Determinants of the interstate migration: an analysis by both economic and non-economic factors in 2008-2018

Yuzhuo Kang

Econ580

Abstract

Over the past 20 years, millions of Americans migrated between states for various reasons. Though the migration trend allows some states to gain population growth, it also causes population loss, constrains the labor supply, and deteriorates the labor market conditions. It is crucial for local state policymakers to understand the roles of critical factors on migration rates to attract the residents to stay by adjusting the policies. Using the IPUMS-USA and CPS data from 2008 to 2018, this paper evaluates the roles of both economic and non-economic determinants on the state-level out-migration rate and individual level migration probability by applying the linear regression model and logistic regression model. The estimation results show that income, housing prices, unemployment rate, and demographic characteristics are significant determinants of the state migration flows. These factors also have considerable effects on individual migration propensity.

1. Introduction

According to the American Community Survey (ACS), 7.57 million people and nearly 3.8 percent of households migrated to different states in 2018. In the late 20th century, the distance of most state-level migration is short. In the last 20 years, there have been lots of long-distance movers. However, from the late 20th century, the annual migration rate in the U.S keeps decreasing. It dropped from 20% in 1990 to 11% in 2017. Many economists have modeled migration decisions with the assumptions of maximization of the expected net benefits on the choice of destination (Molloy et al., 2011). Individuals will migrate if the expected return at the destination outweighs the moving cost. Migration propensities reflect the differences in the economic conditions and other state characteristics between origin and destination regions.

Migration movements are classified into voluntary and involuntary migration (Clark et al. 1981). Voluntary migration decisions are more dependent on the individual's proactive behaviors to acquire more benefits. For instance, the possible reasons for voluntary migration include better wage offers and the change of marital status. In contrast, involuntary movement is more responsive to external economic activities, which stem from local or aggregate adverse economic recession (Bricker and Bucks, 2016). The factors that affect migration decisions are divided into three main categories. The first category is mostly independent of the individual, such as the economic condition of the state of residence. The second category is related to the individual but not entirely under the individual's control, such as the family background. The last category is individuals' internal characteristics, such as gender and race (Castelli, 2018).

There are two theories on the declining state-level migration trends. The first hypothesis is that the additional wage gains associated with changing employers become less and less, which could not motivate individuals to move (Molloy, 2014). Another theory is that the previous movers provide information to the individuals who prepare to move. It saves costs for the first-time mover to collect information (Kaplan et al., 2017).

Using data from the American Community Survey (ACS) and Current Population Survey (CPS) for all the 51 states from 2008 through 2018, this study focuses on how economic and non-economic determinants impact the interstate out-migration rate and individual migration choice. Economic factors include pre-tax wage, housing affordability, unemployment rate, income tax rate, and education. One standard deviation improvement in labor market conditions decreases the out-migration probability by 5 to 15 percent (Wozniak, 2010). Non-economic factors include demographic characteristics and state-specific features, including race, age, marital status, birthplace, and crime rate.

This paper uses two models separately to analyze the state out-migration rates and individual migration patterns. The multiple linear regression model applies state-level migration data and shows the significance of the impacts of different factors and one joint interaction effect between housing price and crime rate. The logistic regression model uses individual-level data and demonstrates how the individual migration probability changes as the level of each factor changes when other variables are controlled. The conditions in the origin and destination evolve continually, so the effects of factors also change over time. The impact of total migration time on the migration propensity is also discussed. These two models are closely related because the individual level migration probability could aggregate up to reach the state level migration rate.

The results of this study will inform policymakers information on the significant factors that impact the annual out-migration rate. The ultimate goal is to keep the population growth rate positive and gain increasing economic revenues. Local state policymakers correspondingly adjust the policies to decline the out-migration rates, attract residents of other states, and increase labor supply. For instance, the application of occupational licensing laws motivates workers to move out, deprives the economic gains from the labor force, and decreases the variation in labor market opportunities (Young and Mulholland, 2016). In addition, Cohen et al. (2011) suggest that the fast-growing income tax rate in New Jersey has an associated loss of over 125 million dollars in state total tax revenue.

Another motivation is that this study compares the migration probability of different demographic groups based on different factors, such as race and age. The state policymakers could use demographic information to compare the proportion of demographic groups in the successive years and predict the variation on the out-migration rate.

The estimation results of the state-level linear regression model show that the median wage, median housing price, unemployment rate, education, and age distribution, crime rate have significant impacts on the out-migration rate. The median income level has a greater impact on the gross out-migration flows than the housing affordability and unemployment rate. States with higher median housing price, lower median wage, lower unemployment rates, higher proportions of residents with more than 12 years of education, lower percentages of aged 30-60 residents have higher out-migration rates.

The analysis of the individual migration propensity using the logistic framework is more dependent on personal attributes. The estimation results show that the migration probability has negative associations with income and the number of children in the household. Unemployed workers, unmarried individuals, tenants, and individuals with the bachelor or higher degrees are more likely to migrate to other states. Individuals aged 16-30 are more likely to migrate than middle-aged and elderly adults. Also, migration probability for the individuals who moved more than once in five years is higher than the first-time movers controlling for other variables.

The last section of the paper compares the actual migration trend and the predicted migration trend in the model from 2008 to 2018. The result shows that the migration trend from model prediction is close to the actual trend except for the recession periods.

2. Literature Review

There were lots of interstate migration studies that focus on different determinants of migration decisions. Using CPS data from 1998 to 2008, Cohen et al. (2011) cite the case of New Jersey to claim that high-income tax rates motivate residents to migrate to states with lower tax rates. Controlling for demographic characteristics, the effect of the income tax rate on the migration rate is small because a 1% growth on income tax decreases the migration rate by 0.1%.

The limitation of this study is that the authors did not consider the distinction between high-income and low-income taxpayers. The impact of income tax on their migration propensity is different. The authors also did not control for shocks in different years when studying the impacts over long periods. My study improves by including the lowest and highest income tax rates in the model to explain the different migration motivations of rich and poor people. The state and year fixed effects are also included to consider the roles of unique unobservable factors on the migration flows.

Peng and Tsai (2019) investigate the effect of relative housing prices on the migration rates with city-level panel data during 1994-2016. The authors use quantile regression and panel cointegration methods and claim that the average housing price is cointegrated with the migration rate in the long run. However, the impact of housing prices is asymmetric. When the housing prices are below and above the average level, the associations with the migration rates are the opposite.

The limitation of this paper is that the authors did not consider the difference between the tenants and house owners. My paper will add the dummy variable to differentiate tenants and house owners and indicate the difference in their migration probability controlling for other variables. Another limitation is that the authors did not show the variation on the effects of factors during the long periods. My paper will improve this part by showing if the relationships between the factors and the migration rates change over time.

Sasser (2010) analyzes the impacts of labor and housing market conditions between origin and destination. Using CPS data, the author shows that unemployment insurance claims, average income, and average housing price differentials are significant determinants with the logistic framework. The drawback is that the author assumes migration choice is independent of the states besides origin and destination, but the spatial dependence cannot be overlooked in migration choice. My paper corrects this part by using the general characteristics of each state as the response variable instead of the differences between pair states. My paper also improves by comparing both the median and mean measurements of income and housing prices.

