Housing cost, location choice, and settlement intention among migrants in China *

Wanyang Hu [1†]

[1] *School of Public Affairs, University of California Los Angeles, USA*

Abstract

In many countries, the most important source of urban population growth is domestic migration. Migrant workers benefit from migrating into cities and contribute to urban economic development. Many research efforts have been devoted to understanding the drivers of migrants' inter-city location choice, with particular focuses on the roles of jobs and amenities. By contrast, the role of housing cost has received much less attention. As housing price rises rapidly, and housing affordability becomes a binding constraint in migrants' location choice, many migrants in China have difficulty to stay in large and productive cities. Existing research on housing cost and internal migration in China is limited by the lack of micro data and models subject to endogenous home price. This research models the discrete inter-city location choice by individual migrants, using instrumental variables for city-level home price. I find high home price significantly reduces migrants' propensities of migration and settlement. A 10% increase in a city's home price reduces the city's odds of attracting migrants by 17.6% and decreases the odds of migrants' settlement in that city by 8%. Moreover, the effects of home price are stronger than that of expected wage in migrants' location and settlement decisions. My findings suggest that housing cost has become a major barrier to attracting and retaining migrant labor by cities in China.

Key words: migrant, inter-city location choice, home price, China

September 2018

*I thank the helpful suggestions and comments from Rui Wang and participants of the Symposium on "The role of housing in China's urban transformation", held at the City University of Hong Kong and Lingnan University in June 2018. I would also like to thank the UCLA Ziman Center for Real Estate's Howard and Irene Levine Program in Housing and Social Responsibility for generous funding.

†Corresponding author, wanyang@g.ucla.edu

1 Introduction

Population growth has been regarded as the yardstick to measure urban success (Glaeser, 2008), and much of urban population changes occur through internal migration across cities and regions (Mundial, 2009). Migrants supply labor to urban economy and contribute to urban development through their skills and attenuated agglomeration externalities. Given the importance of labor and population dynamics in urban development, there has been sustained research interests in understanding the relocation of labor across cities and regions. Jobs, amenities, and housing cost are three factors in urban economic theory that influence labor's inter-city location choice, and they are used in urban model to explain the equilibrium distribution of population across cities (Glaeser, 2008). [1] Many current scholarly debates in regional science have revolved around the relative importance of jobs and amenities in determining inter-city migration and urban population dynamics. [2] However, the role of housing cost in inter-city migration has received much less attention (Plantinga, Détang-Dessendre, Hunt, & Piguet, 2013). This is partially due to the fact that it is usually difficult to directly examine the effect of housing cost in migration pattern in urban model based on a spatial equilibrium framework, since wage, population and housing cost are simultaneously determined within the system (Glaeser, 2008; Roback, 1982).

By contrast, a large proportion of research has focused on the relation between housing market conditions and intra-city location choice (W. A. Clark & Onaka, 1985; W. A. Clark & Van Lierop, 1987; Quigley, 1976, 1978). Focusing on a single housing market, or between sub-markets of a metropolitan area, these studies usually evaluate residential location choice within a metropolitan area in conjunction with the selections of employment locations, housing or neighborhood types, travel mode or auto-ownership, etc. (Lerman, 1976; Quigley, 1978; Williams, 1979) Another rich line of research examines the impact of migration on local housing market, instead of vice versa (Akbari & Aydede, 2012; Saiz, 2007; Saiz & Wachter, 2011).

It has been argued that housing cost plays a less important role in long-distance migration compared with jobs and amenities (W. A. Clark & Van Lierop, 1987; Gleave & Cordey-Hayes, 1977). And change in urban population is often attributed to locational shifts in employment opportunities, consumption amenities, the easiness to enjoy natural amenities, or the changing preference accompanied with population aging and rising income (Cebula, 1979; DaVanzo et al., 1980; Greenwood, 2014). Nevertheless, such belief is problematic if we examine this issue through the lens of individual choice behavior. Spatial relocation of population is a disequilibrium process and migrants select the location where one's utility is maximized. Housing cost is a major indicator of local price, and is thus an important determinant of one's expected disposable income

[1] Housing price is usually endogenously determined in urban model by urban population and urban amenity endowment; it is usually regarded as an indication of the price of local non-tradable goods.

[2] Amenities include natural amenities such as climate and air quality (Bayer, Keohane, & Timmins, 2009; Cheshire & Magrini, 2006; Hunt & Mueller, 2004; Mueser & Graves, 1995; Rappaport, 2007), as well as man-made consumption amenities (D. E. Clark & Hunter, 1992; Glaeser, Kolko, & Saiz, 2001; Gottlieb & Joseph, 2006; Lewis, Hunt, & Plantinga, 2002). Expected job opportunities and wage are also important factors in long-distance migration(Dahl & Sorenson, 2010; Davies, Greenwood, & Li, 2001; Greenwood, Hunt, Rickman, & Treyz, 1991; Hunt & Mueller, 2004; Kennan & Walker, 2011; Mueser & Graves, 1995).

in potential destination. Thereby, it ultimately enters one's utility function and affects individual labor's location choice. This is particularly true when housing price is soaring. Research shows that housing cost has been playing a rising role in determining the inter-state migration in the United States since late 1970s (Sasser, 2010). In China, high housing price is becoming a pressing issue in big cities that are the destinations of most urban migrants and the housing affordability crisis has been acting as a binding constraint in migrants' location choice (Chinadaily, 2017). More and more young people choose to leave the top-tier cities due to their inability to afford a home in host cities (Chinadaily, 2017; Freeman, 2017). However, empirically it is less known to what extent has rising housing cost impeded inter-city migration, and whether it may further affect the labor supply and agglomeration economies in cities. To the author's knowledge, few study has quantified the effect of housing cost on migrants' inter-city location choice and settlement in China. To fill the gap, this research uses 2014 Migrant Population Survey data in China and constructs a conditional discrete choice model to test the effect of housing cost in migrants' decisions of destination cities and settlement in China. I use predicted urban land development quota and the share of urban land area with slope above 15 degrees as instrumental variables for a city's home price level. My results show housing cost imposes significant negative effects on domestic migration and settlement, and the effects of housing cost are even stronger than the effects of expected wage. This suggests that high housing cost has become a major barrier to urban population growth in China. The implicit deprivation of labor supply by high housing cost in Chinese cities may adversely affect long-term urban prosperity.

The following section reviews the literature on housing cost and inter-city migration. Section 3 constructs a theoretical model based on spatial equilibrium and the random utility theory. Section 4 introduces data and identification strategy. Second 5 presents the empirical results. The last section concludes the paper.

