International Journal of Religion
2024
Volume: 5| Number 12 | pp. 979 – 992
ISSN: 2633-352X (Print) | ISSN: 2633-3538 (Online)
ijor.co.uk
DOI: https://doi.org/10.61707/ygpv9305

An Analysis of Migration Determinants with An Emphasis on Housing Cost and Income

Chao Li[1], Cang Han[2], Qixuan Zhu[3]

Abstract

This paper investigates the determinants of migration and location choices, focusing on their magnitude and directions. Through applying empirical methods and theoretical model, this paper reveals the connection between the migration decision and individual's economic conditions, and the regional variations of the effects of factors. It is found that individuals in the lowest (first quartile) and highest (fourth quartile) income group are more likely to move, than those in median group. Migrants are attracted to large metropolitans that can provide higher incomes, lower gender discrimination, higher education attainment benefits, and better-quality amenities, but higher housing costs deter this choice. The sensitivity to housing costs varies among cities, with migrants being more responsive to cost increases in large metropolitan areas than in smaller urban areas. A Monte Carlo simulation shows that a 10% increase in New York's housing cost could potentially lead to a population decline by around 0.15%, while in Denver, this negative effect is insignificant.

Keywords: *Migration, Location Choice, Housing Cost, Heterogenous Individuals, Amenity*

INTRODUCTION

Migration driven by a complex interplay of genetics, socio-economic factors, and cultural ties, remains a fundamental aspect of human existence. To better understand the impacts and mechanisms of migration patterns, research on migration involves a multitude of methods and technologies including theoretical and empirical methods. A seminal work in the field of theoretical migration study, Schelling (1971) built one- and two-dimensional segregation models to investigate the impact of ethnic group on an individual's community choice. Roback (1982) constructed an equilibrium model to explain the difference in distribution of workers among locations with various non-tradable amenities. Van Nieuwerburgh and Weill (2010) studied the effect of migration on the mechanism of the increasing house price dispersion using a dynamic equilibrium model. Davis et al. (2021) investigated the U.S. city growth rate through constructing a dynamic general equilibrium model of urban migration. Cun and Pesaran (2022) focused on the size and speed of spatial spill-over effects among mainland U.S. states. Through calibrating an equilibrium spatial macro-finance model, Mabille (2023) explained the dampened first-time home buyers considering migration as a channel. The land-use regulations on labor allocation have been studied by Hsieh and Moretti (2019) and Herkenhoff et al. (2018). Other determinants of population reallocation such as regional productivities and amenities (Komissarova, 2022), job opportunities (Díaz et al., 2023), housing supply elasticities, and credit constraints (Ammar et al., 2001; Giannone et al., 2020) are examined using lifecycle mobility model or spatial equilibrium model.

Complementary to the theoretical method, empirical studies allow researchers to focus on specific factors of migration. Glaeser et al. (2001) proposed a method to measure the amenities of a city and argued there are four particularly critical amenities: variety of goods and services, weather, crime rate, and speed. The relationship between climate and location decision has been examined by Poston Jr et al. (2009) and Eichman et al. (2010). Households have been found to prefer warm coastal areas, while firms tend to choose large, growing cities (Chen and Rosenthal, 2008). Rich cultural amenities make a location more attractive to migrants (Clark, 1992). Crime rate is another foremost consideration when people make picks (Gottlieb and Joseph, 2006). Economic

[1] School of Insurance, Shandong University of Finance and Economics, Shandong Province, China. ORCID: 0009-0005-5455-7458 E-mail: chao92.li@hotmail.com

[2] Finance Department, Business School, Soochow University, SuZhou, Jiang Su Province, China. ORCID: 0009-0004-6633-2312. E-mail: hcang@stu.suda.edu.cn (Corresponding Author)

[3] Finance Department, Business School, Soochow University, SuZhou, Jiang Su Province, China. Email: 20244010002@stu.suda.edu.cn

opportunities (unemployment rate, per capita income) and the moving costs are examined by Davies et al. (2001). The role of earnings (Agius Vallejo and Keister, 2020; Kennan et al., 2011), quality of life (Berger and Blomquist, 1992), and housing price (G. Berger, 1992; Bishop, 2008) in mobility and destination choice has been quite clearly elucidated. Plantinga et al. (2013) extend their work and investigate the effects of housing cost on migration through using multiple housing cost measurements. However, the negative effects of rising housing prices are more pronounced for renters and insignificant for owners (Meng et al., 2023). Andrienko et al. (2020) delve into the potential of big data to deepen insights into migration through a tri-phased analysis.

Among these migration determinants, economic factors such as housing cost and income consistently prove to be the primary drivers in relocation decisions. Over 2022[1], the Bay Area had the largest net outflow of all large metro areas due to the skyrocketing housing cost[2]. The outburst of Covid-19 stimulates the popularization of remote and hybrid work options which accelerates this trend, leading to a demographic change in Silicon Valley.

Motivated by this phenomenon, this study aims to address two key questions: (i) What roles do individual economics status and regional amenities play in migration decisions and location choices? (ii) What are the directions and magnitudes of these effects, and do these effects exhibit regional variations? To address the two questions, we have conducted both empirical and theoretical analysis. A critical contribution of this paper is that this paper emphasizes the investigation of the magnitudes of these influences and their variability across different areas. The empirical method is inspired by Plantinga et al. (2013), where individual's wage and housing cost were estimated for each location conditionally on regional unobserved factors and individual characteristics. The expected wage and housing cost in each unselected area were predicted based on the individual attributes, using the estimated wage and housing cost equations. Determinants of migration decisions and location choices are traceable to area-specific and individual-specific housing cost, wages, and area-specific amenities. To examine the magnitude of these determinants, we built a static stochastic theoretical model based on the utility-based-theory calibrated to empirical findings and facts. Modeled individuals with heterogenous characteristics and preferences on housing service make location choice to maximize their utilities. The effects of housing cost and income were examined with the help of Monte Carlo simulation.

Using the 2022 5-year American Community Survey (ACS) Public Use Microdata Sample (PUMS) household and personal data, we examined the factors of migration and location choices through logistic and conditional logit model. The migration decision has been found to be closely related to individual's characteristics, income, and housing expenditures. Relative to the median income group, both low-income (first income quartile) and high-income (fourth income quartile) individuals are more likely to migrate, influenced by observable (moving costs, financial constraints) and unobservable (job relocations, lifestyle adjustments) considerations. Individuals living in expensive places are more likely to enjoy the local amenity services and tend to stay. In consideration of amenities, both market density and entertainment facilities impose positive effects on location choice. Higher income, reduced gender discrimination, and greater education attainment benefit enhance the likelihood of selecting large metropolitan areas, while higher housing costs dampen migration enthusiasm. Through 100 path of Monte Carlo simulation, we found that the migrants are more sensitive to housing cost changes in large metropolitans compared to small urban areas, consistent with economic principles where higher prices result in more elastic demand. This phenomenon stems from the combination of higher absolute housing costs and the more diverse and denser population in large metro-areas.