Other migration studies discuss the impacts of demographic characteristics. Using NLSY97 data, Gius (2011) quantifies the migration decision as the function of economic factors, individual characteristics, and moving costs. Gius compares the migration probability of different groups of people with the logistic framework and concludes that white people and aged 16-24 individuals are more likely to migrate. One drawback is that the author does not include state-specific features that evolve in the model. Another limitation of this paper is that the author did not consider the interaction between the economic and non-economic factors. My paper will improve this study by adding the interaction term between housing prices and crime rates to show if the impact of housing prices is the same when the crime rate is at different levels.

3.1 Modeling the factors of the out-migration rate at state-level

Table 1 displays the migration reasons based on the calculation from the Annual Social and Economic Supplement report between 2008 and 2018. The primary reason for moving is the housing-related reasons, including cheaper housing prices and the desire to own a home. On average, 43% cited housing reasons for all age groups. The other two prominent reasons are employment and family reasons, such as the acceptance of new job offer and changes in marital status. In the recent ten years, the housing issues are more considerable in migration decisions, which only account for 8 percent of movement reasons 20 years ago. Also, young and middle-aged people are more likely to migrate due to job and family reasons. The migration decisions of older people are more dependent on state features, such as climate and medical security.

Though the survey result shows that employment, housing, and family reasons are three main determinants of the migration decisions, the magnitudes of the impacts of these factors are unclear because the economic condition of every state evolves continuously. For instance, the labor market condition could change due to recession and the advance of technology in different years. More importantly, the survey result overlooks some critical economic factors like income tax rates. In order to show the relative significance and magnitude of impacts of economic and non-economic factors, the multiple linear regression model with Ordinary Least Square (OLS) assumption is applied as the following:

$$M_{it} = \boldsymbol{\beta_1} X_{it} + \boldsymbol{\beta_2}\, HP_{it} * CRIME_{it} + \boldsymbol{\lambda}_i + \boldsymbol{\delta_t} + \boldsymbol{\epsilon_{it}} \qquad (1)$$

X_{it} = [WAGE, LOWTAX, HIGHTAX, HP, UNEMPLOY, EDUC, AGE, WHITE, CRIME]

i: the index of state, $i = 1, \ldots, 51$
t: the index of year, $t = 2008, \ldots, 2018$

In equation (1), the economic factors include wage, income tax rate, unemployment rate, and housing prices, and education. In order to control the impacts of demographic characteristics and state-specific features on the out-migration rate, non-economic determinants,

Including age, race, and crime rate are included in the model. The definition of each variable is in Table 2. The factor levels measure the general characteristics of each state. The disparity of these factors between potential destinations and current location motivates the individuals to move out when the anticipated returns are higher in the destination. Migrants behave rationally and choose the destination that satisfies their maximized utility. It assumes that the annual conditions of the labor and housing market in each state change regularly.

The dependent variable M_{it} is the out-migration rate of the state *i* in year *t*. The state out-migration rate calculates the total outflows of residents from each state. It is expressed with the following equation:

$$M_{it} = \frac{\textit{the number of residents who move to other states}}{\textit{total population of state } i}$$

The residents who move abroad in year *t* and the individuals who move to state *i* in year *t* are not counted toward the total population. The out-migration rate M_{it} is between 0 to 1.

In the regression model, both the mean and median levels of wages and housing prices are used. One model contains average income and housing price in each state, and another model includes median income and housing price. Two measurements are compared to show which one has more explanatory power on the out-migration rate. Also, the purpose of adding both minimum and maximum tax rates in each state is to show how income tax affects the migration propensity of low-income and high-income workers.

There is also an interaction term between median housing price and the crime rate in state *i* in the model. The joint interaction effects between economic and non-economic factors also have critical effects on the out-migration rate. According to the migration study by Withers et al. (2008), racial groups have different reactions to the housing prices in the migration decisions. Many White people prefer housing-expensive regions, but most Hispanic people migrate to states with lower housing prices. The conjecture is that behavioral differences are affected by state-specific features. The reason to select the crime rate is that it is a unique state characteristic that changes over time. The hypothesis is that when the crime rate is at different levels, the impacts of housing price on the migration rate is the opposite. This model also measures the variation on the effects of factors in different years.

ϵ_{it} is the error term, which contains all the omitted variables in the model, and the omitted variables influence the migration rate. The bias of OLS estimators exists because of the correlation between the omitted variables and endogenous regressors, which are determined within the regression model. The bias of OLS estimators increases when the omitted variables have stronger associations with the endogenous regressors.

The inclusion of fixed effects is an effective measure to reduce the omitted variable bias. $\boldsymbol{\lambda_i}$ represents the state fixed effects. It is the dummy variable that accounts for the unobservable amenities and features unique to each state, such as culture and climate, which is constant over time. State fixed effects capture the heterogeneities across the states. $\boldsymbol{\delta_t}$ is the year fixed effect, which represents the unobservable factors that are constant across states but vary at different times. The fixed effects model reduces the omitted variable bias by diminishing the association between the endogenous factors and omitted variables.

3.2. Data source for the state-level regression model

Except for the income tax and the crime rate, the data source of all the factors is from ACS. ACS is an annual individual-level survey data. The sampling scheme of ACS is multi-stage and stratified sampling. The stratification is based on the geographic region and individual sampling probabilities. A systematic sample is drawn to represent each county and mailed survey form. Non-respondents are contacted via telephone, so the response rate is high.

The advantage of the ACS is its large sample size and broad scope. The sample size for each year contains more than 28,000,00 people and 3 million households nationwide, which is larger than other individual-level data. The scope is representative of all geographic areas with proper sampling weights. Other accessible data such as NLSY are not proper for the migration studies because the individuals are interviewed biennially. Migrators may move more than twice in the two-years interval so that the estimation may be biased. Aggregate data cannot show different migration propensity between individuals, such as employed and unemployed workers.

The data source for the income tax rate is the Federation of Tax Administrations. It contains information on the annual minimum and maximum income tax rates for each state from 2000 through 2018. It assumes that all individuals use standard personal exemptions and deductions. Crime rate data is from the Uniform Crime Reporting, the official FBI data. It records the annual crime rate of each state. These two data sources are administrative.

The state-level measurements of wages, housing prices, unemployment rate, education, age, and race are constructed based on the samples in the data. For instance, the unemployment rate of the state i is the total proportion of residents of the state i in the sample who are employed. When the individual data is aggregated by state, it forms the panel data. The state-level panel data contains the out-migration rate and measurement of variables for each state from 2008 to 2018, which is used to investigate the differences in migration flows between states. It also adds the dummy variables to control the fixed effects. The descriptive sample statistic of each factor is shown in Table 3. In the state-level panel data, the total sample size is 559. The observations of Alaska in 2008 and 2009 are dropped because the out-migration rates are more than 50 percent, which could bias the result.

3.3 Regression results of the impacts of factors on the out-migration rate at state-level

The explanatory variables measure the characteristics of the origin state. Table 4 shows the coefficients, standard errors, and significance for the determinants under different combinations from 2008 to 2018. The combinations distinguish by the inclusion of fixed effects, interaction term, the choice of mean and median, and the total number of determinants.