2 Literature review: housing cost and migration

Despite that economic theory predicts high housing cost impedes migration, existing empirical studies on the effect of housing cost in inter-city migration patterns have found quite mixed evidence, including positive, negative, weak or no impact (Berger & Blomquist, 1992; Bishop, 2008; W. A. Clark & Van Lierop, 1987; d'Albis, Boubtane, & Coulibaly, 2017; Gottlieb & Joseph, 2006; Hunt & Mueller, 2004; Jeanty, Partridge, & Irwin, 2010; Lux & Sunega, 2012; Plantinga et al., 2013; Rabe & Taylor, 2012; Sasser, 2010). The inconclusive findings are attributable to two sources of endogeneities. First, there likely exists unobserved urban amenities that correlate with home price and upwardly bias its estimates. Second, any reduced-form testing on the impact of home price on aggregate migration data may suffer from simultaneity bias, since wages, population and housing price are simultaneously determined within the spatial equilibrium framework and housing price itself is a result of migration flow (Glaeser, 2008; Roback, 1982). This leads to biased and inconsistent estimates. Some studies have used exogenous land supply constraints, such as geographic constraints or land-use regulations, as instrument variables for home price to eliminate the endogeneity bias. The rationale is that home price in places with high price elasticities of housing supply tend to be less responsive to influx of migrants and increased housing demands, thus could attract more people (Levy, Mouw, & Perez, 2017; Zabel, 2012). By analyzing aggregate urban migration flow (net migration inflow or population growth) from a macro perspective, these studies lack a clear behavioral interpretation at individual level. There is also research using non-parametric method, such as VAR model and large panel data set to detect the causality between housing market conditions and migration (d'Albis et al., 2017; Saks, 2008; Zabel, 2012). However, VAR model is atheoretical, thus its parameters do not have economic interpretations per se.

Most extant studies on this topic rely on aggregate data of migration flow (in- and out- migration, or population growth) and measures of housing cost, such as median home price or rent index, cost of living index or rent-to-income ratios. In recent years, the availability of individual micro-data allows for the usage of discrete choice model and its extensions to examine this issue. This approach could circumvent the simultaneity between aggregate migration flow and home price in studies using aggregate data, since each individual can be treated as a price taker (Plantinga et al., 2013). Plantinga et al (2013) use nested discrete choice model to investigate the role of individual housing cost in migrants' location choice and find housing cost negatively affects inter-state migration in the United States. Specifically, Plantinga et al use predicted individual housing cost in each potential destination instead of aggregate home price. However, it seems difficult to discuss the policy implications of the estimated coefficient of individual-location-specific housing cost on location choice, because individual migration decisions are made simultaneously with employment and housing consumption decisions. The majority of other studies using discrete choice models primarily focus on intra-city location choice instead of cross-city migration. [3] In addition, most studies focus on the decisions of whether to move or where to move (Fu & Gabriel, 2012; Plantinga et al., 2013), but few have examined how housing cost affects migrants' settlement intention in destination cities. [4]

Despite the evidence found in the US, to the author's knowledge, few research has empirically tested how housing cost affects migrants' inter-city location choice and settlement in China. Understanding whether housing cost has become a major migration barrier that limits urban labor supply and urban human capital agglomeration is becoming an important policy issue in contemporary China when home price has been, and is still rapidly rising in most Chinese cities. Using discrete choice model and instrumental variables for housing cost, this research aims to fill these gaps and to identify the effect of housing cost in migrants' decisions of cities and settlement in China.

[3] These studies use joint choice model to evaluate residential location choice in conjunction with the selections of employment locations, housing or neighborhood types, travel mode, auto-ownership, etc., within a metropolitan area (Lerman, 1976; Quigley, 1976, 1978; Williams, 1979).

[4] A related line of research finds the accessibility to public housing or formal housing in destination cities attracts migrants and increases their intentions to stay in host cities (d'Albis et al., 2017; Liu, Wang, & Chen, 2017; Verdugo, 2015).

3 Theoretical model

This section constructs the location choice model. Let's assume a potential migrant's utility depends on the consumption of housing and a composite good with normalized price of one. Wage is the only income source of migrant, and migrants' consumptions are subject to the budget constraint in potential destination city i. Each migrant has M potential destination cities in the choice set. Use i to indicate cities, $i = 1, 2, 3...M$, k to indicate individuals, $k = 1, 2, 3...N$. The utility function of migrant k in city i takes the Cobb-Douglas form as follows.

$$\max_{C_{ki},H_{ki}} U_{ki} = C_{ki}^{\beta_c} H_{ki}^{\beta_h} X_i^{\beta_x} e^{M_{ki}+\xi_i} \tag{1}$$
$$s.t. C_{ki} + P_i H_{ki} = w_{ki}$$

C_{ki} and H_{ki} are the consumptions of numeraire goods and housing by individual k in city i; X_i and ξ_i are vectors of observed and unobserved urban amenities in city i; M_{ki} is the psychological migration cost of individual k when moving to city i. [5] M_{ki} is measured by two variables that whether the migrant moves from another province, and whether there exists linguistic difference between migrant's origin province and destination province; P_i is the home price in city i, and w_{ki} is the expected wage of individual k in city i. First order condition yields the optimal consumption of C_{ki} and H_{ki}:

$$C_{ki}^* = \frac{\beta_c}{\beta_c + \beta_h} w_{ki} \qquad H_{ki}^* = \frac{\beta_h}{\beta_c + \beta_h} \frac{w_{ki}}{P_i} \tag{2}$$

Inserting Eq (2) into Eq (1) and taking natural logarithm yields the indirect utility:

$$lnV_{ki} = \beta_w lnw_{ki} - \beta_h lnP_i + \beta_x lnX_i + M_{ki} + \xi_i \tag{3}$$

$lnV_{ki} = lnU_{ki} - ln(\frac{\beta_c^{\beta_c}\beta_h^{\beta_h}}{(\beta_c+\beta_h)^{(\beta_c+\beta_h)}})$ is the monotonous change of the utility U_{ki}; $\beta_w = \beta_c + \beta_h$.

In equilibrium, the indirect utility is equalized in all locations and there's no migration across cities; however, cities and regions usually take time to adjust in response to exogenous shocks and to restore equilibrium. As a result, in short-term, there exists utility differences across cities that induce inter-city migration (Jeanty et al., 2010).

Following McFadden's random utility theory, I introduce a random term to represent migrant's idiosyncratic preference for cites, which yields the following utility function.

$$\begin{aligned} U_{ki} &= V_{ki} + \epsilon_{ki} \\ &= U(w_{ki}, P_i, X_i, M_{ki}, \xi_i) + \epsilon_{ki} \\ &= \beta_w lnw_{ki} - \beta_h lnP_i + \beta_x lnX_i + M_{ki} + \xi_i + \epsilon_{ki} \end{aligned} \tag{4}$$

Let U_{ki} be the utility individual k obtains by selecting city i. Migrant k selects city i if $U_{ki} > U_{km}, \forall i \neq m$. Migrants' utility of selecting a city to move depends on a systematic utility, i.e., the

[5]The physical and monetary cost of migration is expected to dissipate in the long term, which thus does not enter the equilibrium utility function or the budget constraint.

indirect utility V_{ki}, and a random component ϵ_{ki}. One's utility is maximized in current location selected. The systematic utility V_{ki} is a function of expected income w_{ki}, home price P_i, urban amenities X_i and ξ_i, and migration cost M_{ki}. w_{ki} and M_{ki} are individual-location-specific variables that depends on the individual and the potential alternative city. ϵ_{ki} is assumed to be i.i.d.Gumbel distribution across individuals and alternative cities. The probability of individual k selecting city i is

$$\begin{aligned} P(y_{ki}=1) &= Prob(U_{ki} > U_{km}, \forall i \neq m) \\ &= \frac{exp(\beta_0 + \beta_w lnw_{ki} - \beta_h lnP_i + \beta_x lnX_i + M_{ki} + \xi_i)}{\sum_{i=1}^{M} exp(\beta_0 + \beta_w lnw_{ki} - \beta_h lnP_i + \beta_x lnX_i + M_{ki} + \xi_i)} \end{aligned} \tag{5}$$

β_i are the coefficient vectors to be estimated.

