Housing Prices and Inter-urban Migration

Cécile Detang-Dessendre[a]

Gary L. Hunt[b]

Virginie Piguet[a]

and

Andrew J. Plantinga[c]

Draft: June 8, 2011

[a]Institut National de la Recherche Agronomique, Dijon, France. [b]School of Economics, University of Maine, Orono, ME, USA. [c]Department of Agricultural and Resource Economics, Oregon State University, Corvallis, OR, USA.

Housing Prices and Inter-urban Migration

Abstract.

Understanding the causes and consequences of human migration has long been of interest to urban and regional economists. Empirical studies build on the theoretical results of Roback (1982) and Mueser and Graves (1995) by estimating the effects of wages, housing prices, and amenities on inter-area migration. Findings with respect to amenities are clear (e.g., Rappaport 2007), and household-level studies consistently find that relative wages or incomes increase the probability that a household will select a given location (e.g., Berger and Blomquist 1992). In contrast, the results for housing prices are inconclusive. Studies that include area-level measures (e.g., median housing price for a metropolitan area) find a mix of negative, positive, and insignificant effects on inter-area migration decisions (e.g., Hunt and Mueller 2004). Many migration studies exclude housing price measures.

The purpose of this paper is to investigate the role of housing prices in influencing inter-urban household migration decisions. An important contribution of the study is the development of a new method for representing housing prices in migration analyses. Following the approach commonly used to model wages in studies of household migration, we identify the form of the utility function for which individual-specific housing prices can be predicted for unselected areas as a function of individual characteristics. Our theoretical results guide the development of an empirical measure of housing costs that accounts for the decision to own or rent and the cost of holding housing capital.

We test our housing cost measure using the 2000 PUMS to identify point-to-point migration decisions for a large sample of college-educated males residing in 291 U.S. metropolitan areas. We estimate conditional logit models of metropolitan area choice, controlling for wages, a large range of amenities, and expected housing costs. Our key finding is that our proposed housing cost measure yields the expected results (higher housing prices reduce the probability that an area is selected), which is robust to alternative specifications and samples. We re-estimate our model using three alternative metropolitan area measures of housing costs: median house price, average apartment rent, and average urban land rent. We find that these measures consistently yield counterintuitive results.

Housing Prices and Inter-urban Migration

1. Introduction

Understanding the causes and consequences of human migration has long been of interest to urban and regional economists. Roback (1982) explained the equilibrium distribution of human population by differences in the non-traded amenities at each location. These amenity differences produce wage and rent differentials that, in equilibrium, leave households and firms indifferent to changing locations. Mueser and Graves (1995) modify the Roback model by making instantaneous adjustment to equilibrium costly for households and firms. Migration emerges in their model as a short-run response to disequilibrium in labor and housing markets. Absent any shocks to exogenous factors such as preferences and technology, the sequence of short-run equilibria in these markets converges to the Roback equilibrium in the long run.

Empirical studies build on these theoretical results by estimating the effects of wages, housing prices, and amenities on migration. Findings with respect to amenities are clear. Area measures of population and migration as well as household location decisions are significantly related to climate (Mueser and Graves 1995, Clark and Murphy 1996, Hunt and Mueller 2004, Cheshire and Magrini 2006, Rappaport 2007, Poston et al. 2009, Eichman et al. 2010), air quality (Seig et al. 2004, Bayer et al. 2008), recreational opportunities (Duffy-Deno 1998, Lewis et al. 2002), cultural amenities (Clark and Hunter 1992), and crime rates (Gottlieb and Joseph 2006). Housing prices and wages are endogenous to area-level migration (Mueser and Graves 1995), and so these variables are typically excluded from analyses with aggregate data. However, it is reasonable to treat households as price-takers in labor and housing markets and, thus, the effects of wages and housing prices on migration decisions can be measured in studies using household data. In such studies, higher relative wages or income are consistently found to increase the probability that a household will select a given location, all else equal (Berger and Blomquist 1992, Davies et al. 2001, So et al. 2001, Hunt and Mueller 2004, Bayer et al. 2008, Bishop 2008, Kennan and Walker 2009, Dahl and Sorenson 2010). In contrast, the results for housing prices are much less clear. Studies that include area-level measures (e.g., median housing price for a county or metropolitan area) find a mix of negative, positive, and insignificant effects on migration decisions (Berger and Blomquist 1992, Hunt and Mueller 2004, Gottlieb and Joseph 2006, Bishop 2008).[1] Other studies do not control for housing prices or do not explicitly measure their effects (Davies et al. 2001, Bayer et al. 2008, Detang-Dessendre et al. 2008, Kennan and Walker 2009, Dahl and Sorenson 2010).[2]

