to the data in order to test theoretical implications of section 2.

In theory, the initial distribution is uniform whereas in reality, all locations are pre-populated with the number of consumers inherited from the previous period. To make the initial distribution as close to a uniform distribution as possible and eliminate the initial heterogeneity, I normalize the inflow by the total outflow from the location of origin.

[7] Data available at https://www.census.gov/data/tables/2013/demo/geographic-mobility/metro-to-metro-migration.html.

The model features a circle. The maximum moving distance possible is .5 regardless of the destination. In contrast, the actual US stretches over a limited expanse of land. The maximum distance differs city to city. Among 381 MSA's, Carson City, NV[8] has the shortest maximum distance possible of 4,187 km, from Bangor, ME. In turn, Honolulu and Bangor have the longest maximum distance possible of 8,293 km, between each other. While the gap between the top and bottom of the maximum range is mitigated by the fact that Alaska and Hawaii are included, this may nevertheless contaminate the estimation results: I may inadvertently underestimate Carson City's size for the reasons other than distance tolerance. Even if there were someone willing to move to the city from 8,293 km away, that worker will not show up in the data because the country cuts off at 4,187 km in. I may overshoot Honolulu and Bangor's size vice versa.

That said, I do not detect any systemic interaction between the maximum range and city size in figure 4. The cap on the distance does not affect the city size.

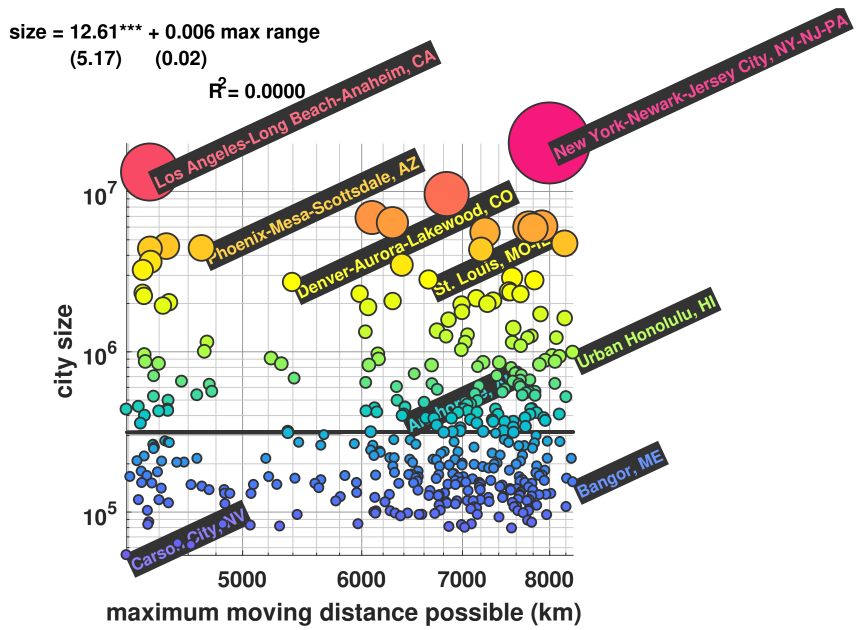


Figure 4. City size over maximum moving distance possible. Location of MSA, be it near the center or border of the country, has no statistically significant bearing on its size. *** denotes coefficient significant at 1%.

While the longest cutoff is about twice as long as the shortest cutoff, consumers perceive the distance on a logarithmic scale. The perceived gap is thus much smaller than twofold as a linear scale implies. I will nevertheless regress city size on inflow **and** the maximum range in section 3.2. The latter captures the said

[8]Coincidentally, Carson City was also the smallest MSA in 2013.

non-economic constraints so that the coefficient on the former will not be watered down by their presence.

I will validate proposition 2.4 in section 3.2 and then back it up with a different type of regression in section 3.3 for robustness.