The fixed effects are excluded from Table 3, Column 1. The different results in the first two columns show that the impacts of all the economic factors except education are opposite when the fixed effects are included. The reason is that the OLS estimators in Column 1 have a large bias because of the correlation between the endogenous factors and the unobservable variables when the fixed effects are excluded. The bias decreases when the fixed effects are added, and the estimated impacts, as shown in Column 2, is closer to the theoretical value.

The explanatory power increases by adding fixed effects because R squared value increases from 0.113 to 0.887. 90% of the variation on the out-migration rate could be interpreted. Table 4, Column 2, shows that the out-migration rate increases when the state-level average wage decreases, income tax rate increases, average housing price increases, and the state unemployment rate falls controlling for other variables.

3.4 Controlling the effects of non-economic factors in the model

In order to control how demographic characteristics factors and crime rate affect the initial estimation results of economic factors, non-economic factors and interaction effects are excluded in the model to regress with only economic factors, as shown in Table 4, column 2.

Table 4 Column 3 shows the regression results after adding non-economic factors and an interaction term between crime rates and median housing price. Compare to column 2, R squared value only increases by 0.022, which shows that demographic characteristics and crime rates can only explain a small fraction of the variance on the out-migration rate.

Furthermore, the estimated coefficients of economic factors in column 2 change due to the inclusion of non-economic factors. The impacts of wage and income tax on migration rate change little. In contrast, the effects of housing price and unemployment rate vary a lot when the state demographic traits are included. It demonstrates that employment-related reasons and median housing price is vital to explain the different migration motivations between movers in different ages. For instance, the migration decision of older people is more dependent on housing prices, and middle-aged workers are more likely to migrate due to job opportunities.

3.5 Compare mean and median measurement

Wages and housing prices are two factors that use the mean or median of the sample to estimate the state-level quantities. The mean and median levels of these variables are compared to show which sample statistics better represents the distribution of the whole sampled data. The purpose is to raise the explanatory power and reduce the bias of the estimated impacts. In Table 4 Column 3, regressors contain average wage and average housing price. In Table 4 Column 4, explanatory variables contain median wage and median housing price.

The comparison between the third and fourth columns in Table 4 shows that using the median level of wages and housing prices increases the R squared value from 0.9088 to 0.9134 compared to mean statistics. Figure 1 shows that the density curve of the estimated migration rate predicted by the median wage and housing price is closer to the actual density curve. It is consistent with the speculation that the distributions of wages and housing prices in the sample are skewed. As a result, the median measurement of wages and housing prices is better.

3.6 The statistics and economic significance of factors

The relative impact of each factor on the migration rate is based on the sample statistics in Table 3. Table 4 Column4 shows that the median pre-tax wage and median housing price have significant effects on the out-migration rate. States with higher housing prices and lower median income levels have higher out-migration rates. Median wage has a more considerable role in migration flows than median housing price. When the annual median wage increases by 10000 dollars, the annual out-migration rate decreases by 4.67%. The out-migration rate only increases by 0.28% if the median housing price increases by 10000 dollars. The annual unemployment rate is also a significant determinant of the out-migration rate, but its impact is small because a 1% growth on the unemployment rate only decreases the migration rate by 0.029%.

The percentage of residents with 12 years or more education and the proportion of aged 30-60 individuals in the state are also significant factors that affect the out-migration rate. Table 4 Column 4 shows that the magnitude of the impact of the proportion of aged 30-60 individuals is greater than education. An additional 1% of individuals with 12 years or more education only increases the out-migration rate by 0.0732%. In contrast, a 1% increase in the proportion of aged 30-60 individuals decreases the out-migration rate by 0.147%. The crime rate is also a significant factor. States with 100 less violent cases per 100,000 residents lower the out-migration rate by 0.038%, so the impact of the crime rate is small. The impacts of non-economic factors are relatively smaller compared to economic factors.

3.7 Explanation with Economic theories on the significance of factors

Falling housing prices are harmful to the financial ability of the household and reduce the mobility of household members, which is called the equity lock-in effect (Peng and Tsai, 2019). The equity lock-in effect raises the interest rate and restricts population mobility due to the impact of negative equity. Negative equity happens when the house value is less than the owed amount of the property's mortgage. When the housing prices fall, the housing turnover rates (the transfer of ownership) fall correspondingly (Smith et al., 2014). Households spend the savings on down payment instead of house transactions, which lowers the likelihood to move. According to the research by Lee et al. (2011), a 10 percent increase in the number of equity-locked in households decrease the mobility rate by 13 percent.

Unobservable differences in local worker quality could cause wage differences across states. Workers face different wages because the returns to the same education quality and skills are different across states (Kaplan et al., 2017). Under this situation, individuals have a higher propensity to migrate because of the potential higher returns on the wage when the locational dispersion of wages is different. "Wage premium" is the term commonly applied in migration studies. It is the location matched parameter and reaches the maximum when the location has the biggest advantage on the returns to skills for each occupation. A higher value of the wage premium raises the migration probability (Kaplan et al., 2017). Each age group has different maximum values of the wage premium, so income has different impacts on the migration choice of different age groups.

The national recession could cause a high state unemployment rate and hurt the local economic conditions. Variation in the locational dispersion of economic opportunities during different business cycles causes the cyclicality of migration. The out-migration rate is lower at the end of the recession compared to the start of recession because the net gain on utilities and benefits are lowest during economic depressions (Saks et al., 2011). It is accordant with the prior estimation that the out-migration rate diminishes by 0.15% when the unemployment rises by 5%.

According to the study by Wozniak (2010), the location choices of college graduates are more dependent on the labor market conditions than less-educated peers. For college graduates, greater local market opportunities with high labor demand reduce the effects of other local conditions. In addition, the salary of college graduates is higher than other less-educated workers. Given a percentage wage change, college graduates have higher migration demands. The lasting impacts of entry labor market conditions on workers who graduated with college degrees are more significant than other workers (Wozniak, 2010).

3.8 Interaction effect between housing price and crime rate on the migration flows

The regression results show that the interaction effect between the median housing price and the crime rate has a significant impact on the out-migration rate. According to the study by Haug (2008), it is essential to consider the role of non-economic motivation on the effect of an economic factor on migration destination selection because it yields different returns. The difference between the preference on expensive and cheap housing in migration selectivity could be explained by the state-specific traits which change over time, such as crime rate.

Figure 2 shows how the state-level crime rate affects the impacts of housing prices on the migration flows. Y-axis represents the out-migration rate, and the X-axis represents the housing prices. Each line with different colors represents the estimated impacts of housing prices at different levels of the crime rate. Three levels of crime rate are mean level and one standard deviation above and below the mean level. One standard deviation of the crime rate is 180 per 100,000 residents. Figure 2 shows that when the crime rate is above the mean level, the housing prices and out-migration rate have a negative association. States with higher housing prices have lower out-migration rates, which demonstrates that migrants do not prefer states with lower housing prices where the crime rate is high. In contrast, when the crime rate is below the mean level, the association is reversed. States with lower housing prices have lower out-migration rates, so individuals are more likely to migrate to states with cheaper housing and low crime rate.

3.9 Robustness checks

The relationship between the out-migration rate and wages, housing prices, and the unemployment rate may not be linear. The effects of these factors are enhanced in the origin states where the median level of these factors is far higher or lower than the national median level, such as California and North Dakota (Sasser, 2011). The squared terms of wage, housing price, and unemployment rates are included in the model to certify the relationship, as shown in Table 5 Column 2. It shows that the squared terms of median wage and median housing price have significant impacts on the migration rate. The R squared value increases from 0.9134 to 0.9196. The median wage and housing price and their squared terms all have positive impacts.