The purpose of this paper is to investigate the determinants of inter-urban migration decisions, with a particular emphasis on the role of housing prices. An important contribution of the study is the development of a new method for representing housing prices in migration analyses. Our proposed approach is inspired by the method commonly used to model wages in studies of household migration (e.g., Hunt and Mueller 2004, Bayer et al. 2008). In the case of wages, a reduced-form wage equation is estimated for each area using observations of wage rates and characteristics of individuals such as age, race, and education. These equations are then used to predict the wage an individual would earn in unselected areas conditional on their attributes. We identify the form of the utility function under which a similar approach can be used to predict individual-specific housing prices for each area using individual characteristics. Our

[1] It is important to distinguish between inter-area migration and intra-area location changes. While housing prices clearly matter for moves in both cases, we are primarily interested in their effect on migration at the scale of metropolitan areas, counties, and states. Intra-area studies that examine effects of housing prices on household location decisions include Chan (2001), So et al. (2001), Engelhardt (2003), Seig et al. (2004), and Ferreira et al. (2010).

[2] Chen and Rosenthal (2008) construct area-level quality of life indices that reflect wages and housing prices. They investigate how changes in these indices for migrants are influenced by individual-level factors such as age and gender.

theoretical results guide the development of an empirical measure of housing costs that accounts for the decision to own or rent and the cost of holding housing capital.

We test our housing cost measure using data from the 5% sample of the 2000 Public Use Microdata Survey (PUMS) to identify point-to-point migration decisions for a large sample of household heads residing in 292 U.S. metropolitan areas. We estimate conditional logit models of metropolitan area choice, controlling for wages, a large range of amenities, and expected housing costs. Our proposed method for measuring housing costs yields the expected result that, all else equal, higher housing costs reduce the probability that a metropolitan area will be chosen. This finding is robust to alternative specifications and samples. We then re-estimate our model using three alternative metropolitan area measures of housing costs: median house price, average 2-bedroom apartment rent, and average per-acre urban land rent. We find that these measures consistently yield counterintuitive results. Potential migrants are likely to base decisions on the costs of housing that they themselves would select, rather than what the average metropolitan area resident would choose, implying measurement error in the metropolitan area housing cost measures. Correlation between this measurement error and unobservable area attributes could be the source of bias in the coefficient estimates on housing costs. In contrast, our proposed measure is based on a projection of individual-level housing costs into individual attributes.

The next section presents a model of household migration that provides the theoretical underpinnings for our housing cost measure and empirical analysis. Section 3 describes the data we use and the specification of choice sets for households. In section 4, we discuss the estimation procedures for the wage equations, housing cost measures, and conditional logit models of metropolitan area choice. Section 5 presents our results and discussion and conclusions are provided in a final section.