The remainder of the paper consists of six sections. Section 1 is an introduction of the background of the current study. Section 2 illustrates the analysis framework involved, where our model is presented. Section 3 shows the data and source thereof we used. The regression and simulation results are presented in Section 4, where an explanation is provided for the results. Section 5 demonstrates the baseline results from the counterfactual experiments conducted. Section 6 is the conclusions from the current study.

Framework

Using 2017-2022 5-year Public Use Microdata Survey (PUMS) data, we analyzed the migration decision and location choice conditional on individual characteristic and regional amenities. Then building and calibrating a theoretical model, we conduct counterfactual experiments to investigate the magnitudes of these effects.

Empirical Method

Following the assumption of Plantinga et al. (2013) and Davies et al. (2001), we suppose that the determinants of location choice are a function of housing cost, income, and regional amenities. The location choice is examined using a conditional logit model. To avoid Independence of Irrelevant Alternatives (IIA) property violation, we categorized the sample Metropolitan Statistical Areas (MSAs) into several groups, where a nested logit model was applied. Both conditional logit model and nested logit model originate from utility-based choice theory. Individual i chooses city j among a set of alternatives, if and only if the utility U_{ij} is greater than the utility of other alternatives. It is expressed by the following form:

$$U_{ij}\left(X_{ij}, A_j, C_i\right) \geq \max\left\{U_{i1}, U_{i2}, \ldots, U_{ik\neq j}\right\}, where\ k \neq j \quad (1)$$

Where X_{ij} stands for the variables which varies across individuals and regions such as income, housing cost, and consumption. A_j represents the amenities of region j, and C_i refers to the individual characteristics of i. The utility function consists of two components: the observable portion V_{ij} and the unobservable portion ε_{ij} which is unknown. The utility function can be expressed by:

$$U_{ij} = V_{ij}\left(X_{ij}, A_j, C_i\right) + \varepsilon_{ij} \quad (2)$$

The observed portion V_{ij} is assumed to be a function of housing cost and wage income. Housing cost and wage are different across each location j for each individual i, implying that the individual does not spend the same amount of income on the housing service for two reasons: first, each MSA has different amenities, housing prices, and rents; second, each individual has heterogenous preference for housing attributes. Consequently, in this case, the housing cost depends on individual's "housing taste" which is a function of individual characteristics and regional attributes. This suggests that the same housing service purchased in one location may not be affordable in another location. This preference for and affordability of housing service can affect households' location choices. Similar to the housing cost, the income is determined by individual characteristics, and each location offers different wages to the same individual.

The solution to the location choice problem requires estimating the utility of all alternatives U_{ik}. This study estimated the housing cost HC_{ij} and wage income received by individual i in each location.

Housing cost and Income Estimation

The housing cost HC is measured by the monthly cost of a house, where HC refers to the monthly mortgage payment and other utility fees if this household is reported as an owner, whereas HC represents the gross monthly rental payment and other fees such as property insurance if the agent is a renter. The regression equation of housing cost is written as follows.

$$\ln HC_{ij} = \alpha_{hj} + \beta_{hj} C_i + \varepsilon_{hj} \quad (3)$$

Where HC_{ij} stands for the housing cost and β_{hj} is a set of coefficients on individual's attributes C_i. HC_{ij} is the estimated housing cost of each city, which is a function of individual's characteristics, implying that the demand of housing service quality is determined by the household's attributes. The constant term α_{hj} captures the unobserved area-level shocks spread through housing market.

Similar to the housing cost equation, the wage income is regressed for each city using observations of wage payments and corresponding individual characteristics, such as age, race, educational level and gender. The estimation equation is expressed as:

$$\ln Wage_{ij} = \alpha_{wj} + \beta_{wj} C_i + \varepsilon_{wj} \quad (4)$$

Where $\ln Wage_{ij}$ stands for the natural logarithm of annual salary payment in metro-area and is individual specific. β_{wj} refers to the set of coefficients of individual's characteristics and varies across city j. Individual attributes, including gender, age, education level, race, and marital status, are represented by the vector of C_i. The observations of individuals in a city are only concerned with the working-age population ranging from 20 to 65. The constant term is an important part of the identification strategy employed in the migration analysis. This study estimated a single wage equation for each city, whereby the constant terms act as fixed effects in our model across all individual observations in each metro area. The equation was used to predict the expected wage of individual i with characteristic C_i in unselected cities.

Migration decision and Location Choice

We first examined the households' migration decisions over the period from 2017 to 2022. The decision to stay or move out is explained by individuals' attributes, income, and housing cost. The equation used to estimate is as follows:

$$Y_i = \alpha_i + \beta_c C_i + \beta_{in} Wage_i + \beta_{hc} HC_i + \varepsilon_c \tag{5}$$

Where Y stands for individual's migration decision regarding the choice of individual i to stay $(Y = 0)$ in current city or to move out $(Y = 1)$. Following Peng et al. (2002), the possibility that an individual chooses to move out is given by:

$$Prob(Y = 1|X = i) = \frac{e^{\beta_c C_i + \beta_{in}(Wage_i) + \beta_{hc}(HC_i)}}{1 + e^{\beta_c C_i + \beta_{in}(Wage_i) + \beta_{hc}(HC_i)}} \tag{6}$$

Where X stands for the individual characteristics, and the housing cost HC and income $Wage$ are divided into four groups based on the quartile levels, where an income level lower than 25th percentile is defined as poverty, while an income between 25th percentile and 50th (the median) is regarded as low income; and a higher level is categorized as high-income group. By analogy, income and housing-related expenses are grouped into four categories: cheap, economy, comfort, and fancy based on the nominal cost.