Controlling for other factors, when the differential on the median wage increases from 10000 to 20000 dollars, the difference in the out-migration rate increases from 5.19% to 12.04%. When the median housing price differential increases from 10000 to 20000 dollars, the disparity on the out-migration rate increases from 0.483% to 1.378%. It demonstrates that the impacts of wages and housing prices are strengthened at the states with very high or low median wage and housing price levels. Also, the estimated impacts of non-economic factors and education do not have apparent change when the squared terms are added in the model.

Another robustness is to control for the out-migration rate in the previous year. The migrants who moved in the previous year provide information on the potential destinations. The information contains the job opportunities and local state amenities (Molly et al., 2011). The information saves the cost for the mover to acquire information, and the movers do not need to migrate based on the testable prediction at the destination (Kaplan, 2017). When the out-migration rate of the state *i* increases, more information is sent back to the state *i*. In addition, the migration rate in year *t-1* also has a relationship with all the past factors in the model. The factors in the previous year could bias the estimated effects of all the variables indirectly because the level of factors in year *t* is determined by the level of factors in the previous year. The inclusion of past-year migration rate reduces the bias of estimation.

Table 5 Column 3 shows that the past year out-migration rate is a significant factor, and the inclusion of it increases the R squared value from 0.913 to 0.928, so the explanatory power is strengthened. 1% increases in the out-migration rate in the previous year only increases the migration rate in year t by 0.092, so its impact is small. The initially estimated impacts of wage, housing prices, and unemployment rate decline, and the coefficients of demographic characteristics and crime rate change little.

3.10 Compare the impacts of factors in different years

In order to show how the magnitude of impacts of factors change over time, the model is estimated using data from two periods: 2008-2013 and 2014-2018, as shown in Table 6 Column 1 and Column 2. In the recent five years, the impact of the median wage and housing price on the migration rate decreases by 0.63% and 0.103% when the median wage and housing price changes by 10000 dollars. Median housing price loses significance in 2014-2018. It demonstrates that the wage gains by way of changing employers reduce, and the house equity lock-in effect is weakened. The effects of income tax and education years only have small differences. The impact of the unemployment rate diminishes by 0.079% because the effect on the migration rate was higher in the recession with deteriorated labor market conditions during 2008-2013. The impacts of demographic characteristics and crime rates vary little over time.

4.1 Modeling the factors of the migration probability at the individual level

In order to investigate the impacts of factors on the interstate migration decision at the individual level, it assumes that the migration decision is only dependent on the utility at the origin and destination and moving cost. One key to evaluating the utility at each state is the location-specific capitals. The utility associated with the location-specific capitals will diminish completely or lose partial due to the limited availability if the individual migrates to another place (Fischer et al. 1997). The capitals include both monetary and non-monetary factors. During the migration process, the weight of different factor changes, which causes the utility to change as well (Haug, 2008). Individual migrates to the state, which yields the highest net growth on utility by comparing all the possible destinations. Using the ancestral model by Gius (2011) as a guide, the utility at a given state is defined as the following equation:

$$V_{ij} = f\,(P_i, S_j, I_{ij}, T_j) \qquad (2)$$

$j = 1, \ldots, 51$

In equation (2), index i refers to each individual, and index j refers to the state of residence. P_i is the vector of individual attributes of person i. S_j represents the vector of fixed attributes of state j, such as the climate. I_{ij} is the vector of factors that are determined by both individual characteristics and the state conditions, such as income. T_j represents the changeable state-specific features, such as the income tax rate. V_{ij} is the utility for individual i at state j, which is a function, represented by f, of the above vectors. Individual i migrate when the differences in utility between origin and destination satisfy the following criterion:

$M_i = (V_{id} - V_{io}) - C > 0$ (3)

The index *o* and *d* refer to the origin and destination state, and C is the moving cost. $V_{id} - V_{io}$ represents the utility disparity for individual *i* between the destination and origin. Individual *i* migrate if the net gain on utility is larger than the moving cost and becomes maximized at destination *d* through the pairwise comparison on all potential destinations. Because the specific value of utility is unmeasurable, the dummy variable M_i in equation (3) is applied to represent whether the individual *i* moves to another state or not. M_i is also the differentials between the net utility gains and the moving cost. This study uses CPS data which contains the migration status in different years. It is a dummy variable indicating if the individual moved in year *t*. If M_i equals to 1, the individual migrated to other states in year *t*. Otherwise, M_i equals to 0.

The model is the logistic regression. Economic and non-economic factors are included in the model to explain how migration probability alters as the level of each factor changes controlling other explanatory variables. The model framework is in the following equation:

$\text{Logit}(P_{ijt}) = \boldsymbol{\beta} X_{ijt} + \boldsymbol{\lambda}_j + \boldsymbol{\delta}_t + \epsilon_{ijt}$ (4)

X_{ijt} = [WAGE, TAX, EMPLOY, RENT, EDUC, AGE, RACE, MARITAL, MOVE, MEMBER, BIRTH]

t: the index of year, $t = 2008, \ldots, 2018$

i: the index of individual

j: the index of state

In equation (4), the response variable is the dummy variable M_i in equation (3). P_{ijt} is the migration probability for individual *i* who live in state *j* in year *t*. P_{ijt} denotes the probability when M_i equals one. In the logistic framework, P_{ijt} changes as the level of factors changes, and the migration probability for individual *i* could be calculated by the following equation:

$$P_{ijt} = \frac{e^{(\beta * X_{ijt} + \lambda_j + \delta_t + \epsilon_{ijt})}}{1 + e^{(\beta * X_{ijt} + \lambda_j + \delta_t + \epsilon_{ijt})}} \tag{5}$$

In equations (4) and (5), X_{ijt} is the vector of all the factors. Economic factors include income, employment status, and marginal income tax rate, housing ownership, and education. Non-economic factors include age, race, marital status, number of children, migration status, and birthplace. The definition of each variable is in Table 7. Demographic attributes also impact the effects of economic factors on migration propensity. For instance, individuals will not move to states with higher income if individuals live with family members. Personal circumstances may have a noticeable change in successive years, which becomes the motivation to move.

$\boldsymbol{\lambda}_j$ is the unmeasurable unique state characteristics which do not evolve. $\boldsymbol{\delta_t}$ is the unobservable year specific features that affect the migration decision, such as the recession. The inclusion of fixed effects reduces the omitted variable bias. ϵ_{ijt} is the error term that contains all the omitted variables. The advantage of the logistic model is its efficiency and accuracy when the dependent variable is the dummy variable. The overfitting problem is alleviated, and the non-linear effects could also be predicted with the logistic framework.

4.2 Data source for the individual-level logistic regression model

The data source for all the variables is from CPS (Current Population Survey). CPS is a monthly survey of about 60,000 households conducted by the U.S Census Bureau Labor statistics. Households are mailed questionnaires. The sampling scheme is multi-stage stratified sampling. The stratification is based on the large counties, social and economic characteristics, and employment status. The selection probability is proportional to its total population.