2. Theory

Individuals are assumed to choose locations conditional on expected wages and housing costs, the amenities of the area, and costs associated with moving. For individual i, the utility in the area j is specified:

(1) $$U_{ij} = U(X_{ij}, z_{ij}; A_j, C_i)$$

where X_{ij} is a vector of housing attributes and z_{ij} is a composite numeraire good. These are choice variables for individual i. The utility derived by the ith individual in the jth area also depends on the individual's characteristics (age, gender, etc.), denoted by the vector C_i, and the amenities in area j, denoted by the vector A_j. If area j differs from the starting location, then A_j includes measures of the dis-amenities associated with moving (e.g., moving costs). Conditional on choosing area j, the individual maximizes utility subject to the budget constraint:

(2) $$P_j X_{ij} + z_{ij} = I_{ij}$$

where P_j is a vector of implicit prices for housing attributes and I_{ij} is the income that individual i expected to earn in area j. Individuals are assumed to be price takers in housing, labor, and goods markets. Furthermore, the price of the composite good is assumed to be constant across areas. In the case of labor markets, individuals form expectations of the future equilibrium wages that they will earn in each area. Following the hedonic price literature, these wages will be a function of an individual's characteristics; thus, we can write income as $I_{ij} = I_j(C_i)$. The form of the income function varies by area because of differences in industrial composition, transportation costs, amenities, and other factors.

In the migration problem, an individual will choose the area that gives the highest utility. Thus, we must solve the utility maximization problem for each area to find the indirect utility function V_{ij}. We assume a quasi-linear utility function of the following form:

$$U_{ij} = u(X_{ij}; A_j, C_i) + z_{ij} \tag{3}$$

This specification assumes additive separability between the numeraire good and goods associated with the migration choice (housing attributes and area amenities). This specification permits individuals to choose different housing bundles in different areas and to make trade-offs among these attributes. However, it restricts individuals from trading off housing attributes and the numeraire good, which by construction gives constant marginal utility. The level of the numeraire good can vary by individual and area. We adopt this specification because it will allow us to specify housing price as a reduced-form function of individual attributes, as we now show.

The solution to the utility maximization problem gives the demands $X^*(P_j; A_j, C_i)$ and $z^*(I_{ij}; C_i)$. With positive consumption of the numeraire good, which we assume here, an individual allocates a portion of their income to housing and any remaining income is spent on the numeraire good. As such, the demands for housing attributes do not depend directly on income. For our empirical application, below, we do not observe the implicit prices for housing attributes, P_j, but we do observe total expenditures on housing, which we denote H_{ij}. Using the results from above, we can write this as

$$\begin{aligned} H_{ij} &= P_j X^*(P_j; A_j, C_i) \\ &= H_j(C_i) \end{aligned} \tag{4}$$

Notice that the function H_j varies by area due to area differences in the implicit prices of housing attributes and amenities. Recalling $I_{ij} = I_j(C_i)$, substitution of $H_j(C_i)$ and $z^*(I_{ij};C_i)$ into the utility function (3) gives the indirect utility function:

(5) $$V_{ij} = V(H_j(C_i), I_j(C_i), A_j, C_i)$$

which is the theoretical basis for the empirical model of migration.

3. Data and Choice Set Specification

Our main data source is the 5% sample of the 2000 Public Use Microdata Survey (PUMS), which includes approximately 5 million U.S. individuals. The survey provides a large number of demographic and socioeconomic variables, including measures of age, income, employment, and educational attainment (Table 1). Residence in 2000 is reported at the level of the Public Use Microdata Area (PUMA). PUMAs are geographic areas, designated by the Bureau of the Census, that contain at least 100,000 people. Respondents to the PUMS are also asked about their residence in 1995. If an individual's residence changed, the former residence is reported at the level of the Migration PUMA (MIGPUMA). MIGPUMAs are agglomerations of one or more PUMAs. In the 2000 PUMS, there are 2,101 PUMAs and 1,050 MIGPUMAs. The PUMS allows us to model point-to-point migration decisions between 1995 and 2000. While the PUMS provides observations of migration decisions at the level of PUMAs and MIGPUMAs, we model migration between metropolitan areas (MAs) for reasons discussed below.