The Likelihood function of observing current distribution of migrants across cities is:

$$L = \prod_{1}^{M} \prod_{1}^{N} P(y_{ki}=1)^{D_{ki}} \tag{6}$$

$D_{ki} = 1$ if city i is selected by individual k. By maximizing the Likelihood function, I obtain estimates of the effects of housing cost in migrants' decisions of cities, i.e., $\widehat{\beta_h}$. However, since home price P_i may be correlated with unobserved urban amenities ξ_i, estimated $\widehat{\beta_h}$ may be biased and require additional identification strategy. I describe the identification strategy in details in Section 4.

In Eq (5), P_i, X_i, and M_{ki} can be obtained from the data. However, I can only observe the wage individual k earns in current city of residence, but not the expected wages in all potential destination cities in the choice set. Following Bayer et al (2009), I use a Mincerian-style wage equation to predict wage in each potential city for each individual, including cities that an individual did not select. Eq (7) is the prediction equation to be fitted for each city to estimate the expected monthly income for each migrant in each alternative city that s/he might choose. [6]

$$\widehat{wage_{ki}} = \beta_{0i} + B_i X_k + \sigma_i \tag{7}$$

$\widehat{wage_{ki}}$ is the expected wage of migrant k in city i; β_{0i} is the city-specific intercept in the fitted wage model which captures unobserved city-level shocks to individual wage; X_k includes migrants' individual characteristics that affect its expected earnings, including employment type(e.g., employee, employer, self-employment, etc.), sector(e.g., public, private, foreign firms, etc.), job type (e.g., management, technician, staff, businessman, etc.), industry (e.g., first, second, third, etc.), gender, education, age, and age squared. B_i is the vector of coefficient estimates in city i; and σ_i is the city-specific residual that follows standard normal distribution. I present the OLS prediction results in section 4.

[6] Monthly income includes salary and the discounted monetary costs of housing and meals covered by employers.

4 Data and identification strategy

4.1 Descriptive statistics and stylized facts

The main data source is the 2014 Migrant Population Survey in China. This dataset uses a stratified, multi-stage probability proportional to size (PPS) sampling strategy to sample around 160,000 urban migrants in over 300 Chinese cities in 31 provinces annually. To ensure the power of prediction of expected wage, I only keep cities with sampled migrant size above 150, which leaves 118 cities in the sample. Since the migrant data is in 2014, the location choice is made prior to 2014. Thus I use urban home price in 2013 to measure housing cost, and urban attributes in 2013 to capture urban amenities that impact migration decision. Urban home price is computed by dividing total income of housing sales by the total areas of housing sales in the city. The data of housing sales is from China Statistical Yearbook for Regional Economy 2014. As shown in the first row of Table 2, the mean home price in 2013 is around 5753 yuan per square meters and the median is 4577 yuan per square meters. Other urban attributes are calculated with data from China City Statistical Yearbook 2014.

Table 1 presents the summary statistics of individual characteristics. 62% migrants in the sample is male; similar with migrants elsewhere, Chinese urban migrants on average are at their prime-age, i.e., 35 years old; most migrants are married and have a local household size of 2-3. 61% migrants move from other provinces. The education profile shows that most Chinese urban migrants have middle or high school degree, and only 6.6% migrants hold college or above college degree. Corresponding to the education level, 83% migrants work in low-end services, small business, or manual jobs in private sector. Self-employment is also an important form of employment for urban migrants, constituting 33% migrants. Table 2 lists the relevant urban characteristics that may affect migrants' location choice. I compute several urban amenity indexes with factor analysis using data from China City Statistical Yearbook 2014, including urban economic conditions, natural and man-made amenities. [7] The average urban GRP growth rate is 8.5% and 41% urban GRP is from tertiary industry. On average, only 6% urban population has bachelor or above bachelor degree. 13.6% cities in the sample are coastal cities.

There are some stylized facts about housing conditions and housing costs among urban migrants in Chinese cities. Table 3 shows the share of migrants in different housing types. In 2014, 66.6% urban migrants rent private housing. About 17.7% migrants live in housing provided by employers, including regular rental housing, affordable/free housing and workplace. Only 11.8% migrants have purchased either private or public housing in destination cities. The distribution of migrant shares living in different housing types are similar in cities at different GRP percentiles.

I further calculate the current monthly housing cost and the share of housing expenditure among total income for urban migrants. In addition to overall median and mean values in full city sample, I also compute the corresponding statistics for cities at different percentiles of GRP. To circumvent the influences of extreme values, I mainly look at median values in the analysis. Table 4 shows

[7] Factor analysis is able to summarize and extract the latent information contained in several highly correlated raw variables, which is widely used in index construction.

Table 1: Summary statistics: individual characteristics

Variable	Definition	Mean	S.D
		N=66456	
$\widehat{wage}$	Predicted monthly income (Yuan); median=3187.42	3363.52	1869.61
sex	Gender, 1=male	0.620	0.485
age	Age, min=15, max=60	35.044	8.559
married	Married=1	0.837	0.370
hhsize	Local household size; author calculation	2.751	1.132
settle	Settlement intention in current city, 1=yes	0.80	0.40
dif_pro	1=migrate from different province	0.612	0.487
lang_dif	1= with linguistic difference [a]	0.492	0.500
edu_2	Primary school	0.117	0.322
edu_3	Middle school	0.499	0.500
edu_4	High school	0.204	0.403
edu_5	Professional school	0.100	0.301
edu_6	College	0.060	0.237
edu_7	Post graduate	0.006	0.077
voc_manage	Management	0.007	0.083
voc_tech	Technical Experts	0.091	0.288
voc_staff	Staff	0.018	0.133
voc_prbus	Small business	0.259	0.438
voc_serv	Service	0.340	0.474
voc_produc	Operation/production/transportation	0.236	0.425
voc_farm	Agriculture/forestry/husbandry/fishery	0.023	0.149
voc_others	Others vocation	0.004	0.064
indu_2	Second industry	0.282	0.450
indu_3	Tertiary industry	0.692	0.462
unit_land	Land owner	0.015	0.122
unit_private	Private sector	0.741	0.438
unit_public	Public sector	0.084	0.277
unit_foreign	Foreign firms and firms using funds from Foreign, Hong Kong and Macau	0.055	0.227
unit_other	Others sector	0.004	0.065
emly_type1	Employee	0.550	0.497
emly_type2	Employer	0.106	0.308
emly_type3	Self-employed	0.329	0.470

Source: Migrant Population Survey data in China in 2014 and author calculation
a. Author calculation, dialect information is from *Language atlas of China.* 2012. Beijing: Commercial Press.