The advantage of CPS over other individual-level datasets is that the household is interviewed for four successive months, and the household is interviewed again after the eight months. It shows the change in the migration status and other related information in consecutive years for each individual in the sample. CPS contains more desired variables for this study, such as the migration status in the last five years. CPS forms the individual-level panel data, which contains the migration status of migrators and associated variables in different years. The descriptive sample statistic of each variable is in Table 8. The sample size is 198367 for each year's CPS data. The total sample size is 2150168. 33039 observations are dropped because the individuals in the dropped sample moved abroad or have missing migration status.

4.3 The connections of state-level linear regression model and individual level logistic model

The state-level linear regression model and individual-level logistic model are closely related. The migration rate of state *i* equals the migration probability of all the residents of state *i*:

Migration rate = $M_{it} = P_{it} = \frac{total\ number\ residents\ who\ moved\ to\ other\ states}{total\ population}$ $\quad i: state, t: year$

Both models use the panel data and investigate the impacts of each determinant on the migration probability. The first model groups all the individuals level data, including migration status and other variables in the sample, by the states and calculates the migration rate and other related variables at the state median level in every state. For instance, the unemployment rate of state *i* is derived by counting the proportion of employed individuals among all the labor forces. It expects that these two models yield similar conclusions on the significance and association between the migration probability and different factors.

4.4 Regression results of the impacts of factors on the migration probability at individual-level

Table 9 shows the significance and estimated impacts of all the factors from 2008 to 2018. The estimated coefficients represent the effects on logarithmic of the odds of migration. The difference between the first and second column specification is the inclusion of fixed effects.

The estimation results are robust to the inclusion of fixed effects because the explanatory power increases, and the residual deviance reduces from 653.3 to 639.8. The state and year fixed effects also help reduce the omitted variable bias.

4.5 Economic Significance and Statistical significance of all the factors

According to Table 9, Column 2, pre-tax wage, employment status, and the ownership of the house have significant impacts on the migration probability. For all the personal attributes, the education attainment, marital status, age, number of children in the household, and the migration status in five years are significant determinants of the interstate moving propensity.

The relative magnitude of the impacts of each factor is based on the sample statistics in Table 8. The coefficients are transformed by exponential function to show the direct change on migration probability for various levels of factors. When the wage increases by 10000 dollars, the migration probability decreases by 6.86%. Controlling other variables, employed individuals are 11.5% less likely than the unemployed individuals to migrate. Also, the interstate migration probability of tenants is 1.23 times the migration probability of individuals who own houses. The result shows that the impacts of housing ownership on the migration propensity are larger than wage and employment status.

For the personal characteristics, the migration probability of individuals with the bachelor or higher degrees is 1.19 times the probability for individuals without college degrees. Individuals who have married are 9.3% less likely to migrate than individuals without marriage. The migration probability of aged 16-30 individuals is 1.14 times the probability of aged 60 or more individuals. Aged 60 or older individuals are 12% more likely to migrate than aged 31-60 individuals. When there is one more child in the household, the migration probability decreases by 9.1%. The migration probability for multiple-time movers is 1.1 times the probability of first-time movers. The result demonstrates that the effect of age is the largest.

4.6 Explanation of the significance with economic theories

Middle-aged workers are more likely to employed than young workers because senior workers could bring enterprises more improvement in economic benefits. The abilities and experiences of middle-aged workers accumulate over time. There exists a significant difference in the variability of wages over time for the same individual (Kennan and Walker, 2011). High wages and the recognition of current employers reduce the necessity for senior workers to migrate to achieve the ideal position. In contrast, young workers need to compare the wage offers in different states to find more employment opportunities. In the sample, the median wage of young workers aged 16-30 is 36000 dollars, and the median wage of middle-aged workers aged 31-60 is 45000 dollars. 15% of the young workers are unemployed, and only 6% of the middle-aged workers are unemployed. The estimation result shows that the migration probability of young workers is 1.27 times the probability of middle-aged workers.

The migration spillover effect means that the growth of the proportion of middle-aged workers causes lower equilibrium migration rates for all the workers in the labor force market (Karahan et al. 2014). The estimation shows that the net out-migration rate declines by 1.47% when the proportion of middle-aged workers increases by 10%. Middle-aged workers are in a less dynamic labor market restricted by locations. According to the study of Karahan et al. (2014), firms prefer to employ aged 30 to 60 workers in the local labor market. The wages for the local workers are lower compared to workers from other states because of the lower moving cost of local middle-aged workers. Middle-aged workers also prefer local companies because it allows staying together with family members.

The migration likelihood of the migrators who moved more than once in five years is higher than the first-time mover controlling for other variables. The individuals who move multiple times are less influenced by the emotional and family ties compared to individuals who never move before. Past migration experiences also impact the migration decisions of multiple-time movers (Gius, 2011). My study shows that multiple-time movers are 10% more likely to migrate than movers who do not move before in five years.

5. The comparison of actual migration trend with predicted migration trend

Figure 3 displays the migration trends of the predicted migration rate from the linear regression model and the actual out-migration rate. The primary motivation is to show if the predicted migration rate follows the trend and variation of the actual migration rate from 2008 to 2018. Another motivation for this comparison is to show the association between economic factors and migration trends. For instance, the migration rate was low when the nation experienced job losses and significant depreciation on housing prices.

In this figure, the trend of the predicted migration rate is close to the actual trend from 2008 to 2018. However, the estimated out-migration rates in 2008 and 2009 in the recession periods have more substantial differences compared to other years. It demonstrates that the out-migration rate is more affected by unobservable state characteristics in the recession periods. Also, Figure 3 shows that the out-migration rates in 2008 and 2009 are lowest in the last ten years, which could be explained by the effects of economic factors. In the recession periods, the median housing prices underwent depreciation, which fell from 208000 dollars to 183411 dollars. The national unemployment rate in these periods increased from 5.37 % to 8.39%. The fluctuation of the housing prices and the unemployment rate had remarkable impacts on the out-migration rate.

The figure also shows that the recovery of the out-migration rate in 2010 is slow. It kept at a low level. The reason is that the recession limited the ability of migrants to relocate because the effects of the recession were more severe in cyclically and trade-sectors which hire lots of migrant workers, and these effects recover very slowly. In addition, in order to minimize the disruption in the local labor market, some local state migrant-receiving institutions stopped migrator recruitment. Also, in the 2008 financial crisis, the slow adjustment of the labor market to the shock hindered workers from moving to states with better jobs (Kaplan et al., 2017).

6. Conclusion and future research plans

This paper analyzes the impacts of economic factors and non-economic factors on the state out-migration rate and individual migration probability. Using ACS data and linear regression model, this study shows that median pre-tax wages, median housing prices, and unemployment rate, education, age distribution, and crime rates have significant impacts on the migration rate. Median wages and the proportion of aged 30-60 individuals have more considerable influences than other determinants. Housing price has an inverse association with the migration rate when the crime rate is at different levels.

Also, using CPS data and logistic framework, this paper demonstrates that wage, employment status, housing ownership, education, marital status, age, number of children, and total migration times are significant determinants of the migration probability. The migration probability of unmarried and unemployed individuals, tenants, and individuals with higher degrees is higher. Aged 30-60 individuals are less likely to migrate than young and older people controlling other factors. Education attainment and housing ownership have more considerable roles in the migration propensity. The multiple-time movers are more likely to migrate than first-time movers.