We estimate our basic migration model with a sample of 24,604 working-age, college-educated male MA residents, representing approximately 13.5 million individuals. Given our emphasis above on wages as a key determinant of migration decisions, we focus on noninstitutionalized adults of working age (25-64). Results from Hunt and Mueller (2002) indicate that labor market relationships differ for males and females and so we limit our attention to male migrants. Finally, to capture the most mobile segment of the working-age population, we study individuals with an educational attainment of 4-year college degree or higher. Preliminary estimation revealed that pooling samples of individuals with different education levels was not justified. As a check on the generality of our findings, we discuss estimation results done with samples of male migrants with an educational attainment of some college.

Given the emphasis on working-age adults, it is necessary that our set of origins and destinations conform to distinct labor markets. MAs are likely to satisfy this criterion because they are delineated so that the communities within them exhibit a high degree of social and economic integration. Although we would wish to include non-MA residents, it is difficult to identify distinct labor markets in non-MA areas. There is no counterpart that we know of to the MA – areas exhibiting a high degree of social and economic integration – developed for non-MA areas. Even if these were available, they would need to match reasonably well the geographical scale of MIGPUMAs. For non-MA areas, MIGPUMAs are frequently too large to reasonably correspond to labor market areas.[3] Our sample is thus restricted to MA residents, defined as individuals who lived in an MA in 1995 and 2000.

In 2000, there were 324 MAs in the U.S. This figure includes 251 metropolitan statistical areas (MSAs), 12 New England Consolidated Metropolitan Areas (NECMAs), and 61 Primary Metropolitan Statistical Areas (PMSAs). We matched each of the 324 MAs to one or more MIGPUMAs. If a portion of a MIGPUMA lay outside of an MA boundary, we retained the MIGPUMA only if at least 75% of its population lived within the MA. In the case of 24 MAs,

[3] For example, a single MIGPUMA in eastern Oregon encompasses an area greater than the combined area of Connecticut, Rhode Island, Massachusetts, New Hampshire, and Vermont.

no matches to MIGPUMAs could be made using this criterion. These MAs had relatively small populations (on average, approximately 126,000 persons) and were dropped from the analysis. We also excluded eight MAs in Alaska, Hawaii, and Puerto Rico, and one MA (Auburn-Opelika) with missing data, leaving us with 291 MAs, comprised of 576 MIGPUMAs.

Area attributes were developed for the final set of 291 MAs. The main data source is the State and Metropolitan Area Data Book 1997-98 (U.S. Bureau of the Census 1998), which provides observations of demographic, social, and economic variables for all MAs and years ranging from 1990 to 1997. We use lagged area measures (as close to 1995 as possible) to explain migration decisions occurring between 1995 and 2000. Additional measures are constructed using county data from U.S. Bureau of the Census (1994) and McGranahan (1999). Because MAs are agglomerations of counties, we can compute MA averages using county-level observations. Table 2 provides a list of and sources for the area measures that were developed.

4. Methods

In this section, we discuss the estimation of wage and housing cost equations. In all models, the explanatory variables are sets of individual characteristics. For these estimations, we use full samples of individuals residing in each MA. That is, we do not restrict our sample to working-age, college-educated males, as we for the migration analysis, in order to increase variation in the data. Following this discussion, we present the methods for modeling migration decisions.

Wage Equation Estimation

The PUMS provides information on the wages earned by individuals in 2000. Of course, we observe wage only for the location where an individual lived and worked. Estimates of wages in unselected MAs are needed for estimation of the migration model. To this end, we estimate a log-linear wage equation for each MA using data on all individuals in the PUMS who resided there in 2000. The dependent variable (*Wage*) is the natural log of average weekly salary wages in 2000 and independent variables are a vector of individual attributes (C_i in section 2) that includes gender, age, race, marital and family status, language, educational attainment, usual hours worked, type of position, and sector of employment (Table 1). In each MA, we dropped individuals in the top and bottom 1% of the wage distribution to reduce the influence of very high and very low wages on our estimates.