3.2 Regression on Distance Elasticity of Inflow

Empirical testing of proposition 2.4 requires two rounds of estimations. First, I regress the city size on its inflow over $\log x_i$ and compute coefficient β_i of $\log x_i$ as I did in figure 1, not only for St. Louis but for all 381 MSA's. Then I further regress city sizes on α_i and β_i thus obtained. To this end, I write (22) as

$$\log s_i = \gamma_0 + \gamma_1 \alpha_i + \gamma_2 \beta_i \ \left(+\gamma_3 \alpha_i \beta_i + \gamma_4 \log(\text{maximum distance possible}_i) \right). \quad (24)$$

Note that γ_1 and γ_2 are coefficients of α_i and β_i, which themselves are coefficients. Since I take a log of x_i and s_i, β_i measures a percentage increase in size over a 1% increase in distance, i.e., the distance elasticity of inflow. It is more suited for empirical use than the rate of decline in inflow by distance originally cited in proposition 2.4.

I included the interactive term $\alpha_i \beta_i$ as a regressor. This product tends to be high among large cities than among small cities. The maximum moving distance possible is included as well.

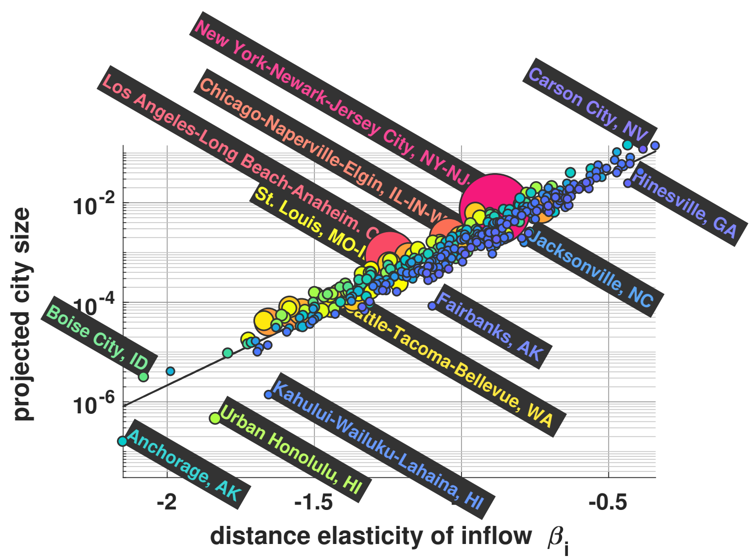


Figure 5. Projected city size over distance elasticity of inflow. Dot size and color are proportional to city size.

	intercept	α_i	β_i	$\alpha_i\beta_i$	$\log(\max x_i)$	R^2	adjusted R^2
coefficient	12.79***	.2183***				.3641	.3624
t-statistic	285.53	14.73					
coefficient	11.22***		-1.394***			.1736	.1714
t-statistic	66.37		-8.92				
coefficient	20.35***	.9360***	6.914***			.7000	.6985
t-statistic	55.15	25.76	20.58				
coefficient	20.69***	1.101***	7.106***	.1287***		.7224	.7202
t-statistic	57.37	23.92	21.83	5.51			
coefficient	12.61***				.005915	1.196e-06	-.002637
t-statistic	5.17				.02		
coefficient	7.054***	1.109***	6.525***	.1462***	1.487***	.7771	.7747
t-statistic	4.84	26.85	21.88	6.95	9.60		

Table 1. The impact of the distance elasticity of inflow on city size. Reported values represent γ's in (24). *** denotes coefficient significant at 1%.

Table 1 and figure 5 report the results.

Table 1 presents empirical evidence for proposition 2.4. Coefficient α_i has a positive impact on city size as expected. It simply means that a type with a high average propensity to move tends to create a large MSA. However, even after I controlled for this size difference in MSA's, the distance elasticity of inflow β_i still exerts a positive effect on the size. The city size indeed increases by as much as 6.5% when the distance elasticity of inflow grows by 1%. It would have been 0% if distance moved did not depend on city size. A city is large not only because it is a destination for a large fraction of out-migrants but also because the number of its in-migrants declines only gradually with distance. This attests to the strong presence of heterogeneity among consumers by type. Hence, perfect mobility is statistically unlikely.