Table 2: Summary statistics: city characteristics

Variable	Definition	mean	S.D
		N=118	
HP	Urban home price in 2013 (yuan per m^2); HP=total income of housing sales/total areas of housing sales [f]; median= 4576.72	5752.687	3519.336
grp_rate	GRP growth rate [a]	0.085	0.023
per_grp3	Percentage GRP in tertiary industry [a]	0.413	0.103
human_cap	Percentage of total urban population with bachelor or above degrees among working age (15-64) population in 2010 Population Census [b]	0.059	0.042
openness	Economic openness index measures the share of industrial values and numbers of enterprises with non-domestic funds in a city. [a]	0.244	1.147
culture	Cultural index measures urban cultural amenities, including number of theaters, music halls, cinemas, and books in libraries. [a]	0.400	1.190
medical	A composite medical index measures the quality of medical service in a city, including number of hospital, doctors, and beds in hospital. [a]	0.248	1.088
greenness	A composite greenness index measures the green space and parks in a city. [a]	0.256	0.957
trans_hub	Transport hub index measures the passengers and freight traffic in a city. [a]	0.260	1.119
transport	Transport index measures coverage of paved roads, buses, and taxis in the city. [a]	0.271	1.037
port_shore	Whether the city is a coastal city, 1=Yes; 0 otherwise. [c]	0.136	0.344
major_river	Whether the city is adjacent to major river, 1= Yes, 0 otherwise. [c]	0.466	0.501
minor_river	Whether the city is adjacent to minor river, 1= Yes, 0 otherwise. [c]	0.424	0.496
avgjantemp	Average January temperature [d]	1.641	8.799
land_resi11	Urban residential land supply in 2011 (Hectare) [e]	565.854	577.002
land_cons11	Urban construction land supply in 2011 (Hectare) [e]	2502.739	2756.870
pre_quota	Predicted urban land development quota in 2011 (Hectare) [e,g]	2585.530	2318.567
per_slope15	Share of urban land area with slope above 15 degrees	0.078	0.093

Source:

a. China City Statistical Yearbook and author calculation;

b. 2010 Population Census in China;

c. China's History in Maps, http://worldmap.harvard.edu/maps/china-history;

d. China Meteorological Data Service Center, http://data.cma.cn/en;

e. China Land and Resources Statistical Yearbook 2012;

f. China Statistical Yearbook for Regional Economy 2014;

g. Author calculation.

Table 3: Share of migrants in different housing types (%)

Housing type	Rent	Own
Private housing		
Regular private housing	66.58	11.41
Lodging	0.99	-
Informal housing	0.46	-
Public housing	0.33	0.35
Housing provided by employers		
Regular	5.17	-
Affordable/free housing	10.41	-
Workplace	2.12	-
Self-built housing	2.18	

Total observations: 145,920

Table 4: Migrants' housing cost

		All	top 1%	top 10%	25%-75%	bottom 25%
Monthly housing cost (Yuan)	Median	500	580	500	500	500
	Mean	794	988	935	691	810
Rent/Income per month	Median	0.147	0.175	0.156	0.143	0.149
	Mean	0.208	0.271	0.220	0.200	0.214

Note: a) The percentile is based on per capita Gross Regional Product (GRP per capita);
b) Rent is monthly rent per household in 2014;
c) The table only includes observations whose monthly rent is non-zero;
d) The sample excludes 150 outliers whose rent/income exceeds 1;
e) The figures are real values in the current city a migrant lives in.

that the median monthly housing costs of migrants is around 500 yuan in 2014, and it is similar across cities with different economic conditions. The median percentage of housing expenditure in income per month is about 15 %. Both the raw monthly housing cost and the ratio of housing expenditure to income indicate that housing cost has not a become huge burden for Chinese urban migrants comparing with the 30% international standard, despite the housing burden is slightly higher in more developed cities. However, it is noteworthy that monetary measures of housing cost as percentage of monthly income cannot adequately reflect the actual housing burdens of migrants, since many urban migrants choose to compromise their living conditions, such as unit size, distance to jobs, or neighborhood amenities, in order to reduce monetary housing cost, especially in big cities where home price is high (Niu & Zhao, 2017). Therefore, despite the low monetary burden we observe, it should not be regarded as an indication of optimistic housing conditions among urban migrants in China. Unfortunately, data limitation prevents us from further understanding the actual housing conditions of urban migrants behind these figures.

Table 5 lists the median monthly costs by housing types. Compared with privately-owned, privately-rental, and employer-provided housing, migrants in much fewer cities live in public hous-

Table 5: Median housing cost by housing type

Variable	Obs	mean	Median	S.D	min	max
Median monthly cost (Yuan)						
Public Housing: ownership	52	1225.48	975.00	1116.78	100	7000
Private Housing: ownership	231	1564.66	1500.00	786.96	60	6900
Public Housing: rental	69	546.30	350.00	535.89	50	3000
Private Housing: rental	264	514.88	487.50	301.88	120	3000
Employer-provided housing	256	429.65	300.00	509.63	50	4000
Self-built housing	88	671.14	500.00	837.97	30	5100

ing and self-built housing. The average costs of rental housing is about 500 yuan per month, and the monthly costs of owned housing range from 1200 to 1500 yuan. Self-built housing and employer-provided housing have similar median monthly costs with rental housing.

4.2 Identification strategy

One source that biases the estimate of the effect of housing cost in location choice, i.e., $\widehat{\beta_h}$ in Eq (4), is the correlation between home price and unobserved or omitted urban amenities ξ_i, i.e., $Cov(P_i, \xi_i) \neq 0$. To address such bias, I employ two strategies. I first use exogenous source of variations that affect home price but are not related to urban amenities to instrument for a city's home price P_i. Following Liang et al (2016), I first use predicted land development quota as instrumental variable of urban home price (Liang, Lu, & Zhang, 2016). Land use regulation has been shown to be an important factor that affects urban home price through influencing the supply elasticities of urban land and housing (Gyourko, Saiz, & Summers, 2008). In China, construction land development quota is one major determinant of urban land supply. It is determined by central and provincial government in the land use plan, which is thus exogenous to city conditions (Liang et al., 2016). However, construction land development quota data is unavailable at city level. Thus, I construct a predicted urban construction land development quota by multiplying province-level construction land development quota by the share of residential land supply of each city in the respective province, see Eq (8). q_j is the total construction land development quota in province j, $\widehat{q_{ij}}$ is the predicted land development quota in city i of province j, l_{ij} is the residential land supply in city i of province j, and I_j is the total number of cities in province j. To eliminate the reverse causation, I use two-year lagged land development quota data in 2011 to instrument for home price in 2013. The second last row in Table 2 shows that the predicted construction land development quota in the city has a mean of around 2586 hectares, which is very close to the actual urban construction land supply, 2502.74 hectares. This increases the confidence in its relevance to urban land supply.

$$\widehat{q_{ij}} = \frac{l_{ij}}{\sum_1^{I_j} l_{ij}} * q_j \tag{8}$$

However, despite being determined exogenously by central and provincial government, construction land development quota may still reflect the city's land and housing demand, which thus may correlate with urban amenities. To remedy for this, I use a second instrument variable, the share of urban land area with slope above 15 degrees, as a natural geographic constraint of urban land supply to instrument for home price, which has been argued to be a more robust instrument that is exogenous to urban conditions (Saiz, 2010). As shown in the last row of Table 2, on average, 8% of urban land area in the sample have slope above 15 degrees.