The MA-specific wage equations are used to estimate the wage each individual would earn in each MA, conditional on the individual's gender, age, race, etc. These estimates are denoted $Wage_{ij}$. Ideally, we would have estimated the wage equations with 1995 data, so that wage predictions are lagged with respect to the migration decision. However, PUMS data are available in either 1990 or 2000. If the parameters of the wage equation are not changing appreciably over time, then either data set can be used. If they are changing, then the 1990 data have the disadvantage that the wage prediction and the migration decision are separated by 5 to 10 years. The 2000 data are preferable in this respect (the separation is between 0 and 5 years), but have the shortcoming that some of the information may have been unobservable to potential migrants in 1995. A further consideration is that definitions of PUMAs and MIGPUMAs differ somewhat between the 1990 and 2000 PUMS. This complicates the use of wage data from the 1990 PUMS and the migration data from the 2000 PUMS, and partially explains our decision to estimate the wage equations with 2000 data.

Housing cost estimation

A similar approach is used to predict the cost of housing for each individual in each MA. In developing the housing cost measure, the first issue to contend with is the choice an individual makes between renting and owning. We allow the likelihood that an individual owns or rents to differ by area because, for example, the same individual may rent an apartment if they live in New York City, but buy a house if they live in Miami, Florida. We assume that the probability of ownership depends on individual attributes. This formulation implicitly accounts for the role of income in influencing home ownership since, as in section 2, income is a function of individual attributes. The PUMS includes a variable indicating whether a household head lives in a rented (*Renter*) or owned home (*Owner*). Homes acquired with a mortgage or other lending arrangements are classified as owned. Using the full sample of household heads for each area, we estimate probit models for the binary ownership decision. This yields area-specific functions for the probability of ownership that depends on individual attributes. These functions are used to estimate the ownership probability, denoted α_{ij}, for each individual *i* and area *j*.

Separate housing price equations are then estimated for rented and owned homes.[4] The PUMS indicates the monthly rent paid by household heads or, for owners, the value of their housing unit. We multiply the monthly rent by 12 to obtain the annual rent. The logs of the annual rent (*Annualrent*) and homeowner value (*Ownervalue*) variables are regressed on the corresponding set of individual attributes using the full sample of household heads in each area. As with the wage and ownership analysis, this yields functions that are used to estimate annual values of rented and owned homes for each individual and each area. These estimates are denoted $Annualrent_{ij}$ and $Ownervalue_{ij}$.

[4] We do not distinguish between apartments, single detached houses, etc. Thus, all types of housing may be included in rented and owned homes.

The final step is to derive an annual housing cost measure (H_{ij} in section 2). To do so, we must express the value of owned housing as an annualized equivalent, which requires an estimate of the cost of holding a unit of housing capital. Following the Jorgensonian cost of capital formulation, we specify this as the financial cost of holding housing capital less the rate of housing price appreciation.[5] The first term is approximated with the January 1, 2000 rate of return on 3-month Treasury bills (5.33%) and the second is estimated as the average annual percentage change in the metropolitan area median house price between 1990 and 2000 (*House value change*). The difference between these two terms gives a metropolitan area-specific capital cost r_j. Our annual housing cost measure is, thus,

(6) $$Housingcost_{ij} = \alpha_{ij} r_j Ownervalue_{ij} + \left(1 - \alpha_{ij}\right) Annualrent_{ij}$$

Consistent with the theoretical development in section 2, $Housingcost_{ij}$ is estimated for each area and individual using individual attributes and functions specific to each MA.

Migration decisions

We estimate a nested logit model of migration over the period 1995 to 2000 using our sample of working-age, college-educated males. Individuals decide whether to remain in the same location (the MA where they lived in 1995) or move to a new MA. The stay/move decision is assumed to depend on individual attributes (C_i). Conditional on moving, the individual must select an MA and will do so to maximize utility. According to (5), the maximum utility from location j depends on expected wage ($Wage_{ij}$), expected housing cost ($Housingcost_{ij}$), individual characteristics (C_i), and area attributes (A_j). We specify the utility that individual *i* obtains from the *j*th MA as:

[5] We assume that marginal tax rates are constant across individuals and areas and that there are no investment tax credits or depreciation allowances.