The product of the two regressors above, $\alpha_i\beta_i$, also influences the city size. Even when a city can attract many residents nearby, if there is no sustained inflow from around the country, the city will not be large. Conversely, even when a city has a constant inflow of residents over X, if its potential movers have a low propensity to move out, then the city will not be large either. This applies in particular to four MSA's in Alaska and Hawaii as I will discuss below.

As figure 4 and table 1 show, the maximum moving distance possible per se has virtually no impact on the city size. However, if two MSA's have the same distance elasticity of inflow, the one with the longer maximum distance possible will have a larger size, for the reason I explored in section 3.1. Thus, its inclusion is deemed necessary to obtain an accurate reading of γ's.

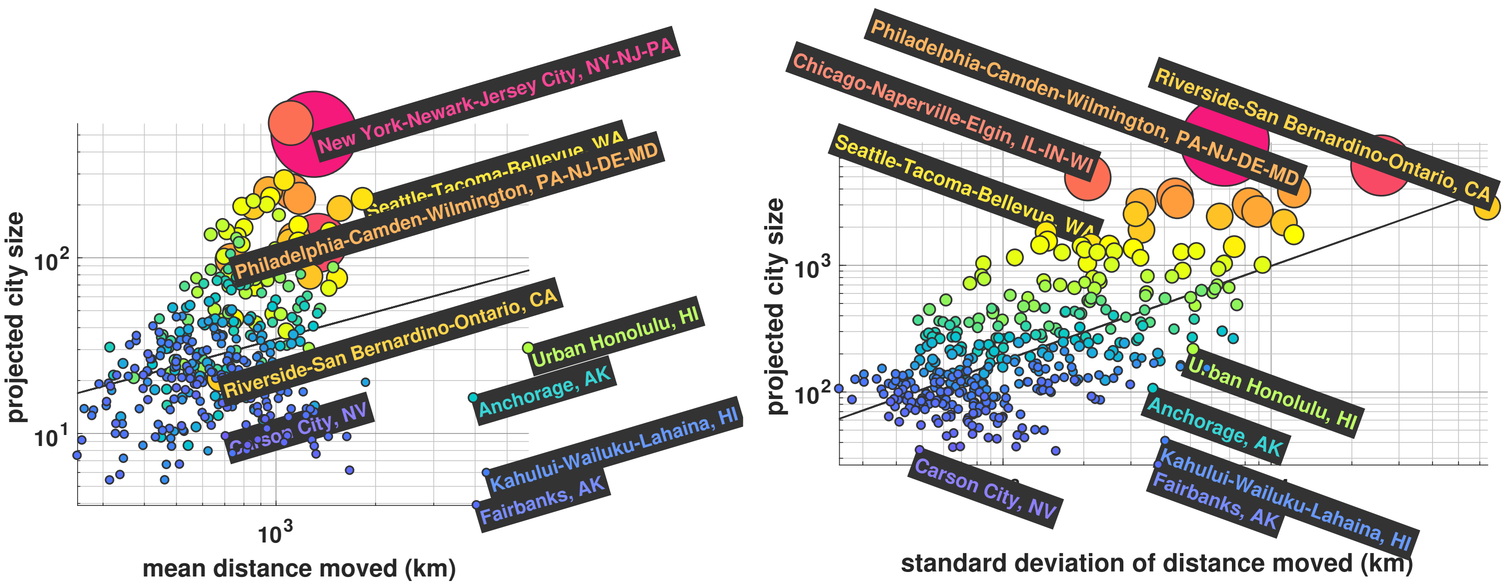
In figure 5, all four entries from Alaska and Hawaii cut below the expected size. It is a systemic pattern that emerges from their geographic disposition rather than for economic reasons. Given their size, these cities should have lower β_i and they would have if they were surrounded by other cities nearby. In reality, they are surrounded by Canada or the Pacific, neither one of them provides an inflow. Inevitably, β_i cannot drop despite their size to account for distant inflows from the lower 48. Namely, their distance elasticity of inflow is high because of distance rather than inflow. Other MSA's in remote locations suffer the same symptom as well but to a lesser degree.