It is also noteworthy that if X_i fails to capture all urban amenities that attract or dispel migration, the estimated coefficient of $\widehat{\beta_h}$ may still be biased even if I instrument for home price. To solve this, I use the difference between average urban wage and aggregate home price as an alternative approach to measure urban amenities. The rationale is that, in long-term equilibrium when there's no migration, the indirect utility $V_{ki}(w_{ki}, P_i, X_i)$ is a constant in every location. Thus the difference between home price and wage, $(P_i - w_{ki})$, should be equivalent to the monetary value of the differentials in urban amenities X_i across cities. [8] I believe a combination of instrument variable and this alternative measurement of urban amenities could better eliminate the bias from omitted amenity variables.

4.3 Predicted wage

According to Eq (7), I predict each migrant's expected wage in each potential destination city. I summarize migrants' predicted monthly wage in the first row of Table 1. The expected wage per month has a mean of 3363.52 yuan, and a median of 3187.42 yuan, but it also has very large variations across individuals and locations. Table 6 shows the empirical results of an illustrative model of national-level wage regression. [9] As expected, wage rises with education level; low-end service vocations yield lower wages compared with other vocations; interestingly, certain manual jobs pay more wages than vocations such as staffs or small businessmen. Non-domestic firms pay the highest wage, with private sector ranking the next. All else equal, self-employed migrants have higher wage premiums than migrants who are regular employees.

5 Empirical results

5.1 Housing cost and migrants' inter-city location choice

This section presents empirical results of the location choice model. I first use two naive conditional logit models without instrumenting for home price in column (1) and (2) to obtain baseline results. Specification in column (1) only includes urban aggregate home price in 2013, predicted wage, and two migration cost measurements, dif_pro and lang_dif. Model in column (2) adds a set of

[8] According to Eq (4), urban amenities would be better captured by the difference between home price, migration cost, and wage, i.e., $(P_i + M_{ki} - w_{ki})$. However, it is very difficult to estimate the equivalent monetary value of the psychological migration cost M_{ki}, thus I did not include M_{ki} in the empirical analysis. Despite this, I believe the exclusion of the monetary value of M_{ki} will not influence the results to a great extent given the dominant roles of wage and housing cost in migration decisions.

[9] I have run the same model for each city in the sample to predict individual-city-specific wage.

Table 6: Illustrative model of national-level wage regression

Dependent variable: Ln(monthly wage (yuan))	Coefficients	t-statistics
Gender, Male=1	0.228***	(81.69)
Age	0.0399***	(38.01)
Age Square	-0.000543***	(-36.35)
Education: base group= no education		
Primary School	0.0820***	(6.30)
Middle school	0.170***	(13.41)
High School	0.243***	(18.79)
Professional School	0.337***	(25.02)
College	0.496***	(34.18)
Post Graduate	0.756***	(27.30)
Vocation: base group=unemployed/no stable jobs		
Management	0.425***	(16.59)
Technical Experts	0.327***	(30.53)
Staff	0.226***	(16.33)
Small Business	0.227***	(21.80)
Service	0.194***	(20.17)
Operation/production/transportation	0.252***	(24.94)
Agriculture/forestry/husbandry/fishery	-0.00313	(-0.11)
Others	0.183***	(7.62)
Industry: base group=agriculture		
Secondary Industry	0.00539	(0.22)
Third Industry	-0.0784**	(-3.26)
Unit type: base group=no unit		
Land owner	-0.0285	(-1.50)
Private	0.0826***	(15.16)
Public sector	0.0457***	(6.38)
Foreign/Hong Kong/Macau/Taiwan-owned	0.103***	(14.16)
Others	0.0389	(1.64)
Employment identify: base group=others		
Employee	0.0162	(1.34)
Employer	0.421***	(30.34)
Self-employed	0.156***	(12.56)
Cons	6.538***	(157.18)
N	129643	
Adjusted R-squared	0.223	

t statistics in parentheses $+p < 0.10, *p < 0.05, **p < 0.01, ***p < 0.001$

Note: To ensure the power of prediction, I only keep cities whose migrant size is above 150.

variables measuring urban economic conditions, natural and man-made amenities that potentially affect migration decisions. Results in both column (1) and (2) show positive associations between urban home price and the log odds of selecting current cities among migrants. These results are not surprising since without any identification strategy, the raw measurement of home price is endogenously determined by both observed and unobserved urban amenities, which upwardly bias the estimates of the effect of housing cost in location choice.

To address such bias, I instrument for urban home price with two instrumental variables. Column (3) and (4) show results using predicted urban land development quota in 2011 as IV for home price. Column (5) and (6) present results using the share of urban land area with slope above 15 degrees as IV. In column (3) and (5), I explicitly control for a set of urban amenity variables; in column (4) and (6), I measure urban amenities with the difference between average home price and average wage to better eliminate unobserved and omitted amenity variables that correlate with home price. Combining this measurement of urban amenities and instrument variables of home price, I believe the estimated coefficients of home price are less subject to biases.

The lower panel in Table 7 shows the first-stage results. [10] Predicted urban land development quota negatively affects urban home price, which is as expected that more urban land supply depresses urban home price. The results are statistically significant at 1% significance level in both specifications. The F-stats is far above the conventionally used threshold for first-stage regression, indicating a strong relevance of the IV with home price. As shown in the conditional logit model in the upper panel of Table 7, housing cost reduces migrants' probability of selecting cities. Estimates in column (4) suggest that, *ceteris paribus*, a 10% increase in urban home price level reduces the odds of migrating to current city by 43.55%. Moreover, the effect of housing cost is even much larger than the effect of expected wage in migrants' location choice, that a 10% increase in expected wage can only increase the odds of migration by 14% while holding other factors constant.

Column (5) and (6) show results using the share of urban land area with slope above 15 degrees as IV for urban home price. More slopes in urban area reduce the available land for housing construction, especially the construction of high-density housing, which thus raise home price in the city. The first-stage results indicate significantly positive effects of the share of urban land area with slope above 15 degrees on urban home price. The F-stats also suggests a strong prediction power of slope as IV for urban home price. Similar with results in column (3) and (4), I also find negative effect of housing cost on migrants' location choice using slope as IV. As shown in column (6), a 10% increase in home price decreases the odds of migration by 17.6%, whereas expected wage has a much weaker effect on migrants' location choice, that a 10% increase in expected wage increases the odds of selecting current city by 9%.

Other migration drivers also yield interesting results. Expected income has consistently positive effect on location choice. Linguistic difference impedes migration as expected, but distance encourages migration instead of discouraging it. This suggests that physical distance is no longer a

[10] Since the main model is a nonlinear conditional logit model with binary dependent variable, I use control function method to estimate the IV regressions instead of traditional 2SLS to obtain consistent estimates, as recommended by Wooldridge (J. Wooldridge, 2007; J. M. Wooldridge, 2010).

major migration barrier, but cultural/language difference remains to limit the free flow of migration across cities. The share of tertiary industry in total urban GRP, urban human capital concentration, economic openness, cultural amenities, quality of medical services and urban transportation facilities all increase the probability of migrating to cities. GRP growth rate and urban green space do not attract migrants to cities. Coastal cities and cities with major rivers significantly increase migrants' propensities of relocating than other cities with similar attributes. The measurement of urban amenities using the difference between home price and average wage has expected positive effects on migration.