Following Davies et al. (2001) and Hosmer Jr et al. (2013), we applied a conditional logit model to estimate the location choice of individuals conditional on moving. To avoid IIA violation, we also employed the nested logit model. The MSAs are categorized into four groups based on the population. The location choice is assumed to be affected by the individuals' characteristics, regional amenities, income, and housing cost. The location choice equation is specified as follows:

$$U_{ij} = \beta_A A_j + \beta_C C_i + \beta_{in} Wage_{ij} + \beta_{hc} HC_{ij} + \varepsilon_v \tag{7}$$

Where U_{ij} stands for metro-area choice made by individual i to move to city j out of the location set J. A_j denotes the regional heterogenous amenities comprising air quality, commuting time, market density, entertainment facilities, and crime rate. It differs across all areas and equally affects all residents. Income and housing cost are alternative specific variables which vary across individuals and regions. The probability $\Pr(U = j|X = x)$ of migrant i moving into location j can be expressed as:

$$\Pr(U = j|X = x) = \frac{e^{\beta_A A_j + \beta_C C_i + \beta_{in} Wage_{ij} + \beta_{hc} HC_{ij}}}{\sum_{k=1}^{J} e^{\beta_A A_k + \beta_C C_i + \beta_{in} Wage_{ik} + \beta_{hc} HC_{ik}}} \tag{8}$$

The coefficients are estimated by maximum likelihood function. In this study, we estimated the location choice of working-age population aged from 20 to 65, with each individual having 49 choices of the sample MSAs.

The equation estimation above showed the determinants influencing migration choice and location choice. To further investigate the magnitudes of these effects and explore the potential regional variations, we performed counterfactual experiments by constructing a theoretical model as presented in the following subsection.

Theoretical Model

Consistent with empirical theory, rational households make decisions to maximize their utility function. The utility obtained by agent i from region j is defined as U_{ij}, where individual chooses city j if and only if the utility satisfies the following condition: $U_{ij} \geq U_{i,k \neq j}$, for all $k, j \in J$. Where J is the set of all alternative choices. Each heterogenous individual is endowed with different preferences for consumption and characteristics such as age, gender, race, education etc. Households derive utility from non-housing consumption (consumption) and housing service in city j, and its function follows a Cobb-Douglas form and is specified as:

$$U_{ij} = c_{ij}^{\chi_i} h_j^{1-\chi_i}, \qquad 0 < \chi_i < 1 \tag{9}$$

Where c_{ij} stands for the consumption an individual i can enjoy in city j, and h_j refers to the housing service quality provided by city j. A higher h_j means a better quality of amenities enjoyed by residents. The housing service h_j is a function of city amenities A_j such as air quality, supermarket density, crime rate and entertainment facilities.

$$h_j = \boldsymbol{\beta} A_j \tag{10}$$

Where $\boldsymbol{\beta}$ is the set of coefficients on non-tradable city amenities A_j. The preference for consumption $\chi_i \in (0,1)$ differs across all sample individuals. It is randomly distributed to each individual and follows an I.I.D normal distribution $\chi_i \sim N(0.5, \sigma^2)$. This study excluded the extreme situation ($\chi = 0$), where people have to spend all their income on housing service.

The utility function is maximized subject to the budget constraint:

$$c_{ij} + HC_j(h_j) = Wage_{ij} \tag{11}$$

The $Wage_{ij}$ in region j was estimated and predicted by the estimated wage equation above, and housing cost HC_j was calibrated using the housing price data. To capture the fact that the relatively low-income group have the ability to live in a city with high housing prices, the model was predicated on each city providing a range of different house bundles, varying from high-quality to low-quality with the corresponding housing cost.

The simulation process goes as follows: first, we generated a certain number of individuals randomly endowed with different characteristics, such as the consumption preference χ, age, gender, education level, and race. Subsequently, conditional on those characteristics, the expected wage incomes in each location were predicted. Given the wage income and housing cost, sample individuals evaluate the utility of all options to choose the one providing the highest utility.

Data

The data used to conduct analysis are from the American Community Survey (ACS) of the U.S. Census Bureau. The 5-year Public Use Microdata Survey (PUMS) 2017-2022 document not only the individual level variables such as age, class of work, migration status, marital status, and point-to-point migration status, but also household level variables such as housing tenure status, utility payment, monthly mortgage, and rents. This point-to-point migration data set makes it possible for us to collect and group the data needed for the selected MSAs. As this study involved the estimation of wage income, we used the subset of PUMS sample based on the following criteria: working-age respondents aged older than 20 and younger than 65 and the migration data among 49 selected MSAs.

The wage income is the main source of income of working-age populations. (Figure 1) showed the density function of nominal wage and natural logarithm of the wage of the 49 sample cities. As shown in (Figure 1), over 90 percent of the total population's annual wage is less than 100,000 dollars and over 99 percent of the total population's annual wage is less than 350,000 dollars. The average annual wage is approximately 50,000 dollars.

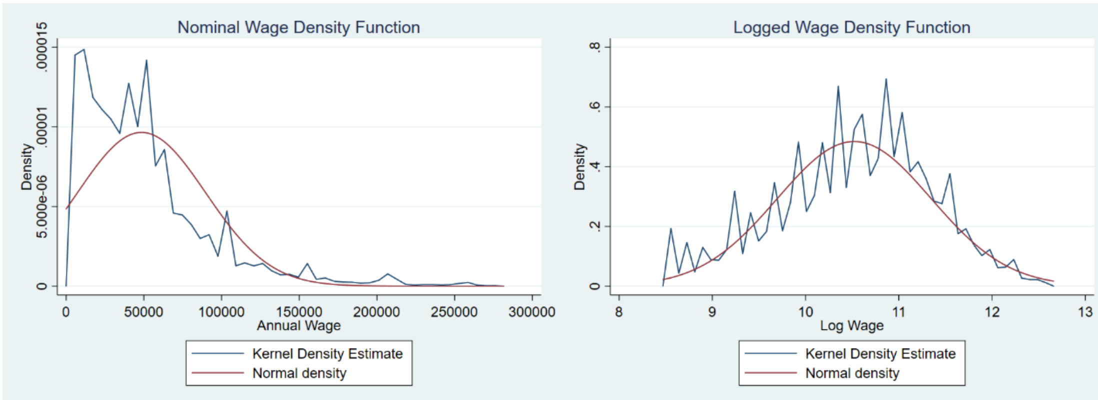

Figure 1. The Wage Density Function of Sample Population

Note: The figure of wage data was constructed using the 2022 American Community Survey 5-year PUMS.

Sample MSAs were selected based on the populations and are the 49 most populated cities of each state in contiguous United States. These 49 cities include huge metropolitan areas such as New York and small cities such as Delaware and Wilmington. The total population of these 49 cities accounted for roughly 11.3 percent of the U.S. population in 2022. These sample cities are categorized into four groups according to the definition of city size of OECD (Organization for Economic Co-operation and Development): large metropolitan areas having a population of 1.5 million or more; metropolitan areas with a population between 500,000 and 1.5 million; medium-sized urban areas comprising a population between 200,000 and 500,000; and small urban areas containing a population between 50,000 and 200,000. Of our sample MSAs, the population varied widely from small cities like Burlington to large metropolitan areas like New York City.