None of these would matter if one assumed perfect mobility. New York City may be composed exclusively of people from New England or of California natives, with no difference in its size in the end. In this case $\beta_i = 0$ across the country and the resultant city size s_i simply matches the exogenous type size n_i. In contrast, the present model anticipates that New York City cannot have the size it has unless it gathers workers 1) of distance tolerant type (high α_i), and 2) from across the country (high β_i). ACS indicates that both regressors are markedly different from 0 and exert a crucial influence on the development of MSA.

3.3 Regression on Moments

To assess the robustness of the findings in section 3.2, I will regress city size on the moments of its inflow. A high α_i translates to $m_i(x_i)$ having a high mean and a high β_i translates to a high variance. If the findings in section 3.2 are valid, these moments should have a positive effect on the size of the destination city as do α_i and β_i.

Table 2 and figure 6 report the result, which is in lockstep with section 3.2. When the mean or the standard deviation of inflow inflates by 1%, the destination's size grows by .5% and .7% respectively. The findings indicates that the pdf of distance tolerance flattens out and shifts to the right with the city size in line with proposition 2.3. For any given x_i, large cities have a higher mean of inflow than small cities, which in turn implies that the former first-order stochastically dominates the latter in y and thus has a high α_i. In addition, as β_i becomes higher, the variance of inflow becomes larger as does the city size, which also supports proposition 2.4. Note that if one assumes perfect mobility instead, the moments of distance moved would be the same for any i and thus orthogonal to the city size, which is unlikely according to figure 6.



(a) (b)

Figure 6. City size over the mean and standard deviation of distance moved among in-migrants. Colors and dots are size proportionate.

	intercept	mean	standard deviation	R^2	adjusted R^2
coefficient	7.365***	.8059***		.1079	.1056
t-statistic	9.39	6.77			
coefficient	6.975***		.8024***	.4005	.3989
t-statistic	19.38		15.91		
coefficient	3.988***	.5123***	.7487***	.4423	.4394
t-statistic	6.02	5.32	15.05		

Table 2. City size regressed on the mean and standard deviation of inflow. *** denotes coefficient significant at 1%.

Four MSA's in Alaska and Hawaii (two each in each state) have a high mean and variance for their size. Their $m_i(\cdot)$ does not take off until later because they only have one city nearby (the one and only other MSA in the same state) and the next hike in value needs to wait till they cross the Pacific or Canada. As in section 3.2, this is largely a geographic artifact and it does not necessarily mean that they gather high-skilled and/or distance-proof labor. Aside from them, among large cities, Philadelphia and Riverside have roughly the same mean distance moved. However, Philadelphia consists mostly of locally sourced labor, whereas

Riverside takes in workers of various origins.

While both of them are significant, table 2 indicates that the standard deviation exerts more influence on city size than the mean does. The size responds **more to how diverse the locations of origin are** than to how far people moved on average.

4 Estimation of City Types

Along with the impact of imperfect mobility in section 3, the present model further makes predictions about the type of city. This section examines what each MSA is composed of by way of trans-tolerance computed in proposition 2.5.

In the absence of direct observation, I assume that y follows the normal distribution with mean μ_i and unit variance for expository purposes. (In practice, any distribution that satisfies the assumption in proposition 2.5 will do). I set μ_i equal to the log of the total inflow the destination receives, and shift it upwards by μ_{171} across the board so that the city of geometric mean size (Tuscaloosa, AL, 171st in rank) will have a mean of zero. I ran kernel density estimation on inflow first to filter out the noise.