In addition to the above estimations using full sample of migrants, I also run the same models with a restricted sample that only includes migrants who relocated to current city within the past three years. The goal is to ensure the estimated wage using data in 2014 could resemble the labor market conditions at the time when migrants made location choice as much as possible. To save space, I present the empirical results using restricted sample in Table 9 in the Appendix. The estimates are very similar with what I found in the full sample estimation, indicating the robustness of the results. [11]

Table 7: Conditional logit model of location choice: aggregate home price

	(1)	(2)	(3)	(4)	(5)	(6)
DEV=Choice	Baseline model		IV1		IV2	
ln(HP)	1.763***	1.299***	-13.08***	-6.004***	-0.367+	-2.033***
	(218.22)	(58.85)	(-13.43)	(-63.70)	(-1.87)	(-13.87)
ln($\widehat{wage}$)	0.203***	0.235***	1.558***	1.388***	0.390***	0.905***
	(11.87)	(10.96)	(16.89)	(52.34)	(13.70)	(32.11)
dif_pro	0.220***	0.0716***	1.029***	-0.0687***	0.180***	0.133***
	(26.87)	(7.31)	(15.69)	(-7.09)	(11.11)	(13.44)
lang_dif	-2.686***	-2.699***	-2.254***	-2.285***	-2.649***	-2.525***
	(-307.75)	(-277.51)	(-71.30)	(-221.33)	(-230.87)	(-234.74)
port_shore		0.128***	3.689***		0.536***	
		(8.28)	(15.25)		(10.61)	
major_river		0.163***	0.716***		0.206***	
		(11.55)	(17.86)		(13.66)	
minor_river		0.0726***	-0.0680***		0.0333*	
		(5.04)	(-3.96)		(2.22)	
avgjantemp		-0.0333***	0.174***		-0.00940**	
		(-39.33)	(12.37)		(-3.19)	
grp_rate		-6.357***	-29.10***		-8.905***	
		(-22.05)	(-18.56)		(-20.29)	
per_grp3		1.184***	21.53***		3.600***	
		(15.31)	(15.61)		(12.26)	
openness		0.153***	1.086***		0.260***	

[11] I also try to impose further restrictions on migrants' residence time in current cities. However, as I further increase the threshold of migrants' residence time in current cities, the number of cities included in the final sample decreases dramatically. This raises concerns to the representativeness of the city sample. Therefore, I only set the restriction of residence time to three years as a compromise.

		(25.48)	(17.11)		(18.82)	
culture		0.0150**	1.159***		0.149***	
		(2.71)	(14.92)		(9.04)	
medical		0.0408***	0.707***		0.119***	
		(8.22)	(15.56)		(11.42)	
greenness		-0.140***	-0.656***		-0.199***	
		(-19.00)	(-18.37)		(-19.66)	
transport_hub		-0.00398	-0.0168***		-0.00665	
		(-0.80)	(-3.34)		(-1.35)	
transport		0.160***	0.169***		0.159***	
		(31.08)	(32.93)		(30.55)	
human_cap		2.343***	31.29***		5.538***	
		(13.67)	(15.88)		(13.33)	
ln(HP-Wage)				1.954***		0.942***
				(79.64)		(25.36)
First-stage residuals			14.36***	8.180***	1.677***	3.889***
			(14.76)	(84.09)	(8.57)	(26.19)
First-Stage			DEV=ln(HP13)			
ln(pre_quota)			-0.00777***	-0.0603***		
			(-60.18)	(-540.42)		
per_slope15					0.339***	0.350***
					(332.64)	(354.93)
Adjusted-R-squared			0.734	0.726	0.736	0.721
F-Stats			1249414	3097749.3	1253229.6	2993850.3
N	8676500	6504816	6504816	4932048	6437240	4867302
No.parameters	4	17	18	6	18	6
Pseudo-R-squared	0.199	0.221	0.222	0.198	0.221	0.184
Log-Likelyhood	-254090.9	-207559	-207449.5	-201361.9	-206788.2	-203985.6
Initial log likelihood	-317135.1	-266562.3	-266562.3	-250936.8	-265602.3	-249849.7
AIC	508189.8	415151.9	414935	402735.8	413612.5	407983.1
BIC	508245.7	415384.6	415181.4	402816.3	413858.7	408063.5

t statistics in parentheses; + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.2 Housing cost and migrants' settlement intention

As an extension of the analysis on the effect of housing cost on migrants' choices of cities, I use similar model specifications and identification strategy to further examine how urban housing cost affects migrants' settlement intention. Eq (9) is a logistic choice model of settlement intention. Y_s is migrant's settlement intention, $Y_s = 1$ suggests an intention to settle in destination city, 0 otherwise. As shown in Table 1, 80% migrants in the sample would like to settle in destination cities in 2014. P_{Y_s} is the probability of choosing to settle. HP_i is the aggregate urban home price in city i in 2013; w_{ki} is the wage individual k earns in city i; M_{ki} is the migration cost of individual k in city i; X_i is a vector of urban amenities that is the same as those used in the location choice model; X'_k is the vector of individual characteristics that may affect one's settlement decisions, including gender, age, age squared, education, marital status, and residence time in current city; ϵ is a random error term that follows standard logistic distribution. γs are the respective coefficient vectors to be estimated. The coefficient of interest is γ_h, the effect of housing cost on settlement intention. It is also noteworthy that i is the current city that a migrant has already migrated to, rather than all potential destination cities.

$$Logit(P_{Y_s}) = \gamma_0 + \gamma_h lnHP_i + \gamma_w lnw_{ki} + \gamma_M M_{ki} + \gamma_x X_i + \gamma_{x'} X'_k + \epsilon \tag{9}$$

The empirical results of Eq (9) is shown in Table 8. It follows the same organizations with Table 7. To save space, I omit the results of urban amenities and individual characteristics. Moreover, I focus on results in column (4) and (6) that use both instrument variables and amenity measures with the difference between home price and wage, which could better alleviate the endogeneity bias and yield reliable estimates. [12] First-stage results suggest that, after eliminating the omitted amenity variables through differencing home price and wage, both predicted land development quota and the share of urban land area with slope above 15 degrees have the expected relations with urban home price. More land development quota results in lower home price, whereas more urban land areas with steep slope raise home price. Regarding to the settlement intention, as shown in column (4) and (6) of Table 8, a 10% increase in home price reduces migrants' odds of settlement in destination city by around 8%. And again, such negative effect is even stronger than the positive effect of expected wage, that a 10% increase in expected wage can only lead to a 1.6% increase in the odds of settlement. Both having an origin of different province and linguistic difference reduce migrants' probability of settlement. As expected, amenity measurement using $ln(Hp - Wage)$ positively impacts migrants' settlement decisions.