This study also indicates that the out-migration rate was lower compared to other periods during the recession because of the high unemployment rate and depreciated housing prices, and the unmeasurable variables have more substantial impacts on the migration rate in the economic depression. The migration rate in the following year of recession is still low because of the slow adjustment of the labor market.

The limitation of the data is that the median wage and housing price are calculated from the ACS samples instead of national-level statistics, so it may deviate from the actual level of these variables and bias the estimation of impacts. The income and housing prices in different years are also affected by inflation, so the differences in the impacts over the years may not be reflected. In addition, the total migration time in the CPS data contains many missing values, so the estimated effect of total migration times on the migration probability is not reliable.

The limitation of the methods is that when accounting for the impact of the housing prices, it is necessary to consider the change in wages. Individuals will still migrate to states with higher housing prices if their income increases a lot. The impact of the unemployment rate may be biased because many workers quit the job before migrating, and their employment status becomes unemployed. Also, the interaction terms between the variables are not enough. The effect of factors like housing prices may change at different levels of other factors. Another limitation is the omitted variable bias. It still exists because the error term contains variables that correlate with the factors, such as the unique state characteristics which change over time.

The limitation of the models is that the logistic regression model does not consider the moving cost. The estimation of migration propensity without the moving cost as a predictor may be biased. It also does not show how the individual migration propensity differs between states.

The estimated impacts may vary between states. In the linear regression model, the relationship between migration rate and some variables may not be linear.

There are plans to improve this study and address the above limitations. The first step is to find data on wages and housing prices, which represent the level of the general population in the U.S. I will also use CPI to make income, and housing prices from 2008 to 2018 be measured in real ($2018) dollars. Another task is to find individual-level data that contains more variables related to migration. The desired variables are total migrations time in more extended periods and a dummy variable indicating if the migrator returns to their hometowns. Migration studies show that many movers go back to their hometown, and my goal is to study why and when this phenomenon happens. I plan to finish this part during the summer.

The second plan is to improve the methods. I will use the housing affordability, the ratio of annual housing mortgage amount to income, and annual unemployment insurance claims to replace housing prices and the unemployment rate in the model. These two estimators may better reflect the effects of housing and labor market conditions on the migration rate. More interaction terms such as wage and income tax will be added in the model to show how the impact of wage varies. In order to solve the problem of omitted variable bias, I will add instrumental variables or other ways to reduce the bias of estimation, and I plan to finish this part during the summer.

During the next semester, I plan to further improve the models. For each observation, I will record the moving distance between two states in the panel data and use the moving distance as the additional predictors. I also hope to show the differences in migration propensity between states. The next step is to convert the migration rate by logarithm and compare the estimated results with the linear regression model. Finally, I plan to analyze the net in-migration rate and compare it with the model and estimation results using the out-migration rate.

TABLES AND FIGURES

Table 1
Main Reasons for individuals migrating across states by age, 2008-2018.

Reasons for moving	Age 16-29	Age 30-59	Age 61-100
Housing-related reasons	42.67%	45.39%	42.59%
Employment-related reasons	19.02%	21.21%	15.17%
Family reasons	30.43%	26.74%	28.88%
Change of climate	0.256%	0.415%	0.908%
Other reasons	4.88%	5.65%	12.2%

Source: Calculations on the Annual Social and Economic Supplement (ACES), 2008-2018

Table 2
Definition of variables and data sources in the state-level multiple linear regression model

Variable	Definition	Data Source
WAGE	mean/median pre-tax annual wage in state i	ACS
HOUSE	mean/median housing price in state i	ACS
LOWTAX	minimum income tax rate in state i	Federation of Tax Administrations
HIGHTAX	maximum income tax rate in state i	Federation of Tax Administrations
UNEMPLOY	unemployment rate in state i.	ACS
EDUC	The proportion of residents who have more than 12 years of education in state i	ACS

AGE1	The percentage of residents in state i whose age is between 16-29	ACS
AGE2	The percentage of residents in state i whose age is between 30-60	ACS
AGE3	The percentage of residents in state i who is older than 60.	ACS
WHITE	The percentage of residents whose race is white	ACS
CRIME	violent crime rate per 100,000 residents in state i	ACS

Table 3
Descriptive Sample Statistics for all the variables from 2008 to 2018 in the state-level multiple linear regression model

Variables	Mean	SD	Range
Response variable			
Out migration rate	5.733%	5.29%	1.51% to 40.06%
Economic factors			
Pre-tax wage	46381.69	5643.448	32807.56 to 88049.92
Minimum income tax rate	2.1803%	1.7555%	0% to 6%
Minimum income tax rate	5.3855%	3.0734%	0% to 12.3%
Housing price	263870	90304.8	124004.8 to 827811.9
Unemployment rate	6.68%	2.38%	2.562% to 14.02%
Non-economic factors			
Proportion of residents with more than 12 years education	0.40234	0.051	0.2889 to 0.6206
Proportion of aged 16-29	0.1734	0.01692	0.1437 to 0.2783
Proportion of aged 30-60	0.3833	0.01796	0.3319 to 0.4374
Proportion of aged 61 and older	0.2511	0.03279	0.1336 to 0.3492
Proportion of white residents	0.8180	0.1237	0.4018to 0.9801
Violent crime rate per 100,000	385.4164	181.357	99.3 to 1437.7

Source: Calculation based on the ACS, UCR, and Federal Tax Administration data.

Note: the wage and housing prices in the sample are at median level, and there are total 559 samples.

Table 4
Regression results between the state-level out-migration rate and each determinant

Independent variable	Dependent variable: out-migration rate			
	1	2	3	4
Average annual pre-tax wage	3.93*** （0.71）	-2.91* (1.23)	-3.49* (1.38)	_
Median pre-tax wage	_	_	_	-4.67** (1.52)
Minimum income tax rate	-0.213 （0.185）	0.0083 (0.032)	0.026 (0.032)	0.0241 (0.0302)
Maximum income tax rate	-0.306** (0.109)	0.027 (0.025)	0.015 (0.029)	0.011 (0.023)
Average housing price	-0.093* （0.042）	0.76 (0.692)	0.201** (0.0638)	_
Median housing price	_	_	_	0.283*** (0.0819)
State Unemployment rate	8.33 (7.90)	-0.396** (0.129)	-0.275 (0.147)	-0.295* (0.149)
Proportion of residents with more than 12 years education	_	0.0328 (0.0301)	0.0276 (0.0252)	0.0732** (0.0259)
Proportion of residents aged 16-29	_	_	0.02182 (0.0363)	0.0347 (0.0370)
Proportion of residents aged 30-60	_	_	-0.1562*** (0.0365)	-0.1469*** (0.0367)
Proportion of residents aged 61 and older	_	_	0.0592 (0.0401)	0.0312 (0.0408)
Proportion of residents whose race is white	_	_	-0.107 (0.103)	-0.034 (0.106)
State crime rate	_	_	0.031* (0.0151)	0.038* (0.0162)
Interaction term included?	No	No	Yes	Yes
State fixed effects included?	No	Yes	Yes	Yes
Year fixed effects included?	No	Yes	Yes	Yes
R-squared	0.1129	0.8868	0.9088	0.9134

Note:
The numerical value in each column for each factor represents estimated coefficients.
The dependent variable is the out-migration rate of state i in whole year t. It is multiplied by 100% to make the coefficient larger and easier to interpret.
The coefficients of the wages and housing prices is multiplied by 10000, and the growth unit is 10000.
The coefficient of the minimum and maximum tax rate is multiplied by 100% to make the coefficient larger.
The coefficient of the crime rate is adjusted to the unit change of crime rate is 100 per 100,000 residents.