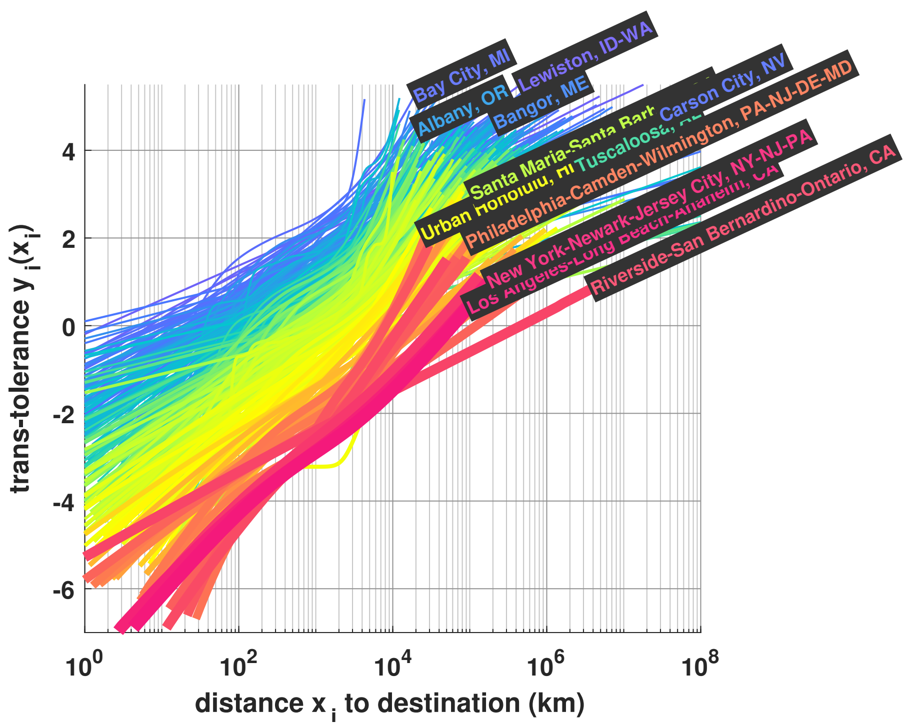


Figure 7. Estimated trans-tolerance $y_i(x_i)$ by MSA. At each x_i, consumers with y above $y_i(x_i)$ are expected to move to MSA, and the ones below the line are likely to stay at their birthplace x_i. Line color and width are proportional to the total inflow into each MSA.

rank	remark	MSA	total inflow	$y_i\left(10^0\right)$	$y_i\left(10^1\right)$	$y_i\left(10^2\right)$	$y_i\left(10^3\right)$	$y_i\left(10^4\right)$	$y_i\left(10^5\right)$
1		Los Angeles-Long Beach-Anaheim, CA	244,099	-9.74	-6.14	-4.28	-2.89	-1.46	0.46
2		New York-Newark-Jersey City, NY-NJ-PA	228,599	-7.85	-5.89	-4.19	-2.85	-1.36	0.51
3		Washington-Arlington-Alexandria, DC-VA-MD-WV	196,434	-9.73	-5.91	-4.04	-2.58	-1.04	0.89
4		Riverside-San Bernardino-Ontario, CA	178,510	-5.28	-4.30	-3.33	-2.39	-1.49	-0.60
5		Dallas-Fort Worth-Arlington, TX	172,896	-10.87	-7.06	-4.64	-2.66	-0.29	2.33
12		Philadelphia-Camden-Wilmington, PA-NJ-DE-MD	126,264	-9.49	-5.51	-3.24	-1.75	-0.22	2.05
34		St. Louis, MO-IL	52,944	-9.17	-5.56	-3.26	-1.28	0.70	2.85
44	max range	Urban Honolulu, HI	41,804	$-\infty$	$-\infty$	-4.32	-3.22	-0.13	5.95
95	1st quarter	Santa Maria-Santa Barbara, CA	22,928	-6.87	-3.47	-1.65	-0.12	1.24	2.97
171	geometric mean	Tuscaloosa, AL	11,911	-2.69	-1.64	-0.63	0.36	1.32	2.27
191	2nd quarter	Sierra Vista-Douglas, AZ	10,576	-2.97	-1.92	-0.90	0.11	1.11	2.14
286	3rd quarter	Albany, OR	5,658	$-\infty$	-2.88	0.61	1.59	3.40	7.76
314	max range	Bangor, ME	4,540	-5.01	-1.96	-0.26	1.22	2.56	4.10
366	min range and size	Carson City, NV	3,062	-0.66	0.08	0.81	1.54	2.27	2.99
371		Bay City, MI	2,770	-5.78	-1.19	1.54	2.45	4.38	8.38
381	last quarter	Lewiston, ID-WA	1,732	-1.79	-0.44	0.88	2.15	3.39	4.64