The metropolitan area measures, including amenities, come from (American Association of Retired Person) AARP website and City-Data website, which are U.S. based interest groups, who collect data from different sources such as government and private institutions, and rank communities and cities using the method called AARP Livability Index, a score system adopted to calculate the average of seven livability category scores: housing, neighborhood, transportation, environment, health, engagement, and opportunity. City-Data researchers gather and process the data from a wide range of sources including American Community Survey, Bureau of Labor Statistics, National Weather Service, US Census TIGER Database, Federal Housing Finance Agency Mortgage Data. In addition, regional amenities such as regional housing price, crime, air quality, commute time and job opportunities are also provided. Based on the data collected, we constructed the regional attributes table, presented in (Table A.1) in Appendix.

The amenities are measured in accordance with the following conditions: market density, which is the number of market and groceries within half a mile; Crime index, which is the ratio of local violent crimes per 10,000 population to national average; and traffic index, which is the ratio of average commute time per year to national average. The air quality index refers to the number of unhealthy days per year divided by national average, implying that high number leads to bad air quality. Entertainment and job index are the ratio of total number of facilities and job positions within a city to those within U.S. Additionally, higher market density areas such as Newark and Philadelphia tend to have higher crime rates.

EMPIRICAL RESULTS

In this sector, we present our empirical analysis and document the stylized facts on the effects of individual characteristics on income, housing cost, migration decision, and location choice.

Wage and Housing Cost Estimation

We estimated the wage and housing cost equation for each sample city by using data from all working-age populations. (Table 1) only presented the regression results of each city category. This paper takes the single, white female with high school degree as the reference group.

Table 1. Illustrative Wage and Housing Cost Equations

	Small Urban		Medium-size Urban		Metropolitan		Large Metropolitan	
Variables	Wage	HC	Wage	HC	Wage	HC	Wage	HC
Age	0.006***	-0.005***	0.005***	-0.004***	0.004***	-0.006***	0.003***	-0.004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Married	0.305***	0.356***	0.257***	0.338***	0.232***	0.267***	0.232***	0.273***
	(0.013)	(0.008)	(0.010)	(0.006)	(0.007)	(0.004)	(0.009)	(0.005)
Separated	0.048	0.104***	0.024	0.088***	-0.045***	0.023**	0.038*	0.037***
	(0.033)	(0.020)	(0.026)	(0.015)	(0.017)	(0.010)	(0.020)	(0.012)
Widowed	0.186***	0.022	0.083***	0.028*	0.065***	0.004	0.006	-0.029**
	(0.034)	(0.020)	(0.027)	(0.015)	(0.018)	(0.010)	(0.024)	(0.014)
Divorced	0.207***	0.105***	0.164***	0.068***	0.151***	0.029***	0.165***	0.049***
	(0.017)	(0.010)	(0.013)	(0.008)	(0.009)	(0.005)	(0.012)	(0.007)
Black	-0.269***	-0.146***	-0.266***	-0.123***	-0.259***	-0.148***	-0.263***	-0.186***
	(0.015)	(0.009)	(0.010)	(0.006)	(0.007)	(0.004)	(0.010)	(0.006)
Asian	-0.171***	-0.053**	-0.076***	-0.099***	-0.133***	-0.078***	-0.142***	-0.015**
	(0.036)	(0.021)	(0.022)	(0.012)	(0.015)	(0.008)	(0.014)	(0.008)
Other	-0.237***	-0.037***	-0.244***	-0.095***	-0.204***	-0.127***	-0.188***	-0.074***
	(0.023)	(0.014)	(0.019)	(0.011)	(0.012)	(0.007)	(0.012)	(0.007)
Male	0.432***	0.014**	0.435***	0.015***	0.341***	0.009***	0.340***	0.004
	(0.011)	(0.006)	(0.008)	(0.005)	(0.006)	(0.003)	(0.007)	(0.004)
Some college	0.201***	0.112***	0.249***	0.120***	0.244***	0.124***	0.337***	0.164***
	(0.015)	(0.009)	(0.012)	(0.007)	(0.008)	(0.005)	(0.010)	(0.006)
Bachelor	0.581***	0.262***	0.692***	0.321***	0.719***	0.359***	0.814***	0.398***
	(0.014)	(0.008)	(0.011)	(0.006)	(0.008)	(0.004)	(0.010)	(0.006)
Graduate	1.003***	0.403***	1.110***	0.507***	1.166***	0.577***	1.213***	0.579***
	(0.017)	(0.010)	(0.013)	(0.008)	(0.009)	(0.005)	(0.011)	(0.007)
Constant	9.405***	6.826***	9.443***	6.820***	9.509***	6.931***	9.510***	6.922***
	(0.024)	(0.014)	(0.019)	(0.011)	(0.013)	(0.007)	(0.017)	(0.010)
# of Obs	35,986	35,986	65,677	65,677	140,408	140,408	88,132	88,132
R-squared	0.188	0.143	0.202	0.175	0.194	0.193	0.188	0.167

Standard errors in parentheses. Levels of significance are presented as follows: *** $p<0.01$, ** $p<0.05$, *$p<0.1$.

The signs of coefficients are consistent with expectations. Wages are higher for married, white males with higher educational levels. Divorced individuals reported higher income than single workers. A possible explanation is that divorced individuals are older than singles, and the average age of divorce is 46. Income increases with education levels and age which both improve human capital.

For workers in small urban areas, gender discrimination in wage income is more pronounced, as indicated by the gender coefficient of 0.432 greater than 0.340 in large metropolitan areas. The advantage of higher educational level becomes more significant as we move from small urban to large metro areas. For each education degree, the positive effect is amplified by the city's size. Take graduate degree for instance, its coefficient in small city is 1.003 and increases to 1.213 in large metro-areas. This situation can be explanation by the diverse industry structure, where certain job opportunities are exclusively available in larger cities. The positive effects of age and marital status on wages exhibit a gradual decline as we transition from small to large cities.

For housing cost analysis, most signs of coefficients are consistent with the wage equation, suggesting a strong association between housing expense and income. As individuals age, their housing expenses tend to decrease. One possible explanation is that older people often have paid off their mortgages or benefit from lower rent or insurance premiums. Male respondents generally report higher housing costs than females. However, in large metropolitans, the gender effect becomes insignificant due to independence and higher income of female workers. The positive and significant coefficient related to marriage suggests that married individuals tend to have higher housing costs than single people, which is attributable to a larger family size as well as joint expenses. Compared to white individuals, black and Asian population spend less on the housing services. The impact of education attainment is positive across all cities. Individuals with higher education level are more likely to secure higher payment jobs, and thus allocate more resources to housing.