The unit of income tax rate and state unemployment rate is in 100%.
Standard errors are in parentheses under each coefficient.
Significance is represented by * at 5% level, ** at 1% level, and *** at 0.1% level.

Table 5
Robustness checks of the state-level linear regression model using squared terms

Independent variable	Dependent variable: out-migration rate		
	1	2	3
Median pre-tax wage	-4.67** (1.52)	-4.06** (1.29)	-4.27** (1.36)
Minimum income tax rate	0.0241 (0.0302)	0.0250 (0.0339)	0.0238 (0.0295)
Maximum income tax rate	0.011 (0.023)	0.014 (0.026)	0.0102 (0.021)
Median housing price	0.283*** (0.0819)	0.217*** (0.0712)	0.247*** (0.0802)
State unemployment rate	-0.295* (0.149)	-0.289* (0.145)	-0.278* (0.141)
Out-migration rate in the previous year	_	_	0.092*** (0.023)
Squared median pre-tax wage	_	-0.831* (0.37)	_
Squared median housing prices	_	0.206* (0.0935)	_
Squared State unemployment rate	_	-0.141 (0.103)	_
Proportion of residents with more than 12 years education	0.0732** (0.0259)	0.0728** (0.0252)	0.0634** (0.0229)
Proportion of residents aged 16-29	0.0347 (0.0370)	0.0339 (0.0361)	0.0381 (0.0368)
Proportion of residents aged 30-60	-0.1469*** (0.0367)	-0.1609*** (0.0375)	-0.1562*** (0.0371)
Proportion of residents aged 61 or older	0.0312 (0.0408)	0.0307 (0.0401)	0.035 (0.0416)
Proportion of residents whose race is white	-0.034 (0.106)	-0.029 (0.108)	-0.031 (0.0112)
State crime rate	0.038* (0.0162)	0.0407* (0.0171)	0.0389* (0.0166)
Interaction term included?	Yes	Yes	Yes
State fixed effects included?	Yes	Yes	Yes
Year fixed effects included?	Yes	Yes	Yes
R-squared	0.9134	0.9196	0.9282

Note:

The numerical value in each column for each factor represents estimated coefficients.
Standard errors are in parentheses under each coefficient.
Significance is represented by * at 5% level, ** at 1% level, and *** at 0.1% level.
The coefficients of median wage and housing price are multiplied by 10000 and the growth unit is 10000.

Table 6
Comparison on the effects of factors between 2008-2013 and 2014-2018

Independent variable	Dependent variable: out-migration rate	
	1	2
Median pre-tax wage	-4.91**	-4.28**
	(1.71)	(1.39)
Minimum income tax rate	0.0239	0.0244
	(0.0291)	(0.0303)
Maximum income tax rate	0.0108	0.0112
	(0.022)	(0.023)
Median housing price	0.324***	0.221
	(0.0981)	(0.116)
State unemployment rate	-0.336*	-0.257*
	(0.164)	(0.120)
Proportion of residents with more than 12 years education	0.0735**	0.0730**
	(0.0261)	(0.0256)
Proportion of residents aged 16-29	0.0346	0.0349
	(0.0369)	(0.0372)
Proportion of residents aged 30-60	-0.1483***	-0.1461***
	(0.0375)	(0.0359)
Proportion of residents aged 61 or older	0.0308	0.0318
	(0.0397)	(0.0413)
Proportion of residents whose race is white	-0.036	-0.031
	(0.111)	(0.102)
State crime rate	0.0352*	0.0431*
	(0.0146)	(0.0178)
Interaction term included?	Yes	Yes
State fixed effects included?	Yes	Yes
Year fixed effects included?	Yes	Yes
R-squared	0.9157	0.9121

Note:
The first column is the impacts from 2008-2013, and the second column represents 2014-2018.
The numerical value in each column for each factor represents estimated coefficients.
Standard errors are in parentheses under each coefficient.
Significance is represented by * at 5% level, ** at 1% level, and *** at 0.1% level.
The coefficients of median wage and housing price are multiplied by 10000 and the growth unit is 10000.

Table 7
Definition of variables and data sources in the second logistic regression model

Variable	Definition	Data Source
WAGE	pre-tax annual wage	CPS
TAX	The marginal income tax rate for an individual or for a couple filing a joint tax return.	CPS
EMPLOY	If the individual is employed (Yes=1, No=0)	CPS
RENT	If the individual rents the current living housing unit (Yes=1, No=0)	CPS
EDUC	The education attainment of individual (1=Bachelor or up, High school or less=0)	CPS
AGE	The indication of the age of respondent (0=aged 31-60, 1=aged 60 or older, 2 = aged 16-30)	CPS
RACE	Binary variable indicating the race of individual (White only=1, Other =0)	CPS
MARITAL	If individual is married (Yes=1, No=0)	CPS
MOVE	If the individual move more than once in five years (Yes=1, No=0)	CPS
MEMBER	The number of children in the household	CPS
BIRTH	If the birthplace of individual is in the U.S (Yes=1, No=0)	CPS

Table 8
Sample Statistics for all the variables from 2008 to 2018 in the individual-level logistic regression model

Variables	Mean	SD	Range
Response variable			
Individual Migration probability	2.503%	2.097%	0 to 1
Economic factors			
Pre-tax wage	45511.55	49956.72	1 to 1699999
Marginal tax rate	7.7842%	6.787%	0 to 40%
Employment status	93.58%	12.05%	0 to 1
Housing ownership	31.682%	36.52%	0 to 1
Non-economic factors			
If individual is white only(race)	68.12%	27.13%	0 to 1

Marital status	39.46%	30.12%	0 to 1
Education attainment	20.26%	12.96%	0 to 1
The number of children in the household	1.92	1.88	0 to 9
If the individual born in the U.S	85.72%	12.31%	0 to 1
The age of individual	0.956	0.862	0 to 2

Note:
Calculation is based on the March CPS data.
For the dummy variables, the mean value represents the proportion of individuals with level of dummy variables equals to 1. Dummy variables have range 0 to 1.
Age variables has three level 0, 1, 2, and the mean age represents the estimated age level of whole samples by taking average on these three levels.