Table 3. Estimated trans-tolerance values of select MSA's at $x_i = 10^0, \cdots, 10^5$. For instance, a St. Louisan born 100 km away from St. Louis must have had the tolerance value greater than $y_{\text{St. Louis}}(100) = -3.26$. Those who drew a value below -3.26 stay 100 km away from the city.

Figure 7 plots $y_i(x)$ for all the destinations and table 3 lists $y_i(x)$ for select destinations at $x = 10^0, 10^1, \cdots, 10^5$ km.[9] All in all, large cities register lower trans-tolerance values than small cities. Moreover, large MSA's trans-tolerance tends to take off more slowly than small MSA's.

For instance, a resident $x_i = 10^3$ km away from her type destination has to have only $y = -2.89$ or above to move to Los Angeles, whereas that of Lewiston, ID and WA needs to exceed 2.15. In other words, a mover with a tolerance level of -1.46 could have been born as far as 10^4 km away from Los Angeles but the same person has to have been born less than 10 km away from Lewiston. If $y_i(x)$ is assumed (and likely) to be negatively correlated with the skill level for instance, then those who move to Los Angeles are more skilled than those who move to Lewiston, ID and WA.

A slow rise in $y_i(x_i)$ among large MSA's are most notable in Riverside, CA. The city keeps picking up in-migrants well past 10^4 km, at which point the trans-tolerance of other MSA's such as Philadelphia starts to climb faster. Riverside

[9]Some destinations record $-\infty$ in their immediate vicinity. These represent trans-tolerance values for residents who were born only 10^0 or 10^1 km away from their type destination, which are expectedly low. While x starts from 0 in theory, most MSA's are more than 100 km apart from each other. I do not have enough data density from which to extrapolate those residents' accurate trans-tolerance value. They are likely to be recorded as movers **within** the same MSA in practice.

seems to attract workers from a wider range of locations for its size, in keeping with the previous analysis in section 3.3. The opposite applies to small MSA's, whose in-migrants are locally sourced. Trans-tolerance of Bay City, MI makes a steep ascent early on at around $x_{\text{Bay City}} = 10^2$. The majority of its incoming residents are from nearby Saginaw, followed by Midland, Flint, Lansing and Detroit, all located in Michigan. Past these source locations, it becomes extremely unlikely that a worker changes her residence to Bay City.

5 Interpretation of Results

Sections 3 and 4 have established the pivotal role that geographic mobility plays in the determination of the city-size distribution. The city-size distribution simply becomes uniform if one assumes perfect mobility in the current model.

Exactly what is it about imperfect mobility that gives the city-size distribution a heavy tail? To answer this question, let us revisit the assumptions about $f(x_i)$, $y(x_i)$ and $\rho(\cdot)$ made in section 2.7. I will let $f(x_i)$ pick up type differences as before but the same argument goes for the other two as well.

City i is larger than city j because $f_i(\cdot)$ first-order stochastically dominates $f_j(\cdot)$ from proposition 2.3. According to figure 3(a), the city-size distribution is heavy tailed because the majority of types have its distance-tolerance CDF taking off long before it reaches the trans-tolerance value. Thus, few people move out and the resulting city size is small. However, there exist certain types whose CDF starts to rise near or past the trans-tolerance value. While these types are very rare, if a city happens to be of this type, then almost all of its potential workers do move to the city, resulting in an explosively large population. Thus, large cities are few and far between because 1) most people are distance **in**tolerant but 2) those who **are** tolerant will move to a select few cities **with a high probability** regardless of how far their birthplace is.