Migration Choice

Individual's migration decision is estimated separately for different sizes of city. The decision of migration is mainly explained by the individuals' attributes. The results for the logit regression model are presented in (Table 2).

Table 2. Logit Regression Results for Migration Decision

Variables	Small Urban	Medium-size Urban	Metropolitan	Large Metropolitan
Age	-0.063***	-0.066***	-0.064***	-0.062***
	(0.001)	(0.001)	(0.002)	(0.002)
Work Hours	-0.006***	-0.004***	-0.005***	-0.008***
	(0.001)	(0.001)	(0.001)	(0.002)
Separated	0.390***	0.169	0.045	-0.167
	(0.091)	(0.103)	(0.106)	(0.132)
Divorced	0.220***	0.292***	0.129**	0.319***
	(0.047)	(0.056)	(0.056)	(0.081)
Widowed	0.553***	0.523***	0.809***	0.788***
	(0.115)	(0.138)	(0.124)	(0.177)
Married	-0.521***	-0.271***	-0.333***	-0.120**
	(0.036)	(0.040)	(0.038)	(0.055)
White	1.310***	0.896***	0.894***	0.956***
	(0.052)	(0.047)	(0.047)	(0.066)
Asian	1.374***	1.039***	0.972***	0.545***
	(0.098)	(0.088)	(0.083)	(0.095)
Other	0.798***	0.469***	0.338***	0.148
	(0.076)	(0.080)	(0.076)	(0.094)
Male	0.496***	0.574***	0.412***	0.412***
	(0.030)	(0.035)	(0.033)	(0.047)
Some college	0.070*	0.185***	0.243***	0.634***
	(0.039)	(0.047)	(0.048)	(0.065)
Bachelor	-0.316***	-0.042	0.091*	0.363***
	(0.039)	(0.046)	(0.047)	(0.064)
Graduate	-0.202***	0.282***	0.181***	0.603***
	(0.050)	(0.057)	(0.055)	(0.075)
Q2 income	-1.438***	-1.583***	-1.268***	-1.118***
	(0.057)	(0.066)	(0.064)	(0.100)
Q3 income	-3.829***	-3.804***	-4.093***	-4.275***
	(0.129)	(0.148)	(0.208)	(0.287)
Q4 income	1.751***	1.215***	1.695***	1.553***
	(0.241)	(0.268)	(0.354)	(0.314)
Q2 HC	-2.433***	-2.325***	-2.581***	-2.225***
	(0.085)	(0.092)	(0.106)	(0.131)
Q3 HC	-4.610***	-4.434***	-4.835***	-4.542***
	(0.229)	(0.255)	(0.340)	(0.298)
Q4 HC	-2.616***	-2.192***	-2.699***	-2.764***
	(0.245)	(0.273)	(0.357)	(0.322)
Constant	2.717***	2.825***	2.053***	2.888***

	(0.087)	(0.092)	(0.090)	(0.131)
Observations	32,549	24,788	29,913	12,219

Standard errors in parentheses. Levels of significance are presented as follows: *** p<0.01, ** p<0.05, *p<0.1.

As previously mentioned, both income and housing costs were divided into four groups based on quartile levels, ranging from the lowest to the highest values. Our reference group consisted of single, black females with a high school education earning income in the first quartile. Most of the coefficient signs align with our expectations. The negative coefficient for age suggests that households become less likely to move as they grow older. Stability in job positions often leads to a higher likelihood of staying in one place. Compared to single individuals, various marital statuses (such as separated, divorced, and widowed) increase the probability of migration. These marital status changes may prompt individuals to seek new living arrangements. Education attainment imposes various effects on the migration tendencies across different urban areas. In small urban areas, a higher education degree correlates with a greater likelihood of settling down. Conversely, in other urban areas, individuals with better educational backgrounds tend to move out. This finding points to this fact: larger cities offer diverse job opportunities, encouraging educated individuals to explore different options by migrating among cities.

The impact of income and housing costs remains consistent across various urban areas. In comparing different income quartiles, we observed a diversity of migration tendencies. Compared with the lowest income group (first quartile income group), second and third quartile income groups were found to be less likely to move out, while the rich group (fourth quartile income group) exhibited a higher likelihood of migration. One feasible explanation is that city-to-city migration involves both observable costs (such as moving expenses) and unobservable costs (such as job relocation and lifestyle adjustments). The incentives of the second and third quartile income group are the weakest, since they neither face the same extreme financial constraint as the poor group, nor are they as financially secure as the rich group (fourth quartile) who can easily cover these costs.

The coefficients of housing expenses consistently show a negative impact on migration probability, implying that individuals living in higher-quality places with elevated housing costs enjoy the current housing service and are less likely to move. Conversely, individuals with the lowest housing cost have the strongest incentive to move out.

Choice of Location

The location choice equation was estimated using conditional and nested logit model. In the nested logit model, the alternatives were grouped based on the city size. The results are presented in (Table 3).

Table 3. Metro-area Choice Estimation

Variables	NestedLogitModel options	ConditionalLogitModel choice
Markets	0.034***	0.026***
	(0.008)	(0.005)
Crime	-0.040	-0.038
	(0.065)	(0.047)
Commuting	0.445***	0.394***
	(0.114)	(0.073)
Air	-0.006	-0.011
	(0.021)	(0.016)
Entertainment	0.040***	0.030***
	(0.014)	(0.009)
Jobs	-0.011***	-0.009***
	(0.004)	(0.002)
Housing Cost	-2.083***	-1.682***
	(0.208)	(0.060)
Income	11.356*	10.655**
	(6.198)	(4.680)

Constant		
Observations	117,600	117,600

Standard errors in parentheses. Levels of significance are presented as follows: *** p<0.01, ** p<0.05, *p<0.1.

Both models yield similar results, which means the size of a city is not a determinant of location choice. The coefficients on area-specific variables illustrate the attractiveness of various city attributes. Market density and entertainment facilities impose significant positive effects on the probability of choosing a particular city. A city becomes more attractive when it can provide higher income to an individual. The crime rate was observed to impose no significant effect on location choice. One possible reason is that this paper does not control the different districts within a city. For instance, both Long Island and Manhattan Island are counted as New York city in the analysis, but these two districts have different crime rates. The counterintuitive scenario arises because migrants often choose low-crime-rate areas within a city, even if the overall city has a higher crime rate. The coefficient on job opportunity is negative, suggesting that individuals may not prioritize job availability when selecting a city. The possible reasons behind are that first, the price level or housing cost might be high and negatively affect the quality of life; second, the availability of work in remote areas allows people to choose a city based on personal preferences.