(7) $$V_{ij} = \begin{cases} \alpha_0^{'} C_i + \beta^{'} A_j + \gamma Wage_{ij} + \delta Housingcost_{ij} + \varepsilon_{ij} & j = 0 \\ \alpha_1^{'} C_i + \beta^{'} A_j + \gamma Wage_{ij} + \delta Housingcost_{ij} + \varepsilon_{ij} & j = 1, 2, ... \end{cases}$$

where *j*=0 indicates the individual's origin (the 1995 MA), $\alpha_0, \alpha_1, \beta, \gamma, \delta$ are conformable parameter vectors, and ε_{ij} is a random disturbance with a type I extreme value distribution.

Several remarks on the specification in (7) are in order. First, because the individual attributes in C_i do not vary by MA, they explain only the decision to stay or move. The parameters on these variables (α_0, α_1) must differ to capture utility differences associated with staying or moving. Second, the area attributes in A_j differ across MAs but not across individuals. If one thought that the marginal utility of a given attribute is different among individuals, one approach would be to interact the area attributes with individual characteristics. For example, one might interact an MA-level measure of cultural amenities with an individual-level measure of educational attainment. Alternatively, one can accomplish the same result by estimating models for selected cohorts of individuals, the approach that we pursue below. Finally, the expected wage and housing cost variables differ by both MAs and individuals. There is an important measurement issue to note concerning these variables. The influence of the area attributes on utility is measured by the $\beta^{'} X_j$ term in (7). However, the intercept terms in the wage and housing cost models for the *j*th MA should capture compensating wage differentials related to these same attributes (Blomquist et al. 1988). By including the intercept terms when we compute the expected wages and housing costs, we ensure that $\beta^{'} X_j$ captures the total contribution of the area attributes to utility.[6]

[6] An example may help to clarify this claim. Suppose that a local amenity provides 50 utils and this causes a downward adjustment in wages equivalent to 20 utils. If we neglect the compensating differential (e.g., by omitting the intercept term when we calculate expected wage), then utility implicitly rises by 20 utils. The term for this

Following Train (2003), the probability that individual *i* chooses to stay (*m*=0) or move (*m*=1) is given by:

$$P_{im} = \frac{e^{\alpha_m C_i + \lambda_m I_{im}}}{\sum_{m=0}^{1} e^{\alpha_m C_i + \lambda_m I_{im}}} \tag{8}$$

where λ_m measures the degree of substitutability among alternatives under choice *m*,

$$\begin{aligned} I_{i0} &= (\beta' X_0 + \gamma Wage_{i0} + \delta Housingcost_{i0}) / \lambda_0 \\ I_{i1} &= \ln \sum_{j=1}^{J} e^{(\beta' X_j + \gamma Wage_{ij} + \delta Housingcost_{ij})/\lambda_1} \end{aligned} \tag{9}$$

and *J* is the total number of MAs. Conditional on moving, the probability that the *i*th individual selects the *j*th MA equals:

$$P_{ij|m=1} = \frac{e^{(\beta' X_j + \gamma Wage_{ij} + \delta Housingcost_{ij})/\lambda_1}}{\sum_{j=1}^{J} e^{(\beta' X_j + \gamma Wage_{ij} + \delta Housingcost_{ij})/\lambda_1}} \tag{10}$$

whereas the probability that the individual selects the origin, conditional on staying, is one (i.e., $P_{i0|m=0} = 1$). This model is a case of the partially degenerated nested logit model analyzed in detail by Hunt (2000). To estimate the model parameters, we apply the normalizations $\lambda_0 = 1$ and $\alpha_0 = 0$.