One of the root causes behind a large skewness of the empirical city-size distribution is found in assumption 2.1. If people perceive distance linearly, the trans-tolerance value climbs at a constant rate and cities will come uniform in size. On the contrary, the robust empirical evidence so far indicates that a linear scale cannot explain the observed migration patterns. People sense distance logarithmically. This leaves countering effects on small and large cities: For a small city, the source area of in-migrants is restricted to a very short radius. The difference between 100 km and 110 km **appears as large as** the one between 10,000 km and 11,000 km to them. For a large city, if it is attractive enough or its incoming residents are tolerant enough to push the source radius beyond a certain point, distance plays barely any role in the location choice of workers born far from the city. To them, in reverse, the difference between 10,000 km and 11,000 km **appears as small as** the one between 100 km and 110 km. While the gap between the absolute distance is 10 km over 1,000 km, the gap between the relative distance is the same factor of 1.1 for both. It is the latter concept of distance that people call on when making location choices. Consequently, the source area becomes exponentially large, as does the city size itself. Limpert et al. [LSA01] suggest that something of **multiplicative** (rather than **additive**) nature underlies many economic phenomena. Workers' tendency to gauge distance in **relative** (rather than **absolute**) terms may be one of those fundamental causes that give rise to the heavy-tailed city-size distribution.

6 Conclusion and Extensions

I examined the role that distance moved plays in determining the city-size distribution. Each worker draws a distance tolerance level from the distribution unique to her type. She then makes a decision on whether to stay put or move to a city to tap into urban productivity that the city has to offer. She compares urban productivity with an urban housing market, a fallback value of her utility level when she stays and, exclusive to the current model, how far the city is from her birthplace when making location choices. The city-size distribution arises as a result of factors specific to each industry and city, which determine how many people move to the city and how far they move.

I regressed the city size on several aspects of the underlying distribution of distance tolerance. The empirical data are in accordance with the predictions from the model. The majority of types are distance intolerant. They gather only in small numbers as most of them prefer to stay at their birthplace unless they happened to be born close to the city. Then there are a very few but noticeable types who are willing to move. They gather in large numbers to create gravitationally large cities as most of them will move to the city regardless of where they were born. The data also reveal that consumers recognize distance on a logarithmic scale, which cements the heavy tail of the city-size distribution.

I assumed that each city hosts at most one type. Actual cities host multiple types. Assuming co-location of different types in the same way as Eeckhout et al.

[EPS14] would yield a finer result than above, provided that relocation data are recorded by industry. I do not know of such data.

In order to stay focused on the city size, I left urban productivity as plain as possible. In reality, the distance moved may be correlated with in-migrants' productivity, the aggregate of which defines the citywide productivity. It will be useful to relax the current assumption on urban productivity and have the distance moved explain it.

I assumed that each type knows where his type-matched city is. However, it is not easy to know in advance where that city is located. Skill compatibility is not fully understood until workers actually start working at their destination, which may or may not be their right destination. One may introduce some uncertainty in matching between type and industry.

A Appendix

A.1 Role of Type Distribution

Observe that the inequality (20) will be flipped if n_j is sufficiently larger than n_i. That is, there is a trade-off between the distance-tolerance distribution and the number of potential city residents. Even when type i is tolerant towards relocation overall, its corresponding city size may be trumped by more intolerant type j if type i is outnumbered by type j in the hinterland to begin with.

While the current model considers only two stages of decision making, the time horizon can be extended to allow for intertemporal dependence of type. A child of distance tolerant parents is likely to be born in his type-matched city because there is a good chance that his parents have already moved to the city. If the child's type is same to his parent's, a big city tends to seal its top-tier status through this positive feedback loop (cf. Duranton [Dur07]). In this case, constant n_i will be replaced by $\nu_i(x_i)$ with $\int_X \nu_i(x_i)dx_i = n_i$ so that (18) becomes a law of motion $s_{i,t+1} = \int_X \nu_i(x_{i,t})S_i[y(x_{i,t})]dx_{i,t}$, where the city size $s_{i,t+1}$ in the next period is determined by the current distribution $\nu_i(x_{i,t})$ and size $s_{i,t}$.