Theoretical Simulation Results

In this section, we simulated 10,000 individuals' location choice in the benchmark model. In addition, to examine magnitudes of the effects of income, and housing cost on location choice, we conducted several counterfactual experiments.

Model's fit with data: In the simulation, amenities are exogenous to migrants, while the housing cost and income are related to individuals' characteristics. Large metropolitans can provide a wide range of job positions and better-quality amenities and narrow the wage gap between different races and genders and amplifies the benefit of education attainment. These attributes of large metropolitans make them more attractive to migrants. (Figure 2) compares the percentage of migrants choosing each city calculated from data and from theoretical model. The simulation results capture the essential features of observations. The mean absolute distance (MAE) between the simulation results and the data is 0.00503, and the rooted mean square error (RMSE) is 0.00651. Consistent with data, the simulated results show that around 23% of migrants choose New York City, which makes it the biggest city. Los Angeles, Chicago, Houston, and Philadelphia take the second, third, fourth, and fifth place in our sample.

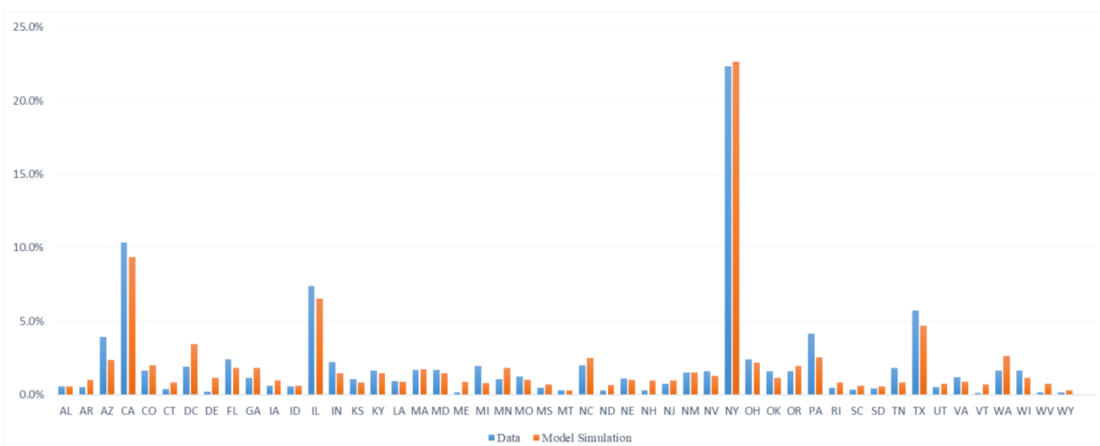

Figure 2. Location Choice, Data vs. Simulated Results

These small deviations of simulation outcomes from data are explainable by the unobserved attributes, considering that this study only examined a very limited scope of regional amenities, and it is possible that some other unmeasured attributes such as education, climate, temperature, or number of parks also impose significant impact on people's choice.

Counterfactual Experiments

In this section, we examined the response of migrants' location choice to the changes of housing cost and income through 100 paths of Monte Carlo simulation of 10,000 individuals' choices. We proposed two alternative scenarios: an increase in housing cost was granted; and each location is able to provide the migrants with a deterministic expected income based on their characteristics.

Raising Housing Cost: We conducted an experiment where we increased housing costs by 10% in one large metropolitan area and one small urban area, while keeping costs constant in other areas. This allowed us to investigate whether the negative effects of increased housing costs remained consistent across different MSAs. To simplify our analysis, we selected New York City as a representative of the large metropolitan areas and Denver to represent the small urban areas. (Table 4) indicated the simulation results, which takes the averages from 100 simulation paths. The first column was the data facts, the second column displayed the baseline results, the third column revealed the results when the housing costs in New York City increased by 10%, and the fourth column showed the results when the housing costs in Denver were raised by 10%.

Table 4. Counterfactual Experiment Results

Areas	Data	Model Simulation	NewYork increases by 10%	Denver increases by 10%	Areas	Data	Model Simulation	NewYork increases by 10%	Denver increases by 10%
AL	0.58%	0.57%	0.57%	0.57%	MT	0.28%	0.30%	0.30%	0.30%
AR	0.53%	1.03%	1.03%	1.03%	NC	2.00%	2.48%	2.48%	2.48%
AZ	3.95%	2.37%	2.37%	2.37%	ND	0.29%	0.65%	0.65%	0.65%
CA	10.37%	9.35%	***9.47%***	9.35%	NE	1.12%	1.00%	1.00%	1.00%
CO	1.64%	2.00%	2.00%	2.00%	NH	0.30%	0.98%	0.98%	0.98%
CT	0.39%	0.83%	0.83%	0.83%	NJ	0.76%	0.97%	0.97%	0.97%
DC	1.90%	3.43%	3.43%	3.43%	NM	1.49%	1.50%	1.50%	1.50%
DE	0.19%	1.15%	1.15%	1.15%	NV	1.60%	1.27%	1.27%	1.27%
FL	2.41%	1.83%	1.83%	1.83%	NY	22.36%	22.65%	***22.50%***	22.65%
GA	1.15%	1.83%	1.83%	1.83%	OH	2.40%	2.20%	2.20%	2.20%
IA	0.59%	0.97%	0.97%	0.97%	OK	1.59%	1.17%	1.17%	1.17%
ID	0.56%	0.62%	0.62%	0.62%	OR	1.60%	1.95%	1.95%	1.95%
IL	7.37%	6.52%	***6.53%***	6.52%	PA	4.17%	2.52%	2.52%	2.52%
IN	2.24%	1.47%	1.47%	1.47%	RI	0.49%	0.85%	0.85%	0.85%
KS	1.05%	0.85%	0.85%	0.85%	SC	0.35%	0.62%	0.62%	0.62%
KY	1.63%	1.48%	1.48%	1.48%	SD	0.42%	0.55%	0.55%	0.55%
LA	0.94%	0.90%	0.90%	0.90%	TN	1.81%	0.83%	0.83%	0.83%
MA	1.69%	1.72%	1.72%	1.72%	TX	5.74%	4.70%	4.70%	4.70%
MD	1.70%	1.48%	1.48%	1.48%	UT	0.51%	0.73%	0.73%	0.73%
ME	0.18%	0.87%	0.87%	0.87%	VA	1.20%	0.88%	***0.90%***	0.88%
MI	1.95%	0.80%	0.80%	0.80%	VT	0.12%	0.72%	0.72%	0.72%
MN	1.05%	1.83%	1.83%	1.83%	WA	1.66%	2.65%	2.65%	2.65%
MO	1.26%	1.03%	1.03%	1.03%	WI	1.63%	1.17%	1.17%	1.17%
MS	0.47%	0.70%	0.70%	0.70%	WV	0.14%	0.75%	0.75%	0.75%
					WY	0.16%	0.28%	0.28%	0.28%

Note: We interpreted the average of the 100 simulation results, using T-test (Student's test) to examine the significance of the changes.