6 Conclusions

This research examines the effects of housing cost on migrants' decisions of destination cities and settlement. I predict expected income for each migrant in each potential destination city. I instrument for urban home price to measure housing cost with predicted urban land development quota and the share of urban land area with slope above 15 degrees. Using discrete choice model, I find significant negative effects of housing cost in migrants' decisions of cities and settlement. The results are robust using different instrument variables, measures of urban amenities, and samples of migrants. The conservative estimates show that, a 10% increase in urban home price level leads to a 17.6% reduction in the odds of migration, and a 8% decline in the odds of settlement in destination city. More importantly, I find the effects of housing cost in migrants' location choice and settlement are even stronger than the effects of wage. My findings suggest that housing cost is not only an important inhibiting factor in long-distance migration and settlement decision, but may exert stronger influences than expected income and urban amenities, at least in contemporary China. It shows that high home price in Chinese cities has already become a major barrier that dispels migrants, despite of the high wages in these cities. Cities with high home price may not only lose its attraction to migrants, but also may experience difficulties to retain migrant labor. With the rapid population aging and increasing scarcity of labor supply, strategies to address housing affordability issue so that to attract and to retain more

[12] Using predicted land development quota as instrument variable may still correlate with omitted amenities if I only include a number of urban amenity measures. Unlike slope, the predicted land development quota is likely to capture the land demand in a city to some extent, which highly relates to migrants' characteristics as well as urban amenities. When solely controlling for a set of amenity variables, the presence of omitted urban amenities in the residuals undermines the validity of predicted land development quota as IV since they may potentially correlate with predicted land development quota. However, such bias will be eliminated when I use the difference between home price and wage to capture all observed and unobserved urban amenities. As a result, I believe results in column (4) and (6) that combine instrument variable and the alternative measure of urban amenities could produce more unbiased and consistent estimates.

Table 8: Housing cost and migrants' settlement intention

	(1)	(2)	(3)	(4)	(5)	(6)
DEV=Settlement	Baseline logit		IV1		IV2	
ln(HP)	0.322***	0.137**	-8.796*	-0.908*	-7.373***	-0.846*
	(17.58)	(3.05)	(-2.41)	(-2.52)	(-3.87)	(-2.25)
ln($\widehat{wage}$)	0.104***	0.0880***	0.337**	0.167***	0.299***	0.166***
	(6.04)	(4.69)	(3.26)	(7.41)	(5.29)	(7.22)
dif_pro	-0.291***	-0.253***	-0.400***	-0.176***	-0.379***	-0.185***
	(-9.43)	(-7.57)	(-5.86)	(-4.01)	(-8.34)	(-4.09)
lang_dif	-0.250***	-0.257***	0.553+	-0.230***	0.426*	-0.234***
	(-8.49)	(-7.93)	(1.66)	(-6.93)	(2.42)	(-7.06)
ln(HP-Wage)				0.432***		0.411**
				(3.57)		(3.27)
First-stage residuals			8.933*	1.202***	7.513***	1.135**
			(2.45)	(3.32)	(3.94)	(3.01)
Individual characteristics	Y	Y	Y	Y	Y	Y
Amenities	Y	Y	Y	Y	Y	Y
First-stage	DEV=ln(HP13)					
ln(pre_quota)			0.00450***	-0.0329***		
			(3.44)	(-34.70)		
per_slope15					0.0712***	0.279***
					(6.52)	(34.83)
Adjusted-R-squared			0.82	0.821	0.819	0.821
F-Stats			12314.4	21247.8	12259.1	21203.6
N	87128	75909	75909	74104	75773	73968
No.parameters	17	30	31	19	31	19
Pseudo-R-squared	0.119	0.12	0.12	0.118	0.121	0.118
Log-Likelyhood	-38445	-33930.2	-33927.4	-32758.5	-33852.7	-32689.1
Initial log likelihood	-43657.8	-38572.5	-38572.5	-37148.5	-38496.6	-37072.5
AIC	76922.1	67918.5	67914.7	65553	67765.4	65414.3
BIC	77072.1	68186.3	68191.9	65718.8	68042.4	65580.1

t statistics in parentheses; + $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

labor is important for city governments to put into policy agenda to ensure long-term urban prosperity.

One limitation of this research is the truncated sample of migrants due to data limitation. Most migrant surveys in China define migrants, or mobile population, based on one's local Hukou status, instead of fully tracking one's migration history. This definition inevitably excludes migrants who move across cities but have already obtained local Hukou in current destination city. Such sample truncation especially excludes many high-skilled migrants. Thereby, due to this data limitation, the sample of migrants is unlikely to be fully representative of high-skilled labor and is inadequate to examine the location choice behavior of high-skilled labor, who may exhibit different sensitivities to housing cost in their location choice and settlement. High-skilled migrant labor is a very important segment of migrants that make crucial contribution to urban development. They also tend to be the most mobile population segment at the same time. Future study on the location choice of high-skilled migrants is necessary with the advancement of data availability.

7 Appendix

Table 9: Conditional logit model of location choice: sample with residence time $\leqslant$ 3 years

	(1)	(2)	(3)	(4)	(5)	(6)
DEV=Choice	Baseline model		IV1		IV2	
ln(HP)	1.751***	1.582***	-19.92***	-5.349***	-0.202	-1.760***
	(139.92)	(47.31)	(-13.19)	(-38.34)	(-0.70)	(-8.02)
ln($\widehat{wage}$)	0.327***	0.324***	2.422***	1.361***	0.502***	0.879***
	(11.83)	(9.66)	(16.00)	(33.12)	(11.25)	(19.41)
dif_pro	0.0748***	-0.0636***	1.367***	-0.180***	0.0525*	-0.0250+
	(6.03)	(-4.32)	(13.46)	(-12.61)	(2.17)	(-1.71)
lang_dif	-2.779***	-2.768***	-1.991***	-2.436***	-2.703***	-2.660***
	(-206.59)	(-187.54)	(-35.26)	(-155.50)	(-149.55)	(-163.07)
port_shore		-0.0194	5.296***		0.415***	
		(-0.81)	(14.15)		(5.57)	
major_river		0.180***	1.010***		0.222***	
		(8.28)	(16.20)		(9.66)	
minor_river		0.107***	-0.0939***		0.0623**	
		(4.91)	(-3.65)		(2.76)	
avgjantemp		-0.0274***	0.284***		-0.00191	
		(-21.23)	(12.96)		(-0.44)	
grp_rate		-4.688***	-38.40***		-7.183***	
		(-10.41)	(-15.93)		(-10.87)	
per_grp3		1.003***	31.39***		3.565***	
		(8.64)	(14.70)		(8.31)	
openness		0.0740***	1.469***		0.188***	
		(8.07)	(14.93)		(9.24)	
culture		-0.0621***	1.642***		0.0814***	
		(-7.16)	(13.69)		(3.35)	
medical		0.0472***	1.041***		0.131***	
		(6.46)	(14.81)		(8.57)	
greenness		-0.0917***	-0.863***		-0.154***	
		(-8.01)	(-15.60)		(-10.12)	
transport_hub		-0.0132+	-0.0304***		-0.0151*	

		(-1.85)	(-4.18)		(-2.12)	
transport		0.204***	0.210***		0.200***	
		(26.73)	(27.90)		(25.90)	
human_cap		2.097***	45.45***		5.500***	
		(8.05)	(14.86)		(8.93)	
ln(HP-Wage)				1.812***		0.901***
				(49.77)		(16.23)
First-stage residulas			21.48***	7.075***	1.790***	3.180***
			(14.24)	(49.33)	(6.24)	(14.34)
Frist-stage	DEV=ln(HP13)					
ln(pre_quota)			-0.00768***	-0.0603***		
			(-38.94)	(-355.01)		
per_slope15					0.340***	0.348***
					(218.66)	(231.88)
Adjusted-R-squared			0.734	0.727	0.737	0.723
F-stats			536345.3	1336305.6	538086.4	1290910.1
N	3701214	2850004	2850004	2169816	2821985	2142549
No.parameters	4	17	18	6	18	6
Pseudo-R-squared	0.207	0.227	0.228	0.199	0.227	0.188
Log-Likelyhood	-107222.2	-90254.9	-90153	-88429.1	-89993.7	-89310.2
Initial log likelihood	-135283.2	-116791	-116791	-110397.7	-116435.9	-109981.9
AIC	214452.3	180543.8	180342.1	176870.2	180023.3	178632.4
BIC	214504.8	180762.4	180573.6	176945.7	180254.7	178707.9

t statistics in parentheses; + p <0.10, * p <0.05, ** p <0.01, *** p <0.001

References

Akbari, A. H., & Aydede, Y. (2012). Effects of immigration on house prices in canada. *Applied Economics*, *44*(13), 1645–1658.

Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management*, *58*(1), 1–14.

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.

Bishop, K. C. (2008). A dynamic model of location choice and hedonic valuation. *Unpublished, Washington University in St. Louis*.

Cebula, R. J. (1979). The determinants of human migration.

Cheshire, P. C., & Magrini, S. (2006). Population growth in european cities: weather matters–but only nationally. *Regional studies*, *40*(1), 23–37.

Chinadaily. (2017). *Young chinese leave big expensive cities due to soaring housing prices.* Retrieved from `http://www.chinadaily.com.cn/china/2017-04/11/content_28875148.htm`

Clark, D. E., & Hunter, W. J. (1992). The impact of economic opportunity, amenities and fiscal factors on age-specific migration rates. *Journal of Regional Science*, *32*(3), 349–365.

Clark, W. A., & Onaka, J. L. (1985). An empirical test of a joint model of residential mobility and housing choice. *Environment and Planning A*, *17*(7), 915–930.

Clark, W. A., & Van Lierop, W. F. (1987). Residential mobility and household location modelling. *Handbook of regional and urban economics*, *1*, 97–132.

Dahl, M. S., & Sorenson, O. (2010). The migration of technical workers. *Journal of Urban Economics*, *67*(1), 33–45.

d'Albis, H., Boubtane, E., & Coulibaly, D. (2017). International migration and regional housing markets: Evidence from france.

DaVanzo, J., et al. (1980). Micro economic approaches to studying migration decisions. *Micro economic approaches to studying migration decisions.*(1201-Nichd).

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.

Freeman, J. (2017). *China's whack-a-mole housing boom.* Retrieved from `https://www.citylab.com/equity/2017/07/hebei-hengshui-property-speculation-urbanization/532581/`

Fu, Y., & Gabriel, S. A. (2012). Labor migration, human capital agglomeration and regional development in china. *Regional Science and Urban Economics*, *42*(3), 473–484.

Glaeser, E. L. (2008). *Cities, agglomeration, and spatial equilibrium.* Oxford University Press.

Glaeser, E. L., Kolko, J., & Saiz, A. (2001). Consumer city. *Journal of economic geography*, *1*(1), 27–50.

Gleave, D., & Cordey-Hayes, M. (1977). Migration dynamics and labour market turnover. *Progress in Planning*, *8*, 1–95.

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.

Greenwood, M. J. (2014). *Migration and economic growth in the united states: National, regional, and metropolitan perspectives.* Academic Press.

Greenwood, M. J., Hunt, G. L., Rickman, D. S., & Treyz, G. I. (1991). Migration, regional equilibrium, and the estimation of compensating differentials. *The American Economic Review*, *81*(5), 1382–1390.

Gyourko, J., Saiz, A., & Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The wharton residential land use regulatory index. *Urban Studies*, *45*(3), 693–729.

Hunt, G. L., & Mueller, R. E. (2004). North american migration: returns to skill, border effects, and mobility costs. *The Review of Economics and Statistics*, *86*(4), 988–1007.

Jeanty, P. W., Partridge, M., & Irwin, E. (2010). Estimation of a spatial simultaneous equation model of population migration and housing price dynamics. *Regional Science and Urban Economics*, *40*(5), 343–352.

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

Lerman, S. R. (1976). Location, housing, automobile ownership, and mode to work: a joint choice model. *Transportation Research Record*, *610*, 6–11.

Levy, B. L., Mouw, T., & Perez, A. D. (2017). Why did people move during the great recession? the role of economics in migration decisions. *RSF*.

Lewis, D. J., Hunt, G. L., & Plantinga, A. J. (2002). Public conservation land and employment growth in the northern forest region. *Land Economics*, *78*(2), 245–259.

Liang, W., Lu, M., & Zhang, H. (2016). Housing prices raise wages: Estimating the unexpected effects of land supply regulation in china. *Journal of Housing Economics*, *33*, 70–81.

Liu, Z., Wang, Y., & Chen, S. (2017). Does formal housing encourage settlement intention of rural migrants in chinese cities? a structural equation model analysis. *Urban Studies*, *54*(8), 1834–1850.

Lux, M., & Sunega, P. (2012). Labour mobility and housing: the impact of housing tenure and housing affordability on labour migration in the czech republic. *Urban Studies*, *49*(3), 489–504.

Mueser, P. R., & Graves, P. E. (1995). Examining the role of economic opportunity and amenities in explaining population redistribution. *Journal of Urban Economics*, *37*(2), 176–200.

Mundial, B. (2009). *World development report 2009: Reshaping economic geography.* The World Bank, Washington DC.

Niu, G., & Zhao, G. (2017). Living condition among chinaâĂŹs rural–urban migrants: recent dynamics and the inland–coastal differential. *Housing Studies*, 1–18.

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.

Quigley, J. M. (1976). Housing demand in the short run: An analysis of polytomous choice. In

Explorations in economic research, volume 3, number 1 (pp. 76–102). NBER.

Quigley, J. M. (1978). Housing markets and housing demand: Analytic approaches. *Urban housing markets: recent directions in research and policy, University of Toronto Press, Toronto*.

Rabe, B., & Taylor, M. P. (2012). Differences in opportunities? wage, employment and house-price effects on migration. *Oxford Bulletin of Economics and Statistics*, *74*(6), 831–855.

Rappaport, J. (2007). Moving to nice weather. *Regional Science and Urban Economics*, *37*(3), 375–398.

Roback, J. (1982). Wages, rents, and the quality of life. *Journal of political Economy*, *90*(6), 1257–1278.

Saiz, A. (2007). Immigration and housing rents in american cities. *Journal of urban economics*, *61*(2), 345–371.

Saiz, A. (2010). The geographic determinants of housing supply. *The Quarterly Journal of Economics*, *125*(3), 1253–1296.

Saiz, A., & Wachter, S. (2011). Immigration and the neighborhood. *American Economic Journal: Economic Policy*, *3*(2), 169–188.

Saks, R. E. (2008). Job creation and housing construction: Constraints on metropolitan area employment growth. *Journal of Urban Economics*, *64*(1), 178–195.

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

Verdugo, G. (2015). Public housing magnets: public housing supply and immigrantsâĂŹ location choices. *Journal of Economic Geography*, *16*(1), 237–265.

Williams, R. C. (1979). A logit model of demand for neighborhood. *The economics of neighborhood*, 17–42.

Wooldridge, J. (2007). Control function and related methods. *NBER Summer Institute*.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Zabel, J. E. (2012). Migration, housing market, and labor market responses to employment shocks. *Journal of Urban Economics*, *72*(2), 267–284.