In this case, $\nu_i(x_i)$ works in the same way as $f_i[y(x_i)]$. While I made $f_i(\cdot)$ exogenous, it can be interpreted as what $\nu_i(x_i)$ converges to in a steady state.

Or even within the life span of one worker, he may relocate a number of times over the course of life. Because the cost of relocation $\rho(\cdot)$ is concave in distance moved, relocation becomes less costly as the second relocation is less draining than the first. (To model this, one needs to forgo the assumption that a type forms a one-to-one correspondence with an industry).

Both extensions will not counteract but only reinforce the result derived from the current model. Thus, I will take the current model's predictions as the most conservative results and leave the extensions above for future research, which should find a wider and more intense impact of imperfect mobility on the city-size distribution.

References

[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.

[BDRN14] Kristian Behrens, Gilles Duranton, and Frédéric Robert-Nicoud. Productive cities: Sorting, selection, and agglomeration. *Journal of Political Economy*, 122(3):507–553, 2014.

[Bow70] Samuel Bowles. Migration as investment: Empirical tests of the human investment approach to geographical mobility. *The Review of Economics and Statistics*, pages 356–362, 1970.

[BW18] Marcus Berliant and Axel H Watanabe. A scale-free transportation network explains the city-size distribution. *Quantitative Economics*, 9(3):1419–1451, 2018.

[DD12] Donald R. Davis and Jonathan I. Dingel. A spatial knowledge economy. *Working paper*, 2012.

[DP19] Gilles Duranton and Diego Puga. Urban growth and its aggregate implications. Technical report, National Bureau of Economic Research, 2019.

[Dur07] Gilles Duranton. Urban evolutions: The fast, the slow, and the still. *The American Economic Review*, 97(1):197–221, 2007.

[Eec04] Jan Eeckhout. Gibrat's law for (all) cities. *The American Economic Review*, 94(5):1429–1451, 2004.

[EPS14] Jan Eeckhout, Roberto Pinheiro, and Kurt Schmidheiny. Spatial sorting. *Journal of Political Economy*, 122(3):554–620, 2014.

[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.

[GI04] Xavier Gabaix and Yannis M Ioannides. The evolution of city size distributions. *Handbook of Regional and Urban Economics*, 4:2341–2378, 2004.

[GKS01] Edward L Glaeser, Jed Kolko, and Albert Saiz. Consumer city. *Journal of economic geography*, 1(1):27–50, 2001.

[HJO16] Boris Hirsch, Elke J Jahn, and Michael Oberfichtner. The urban wage premium in imperfect labour markets. *working paper*, 2016.

[HMP+92] Günter Haag, M Munz, D Pumain, L Sanders, and Th Saint-Julien. Interurban migration and the dynamics of a system of cities: 1. the stochastic framework with an application to the french urban system. *Environment and Planning A*, 24(2):181–198, 1992.

[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.

[Lad67] Jack Ladinsky. Occupational determinants of geographic mobility among professional workers. *American Sociological Review*, pages 253–264, 1967.

[LSA01] Eckhard Limpert, Werner A Stahel, and Markus Abbt. Log-normal distributions across the sciences: Keys and clues. *BioScience*, 51(5):341–352, 2001.

[Man10] Alan Manning. The plant size-place effect: agglomeration and monopsony in labour markets. *Journal of economic geography*, 10(5):717–744, 2010.

[MSW11] Raven Molloy, Christopher L Smith, and Abigail Wozniak. Internal migration in the united states. *Journal of Economic perspectives*, 25(3):173–96, 2011.

[Rau13] Ferdinand Rauch. Cities as spatial clusters. *Journal of Economic Geography*, 14(4):759–773, 2013.

[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.

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