The results illustrated that a 10% increase in housing cost of New York led to a decrease in migration by around 0.15%, a statistically significant change at the 1% level. However, the same increase in housing cost in Denver showed no significant changes, implying a weaker impact on migration patterns. The finding implies that in larger cities, people are more sensitive to the change in price than in small cities. This finding aligns with the economic principles, where higher prices lead to more elastic demand. The possible explanations are as follows. First, the housing cost of New York City is much higher than that of Denver, and therefore, a 10% increase brings a larger absolute increase in New York. Second, New York is densely populated because of the city amenities, and in comparison, the smaller city of Denver has a much smaller population. As a result, the population in Denver is much less diverse than that in New York City and less likely to be affected. In addition, we also observed increases in other cities such as Chicago, Los Angeles, and Philadelphia. It indicates that individuals leaving New York because of increasing housing costs, are choosing other large metropolitans with similar amenities qualities. Personal preference on housing services remains constant.

Deterministic Income: In this scenario, we assumed that each location provides a deterministic income for migrants based on their individual characteristics, meaning that each city offers sufficient job opportunities for each migrant, and that their wage income is perfectly predictable based on their characteristics such as gender, education, race, and other relevant attributes in each city. (Figure 3) showed the simulation results.

As can be seen, most migrants (over 80%) choose the large metropolitans, with New York City alone accounting for 75%. These simulated results were generated based on two assumptions: first, the amenities in a city can be equally enjoyed by all the residents, regardless of wealth, and housing cost; second, the housing cost and amenity externality remain unaffected by immigrants. This simulation reflected the historical migration pattern motivated by industrialization and urbanization (Lagakos et al. 2023). During this period, urban areas had high demand for labor and were capable of providing higher-income jobs relative to farming. In this model, most migrants were attracted to large metropolitans for two reasons: first, better-quality amenities and higher income provided by large metropolitans positively affect households' total utilities, attracting them to move in; and second, as mentioned above, these urban areas can provide the migrants with their expected incomes conditional on their characteristics. The heterogeneity in preference for housing services leads a small fraction of migrants to choose smaller urban areas.

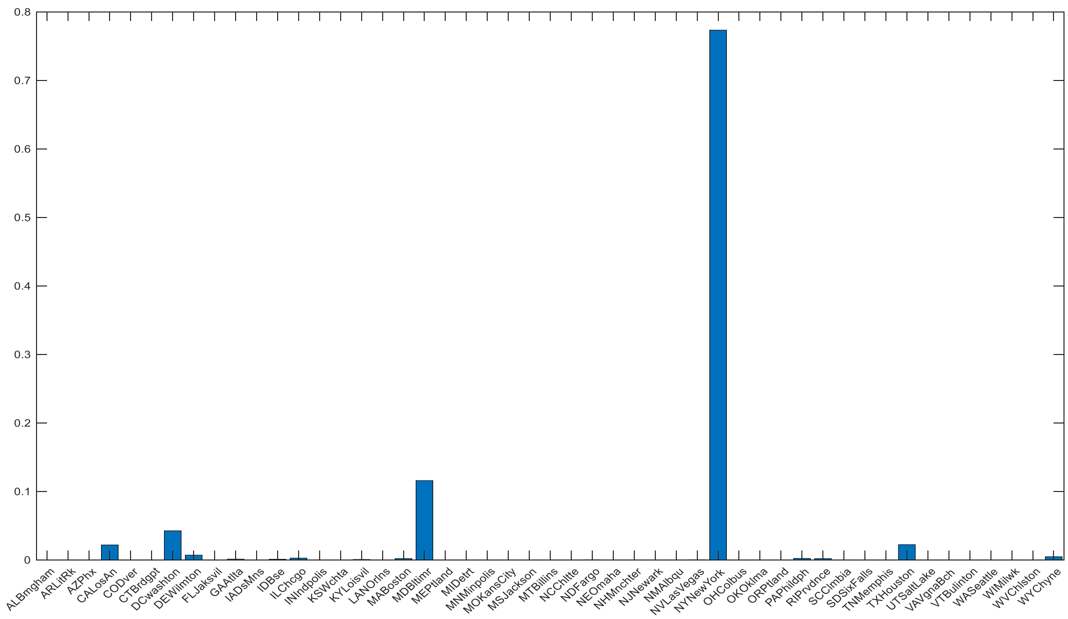

Figure 3. Location Choice of Migrations Under Deterministic Income Scenario

CONCLUSIONS

This paper applies both empirical and theoretical methods to examine the magnitudes and directions of the determinants influencing migrants' location choices. Heterogenous migrants exhibit varying preferences on housing service and individual characteristics (i.e., gender, race, educational level). Migrants make location choices to maximize their utilities which depend on the determinants such as regional amenities, expected housing cost, and wage. Housing cost and wage are both area- and individual-specific variables and are predicted based on individual's characteristics and unobserved regional factors. Furthermore, the magnitudes of these factors are evaluated through counterfactual experiments.

Among the selected amenities, market density and entertainment facilities were found to positively impact on location choice. The higher income, reduced gender discrimination, and enhanced education attainment benefits increases the likelihood of the larger metropolitan areas being picked, while the higher housing costs acted as a hindrance to the choices. This study also revealed the correlation between migration decisions and individual attributes. Compared with the median income group, the low-income (first income quartile) and high-income (fourth income quartile) exhibited a higher likelihood of migration, due to observable (moving cost, financial constraint) and unobservable (job relocation, lifestyle adjustments) costs. Individuals capable of covering the higher housing costs were more likely to enjoy the location amenity service and therefore to settle down. The theoretical simulations revealed that migrants were more sensitive to the housing cost changes in large metropolitans than in small urban areas. This aligns with the economics principles where higher prices lead to more elastic demand. This situation results from the higher absolute housing cost and more diverse and denser population in the large metropolitans.

For improvement on the present study, future research is expected to address other regional amenities such as population distribution, temperatures, urban price levels, and precipitation. Furthermore, large cities comprise a great number of districts, and each district has distinguishing attributes such as crime rate and commute times. This study took into consideration only the metro area as an integral unit, without giving attention to the different regions of a large city. In addition, the migrants were assumed to be price takers, where there is no dynamic interaction between immigration and local housing prices, which might amplify the effect of the income factor.