Table 9
Regression results between the individual-level interstate migration probability and each determinant

Independent variables	Dependent variable: migration probability	
	1	2
Annual pre-tax wage	-0.0683** （0.00216）	-0.0711*** (0.0219)
Marginal Income Tax rate	0.229 (1.29)	0.235 (1.32)
If individual is employed	-0.127*** (0.0308)	-0.122*** (0.0304)
If individual rents the house or owns the house	0.195** (0.0631)	0.199** (0.0639)
The education attainment	0.183*** (0.0392)	0.1806*** (0.0388)
Race of the individual	0.285 (0.461)	0.269 (0.453)
If the individual is married	-0.0938** (0.031)	-0.0972** (0.0318)
The age of individual (16-30 vs 61 or older)	0.122* (0.056)	0.131* (0.059)
The age of individual (31-60 vs 61 or older)	0.106* (0.049)	0.113* (0.052)

The number of children in the households	-0.089* (0.038)	-0.095* (0.041)
If the individual was born in the U.S	0.061 (0.037)	0.046 (0.035)
If individual move more than once in five years	0.0953** (0.046)	0.0929** (0.043)
State fixed effects included?	No	Yes
Year fixed effects included?	No	Yes
Residual Deviance	653.3	639.8

Note:
The numerical value in each column for each factor represents estimated coefficients using logistic regression model with OLS assumption.
The dependent variable is the migration probability of individual i in year t. It is multiplied by 100% to make the coefficient larger and easier to interpret.
The coefficient of the wage is multiplied by 10000, and the growth unit of wage is 10000.
Standard errors are in parentheses under each coefficient.
Significance is represented by * at 5% level, ** at 1% level, and *** at 0.1% level.

Figure 1

Density plot of the predicted out-migration rate with mean wage and housing price and median wage and housing price versus actual value of out-migration rate

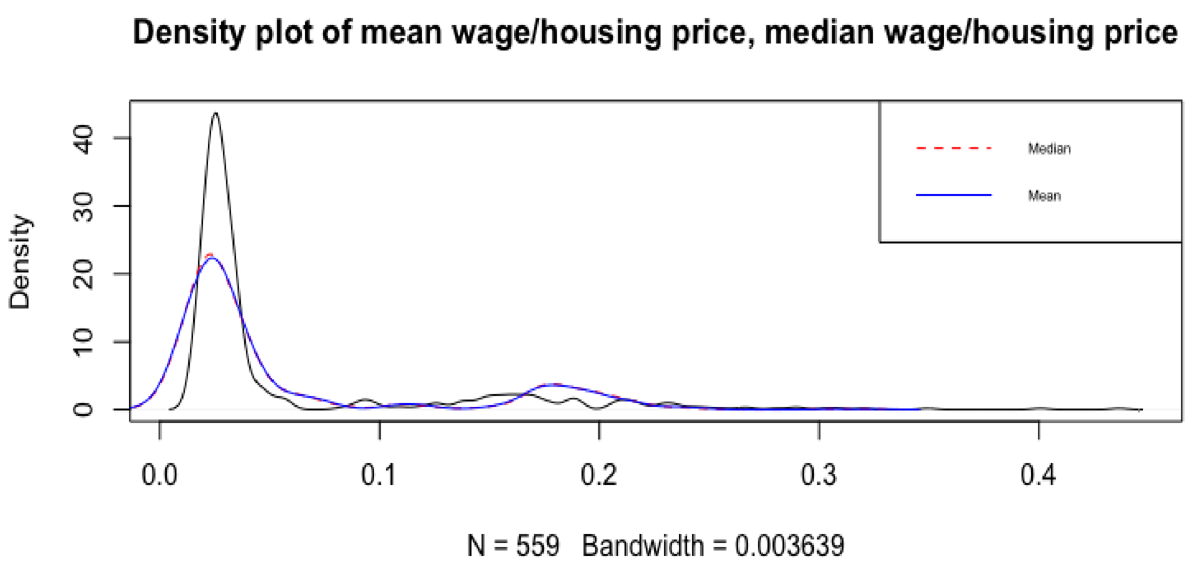

Fig1. Source: Calculation on the IPUMS-USA data thorough 2008 to 2018. N is the total number of observation data points. The predicted out-migration rate is derived using the linear regression model in Equation (4).

Figure 2

The interaction plot between the median housing price and crime rates

Fig2. Source: Calculation on the IPUMS-USA data and Uniform Crime Reporting thorough 2008 to 2018. The color of each regression line represents different level of crime rates.

Figure 3

The comparison of the actual migration trend and the predicted migration trend from the model, 2008-2018

Fig3. Source: Calculation on the migration data from United States Census of Bureau through 2008 to 2018.

Reference

Bricker, J., & Bucks, B. (2016). Negative home equity, economic insecurity, and household mobility over the Great Recession. *Journal of Urban Economics*, *91*, 1-12.

Castelli, F. (2018). Drivers of migration: why do people move. *Journal of Travel Medicine*, Volume 25, Issue 1, 2018, tay040.

Clark, W. A. (1981). Recent research on migration and mobility: a review and interpretation. *Progress in planning*, *18*, 1-56

Cohen, R., Lai, A., & Steindel, C. (2011). The effects of marginal tax rates on interstate migration in the US. *New Jersey Department of the Treasury*, *10*.

Fischer, P. A., Martin, R., & Straubhaar, T. (1997). Should I Stay or Should I Go? S. 49-90 in Tomas Hammar, Grete Brochmann, Kristof Tamas und Thomas Faist (Hg.). *International Migration, Immobility and Development. Oxford: Berg*.

Gius, M. (2011). The effect of income taxes on interstate migration: an analysis by age and race. *The Annals of Regional Science*, *46*(1), 205-218.

Haug, S. (2008). Migration networks and migration decision-making. *Journal of Ethnic and Migration Studies*, *34*(4), 585-605.

Hernández-Murillo, R., Ott, L. S., Owyang, M. T., & Whalen, D. (2011). Patterns of interstate migration in the United States from the survey of income and program participation. *Federal Reserve Bank of St. Louis Review*, *93*(May/June 2011).

Hendershott, P. H., Lee, J. M., & Shilling, J. D. (2015). The 2005–2011 Housing Boom and Bust: Impacts on Housing Turn over and Implications for the Recovery. *Journal of Real Estate Research*, *37*(4), 471-498.

Kaplan, G., & Schulhofer-Wohl, S. (2017). Understanding the long-run decline in interstate migration. *International Economic Review*, *58*(1), 57-94.

Karahan, F., & Rhee, S. (2014). Population aging, migration spillovers, and the decline in interstate migration. *FRB of New York Staff Report*, (699).

Kennan, J., & Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, *79*(1), 211-251.

Martin, P. (2009). Recession and migration: A new era for labor migration?. *International Migration Review*, *43*(3), 671-691.

Molloy, R., Smith, C. L., & Wozniak, A. (2011). Internal migration in the United States. *Journal of Economic perspectives*, *25*(3), 173-96

Molloy, R., Smith, C. L., & Wozniak, A. K. (2014). *Declining migration within the US: The role of the labor market* (No. w20065). National Bureau of Economic Research.

Mulholland, S. E., & Young, A. T. (2016). Occupational licensing and interstate migration. *Cato J.*, *36*, 17.

Peng, C. W., & Tsai, I. C. (2019). The long-and short-run influences of housing prices on migration. *Cities*, *93*, 253-262.

Smith, G., Duda, S., Fassett, T. (2014). The Impacts of Lock-in Effects on Housing Turnover and Implications for a Housing Recovery. *Institute of Housing studies.*

Sasser, A. C. (2010). Voting with their feet: Relative economic conditions and state migration patterns. *Regional Science and Urban Economics*, *40*(2-3), 122-135.

Withers, S. D., Clark, W. A., & Ruiz, T. (2008). Demographic variation in housing cost adjustments with US family migration. *Population, Space and Place*, *14*(4), 305-325.

Wozniak, A. (2010). Are college graduates more responsive to distant labor market opportunities. *Journal of Human Resources*, *45*(4), 944-970.