Notes:

According to the NBC news by Redfin in 2022, almost 50,000 people moved out of Bay Area. High mortgage rates, high home prices, and inflation have driven many homebuyers away.

The average house value in 2019 has reached 1.34 million dollars.

REFERENCES

Agius Vallejo, J., & Keister, L. A. (2020). Immigrants and wealth attainment: Migration, inequality, and integration. Journal of Ethnic and Migration Studies, 46(18), 3745–3761. https://doi.org/10.1080/1369183X.2019.1592872

Ammar, S., Duncombe, W., & Wright, R. (2001). Evaluating capital management: A new approach. Public Budgeting & Finance, 21(4), 47–69. https://doi.org/10.1111/0275-1100.00058

Sîrbu, A., Andrienko, G., Andrienko, N., Boldrini, C., Conti, M., Giannotti, F., ... & Sharma, R. (2021). Human migration: The big data perspective. International Journal of Data Science and Analytics, 11, 341-360.

Berger, G. (1992). Social Structure and rural development in the Third World. Cambridge University Press.

Berger, M. C., & Blomquist, G. C. (1992). Mobility and destination in migration decisions: The roles of earnings, quality of life, and housing prices. Journal of Housing Economics, 2(1), 37–59. https://doi.org/10.1016/1051-1377(92)90018-L

Bishop, K. C. (2008). Location Choice and the Value of Spatially Delineated Amenities.

Chen, Y., & Rosenthal, S. S. (2008). Local amenities and life-cycle migration: Do people move for jobs or fun? Journal of Urban Economics, 64(3), 519–537. https://doi.org/10.1016/j.jue.2008.05.005

Clark, W. A. (1992). Residential preferences and residential choices in a multiethnic context. Demography, 29, 451–466. https://doi.org/10.2307/2061828

Cun, W., & Pesaran, M. H. (2022). A spatiotemporal equilibrium model of migration and housing interlinkages. Journal of Housing Economics, 57, 101839. https://doi.org/10.1016/j.jhe.2022.101839

Davies, P. S., Greenwood, M. J., & Li, H. (2001). A conditional logit approach to US state-to-state migration. Journal of Regional Science, 41(2), 337–360. https://doi.org/10.1111/0022-4146.00220

Davis, M. A., Fisher, J. D., & Veracierto, M. (2021). Migration and urban economic dynamics. Journal of Economic Dynamics and Control, 133, 104234. https://doi.org/10.1016/j.jedc.2021.104234

Díaz, A., Jáñez, Á., & Wellschmied, F. (2023). Geographic Mobility over the Life-Cycle. https://doi.org/10.2139/ssrn.4340078

Eichman, H., Hunt, G. L., Kerkvliet, J., & Plantinga, A. J. (2010). Local employment growth, migration, and public land policy: Evidence from the Northwest Forest Plan. Journal of Agricultural and Resource Economics, 316–333.

Giannone, E., Li, Q., Paixao, N., & Pang, X. (2020). Unpacking moving. Unpublished Manuscript. http://congress-files.s3.amazonaws.com/2023-08/UnpackingMoving_paper_Aug23.pdf

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. Journal of Economic Geography, 1(1), 27–50. https://doi.org/10.1093/jeg/1.1.27

Gottlieb, P. D., & Joseph, G. (2006). College-to-work migration of technology graduates and holders of doctorates within the United States. Journal of Regional Science, 46(4), 627–659. https://doi.org/10.1111/j.1467-9787.2006.00471.x

Herkenhoff, K. F., Ohanian, L. E., & Prescott, E. C. (2018). Tarnishing the golden and empire states: Land-use restrictions and the US economic slowdown. Journal of Monetary Economics, 93, 89–109. https://doi.org/10.1016/j.jmoneco.2017.11.001

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression. John Wiley & Sons. https://www.google.com/books?hl=zh-CN&lr=&id=bRoxQBIZRd4C&oi=fnd&pg=PR13&dq=R.+X.+Sturdivant+(2013)&ots=kM1Ltp5Pda&sig=tjG-117kVLgQYLzkvES6I3eXD60

Hsieh, C.-T., & Moretti, E. (2019). Housing constraints and spatial misallocation. American Economic Journal: Macroeconomics, 11(2), 1–39. https://doi.org/10.1257/mac.20170388

Kennan, M. A., Lloyd, A., Qayyum, A., & Thompson, K. (2011). Settling in: The relationship between information and social inclusion. Australian Academic & Research Libraries, 42(3), 191–210. https://doi.org/10.1080/00048623.2011.10722232

Komissarova, K. (2022). Essays in Spatial Economics and Macroeconomics.

Lagakos, D., Mobarak, A. M., & Waugh, M. E. (2023). The welfare effects of encouraging rural–urban migration. Econometrica, 91(3), 803-837.

Mabille, P. (2023). The missing homebuyers: Regional heterogeneity and credit contractions. The Review of Financial Studies, 36(7), 2756–2796. https://doi.org/10.1093/rfs/hhac077

Meng, L., Xiao, X., & Zhou, Y. (2023). Housing boom and household migration decision: New evidence from China. The Journal of Real Estate Finance and Economics, 67(3), 453–479. https://doi.org/10.1007/s11146-021-09856-y

Peng, C.-Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. The Journal of Educational Research, 96(1), 3–14. https://doi.org/10.1080/00220670209598786

Plantinga, A. J., Détang-Dessendre, C., Hunt, G. L., & Piguet, V. (2013). Housing prices and inter-urban migration. Regional Science and Urban Economics, 43(2), 296–306. https://doi.org/10.1016/j.regsciurbeco.2012.07.009

Poston Jr, D. L., Zhang, L., Gotcher, D. J., & Gu, Y. (2009). The effect of climate on migration: United States, 1995–2000. Social Science Research, 38(3), 743–753. https://doi.org/10.1016/j.ssresearch.2008.10.003

Roback, J. (1982). Wages, rents, and the quality of life. Journal of Political Economy, 90(6), 1257–1278. https://doi.org/10.1086/261120

Schelling, T. C. (1971). Dynamic models of segregation. Journal of Mathematical Sociology, 1(2), 143–186. https://doi.org/10.1080/0022250X.1971.9989794

Van Nieuwerburgh, S., & Weill, P.-O. (2010). Why has house price dispersion gone up? The Review of Economic Studies, 77(4), 1567–1606. https://doi.org/10.1111/j.1467-937X.2010.00611.x