The migration model is estimated with a large sample (24,604 individuals). To reduce the size of the estimation problem, we limit the number of alternatives in the choice set. We include the origin and the selected MA (if different from the origin) and then randomly sample from the unselected MAs to bring the total size of the choice set to 100. This procedure has been

amenity in (7) would then add only 30 utils. In contrast, if we control for the compensating differential by including the intercept term, the amenity term in (7) adds 50 utils, the total contribution of the amenity to utility.

shown to give consistent estimates of the parameters for the model with the full choice set (Ben-Akiva and Lerman 1985).

We use the *Age*, *Married*, *White*, and *Black* variables to explain the stay or move decision (the omitted categories are separated, divorced, single, and other race) (Table 1). In addition to the *Wage* and *Housingcost* variables, area choice is assumed to depend on migration costs and MA-level amenities (definitions and data sources are given in Table 2). Migration costs are measured with the variables *Distance*, equal to the radial distance between the centroids of the most populous county in the origin and destination MAs, and *Population Spread*, equal to the difference in population density between the origin and destination MA. We assume that an individual's origin MA reveals their preference with respect to MA size. Thus, *Population Spread* controls for the cost in utility terms associated with migration to a larger or smaller MA. MA-level amenities include four climate variables (*January temperature*, *July temperature*, *July humidity*, and *Precipitation*), two topography variables (*Mountain*, *Plainshills*), variables for proximity to major water bodies (*Gulf Coast*, *Great Lakes*, *North Atlantic*, *Pacific*, and *South Atlantic*), and variables measuring air quality (*Ozone*), crime (*Violent crime*), and economic opportunity (*Employment growth*). We hypothesize that, all else equal, individuals prefer warmer winters, cooler and drier summers, and less annual precipitation. We expected positive signs on the topography variables, indicating that migrants prefer MAs with hills and mountains, and on the coastal variables.[7] Poor air quality and crime are expected to lower utility, whereas lagged MA employment growth is expected to increase the attractiveness of an area.

We estimate four versions of model that differ according to the housing price variable included. Version I includes our proposed measure, *Housingcost*. For version II, we use the

[7] The omitted topography variable is the proportion of the MA land area classified as plains or tablelands. The omitted coastal variable is an indicator variable for MAs that share no border with a major water body.

median value of owner-occupied housing in 1990 (*Median house value*). This variable has been used in a number of previous migration studies (Clark and Hunter 1992, Bishop 2008, Scott 2010). Version III includes the 40th percentile rental rate for a 2-bedroom apartment in 1995 (*Apartment rent*) and, finally, version IV uses a measure of urban land rent developed by Lubowski (2002) (*Land rent*). The rent variable was constructed by subtracting the value of structures from county measures of housing prices. Versions II, III, and IV include MA-level measures of the housing price. Because our house price, apartment rent, and land rent variables are all measured at the county level, we average them to form MA-level variables.

Summary statistics are presented in Table 3 for the variables used in the migration analysis. The average age of working-age, college-educated male MA residents in our sample is 42 years. Eighty-four percent of this sample is white and 71 percent is married. Turning next to the area variables, we see that approximately 34 percent of the land area in the MAs is classified as mountain or hills and 23 percent of the MAs are located next to a major water body. On average, there are approximately 60,000 violent crimes per 100,000 persons and the ozone standard is exceeded about 2 days per year. The average monthly apartment rent is $516 and the median house value is approximately $80,000, on average, which is similar to the average per acre land rent for urbanized land. Averaged over MAs and individuals, the mean value of our housing cost variable is $6,481 per month, or $77,770 per year, which is similar to the average of the median house value. The average weekly wage is $979, or about $50,000 on an annual basis.

5. Results

Wages and Housing Costs

Although we estimate separate wage equations for all 291 MAs, we present estimates for national-level models in Table 4 to indicate the general nature of the results. Almost of all of the estimated coefficients in the national-level wage equations are significantly different from zero at the 5% level and the signs of the coefficients are consistent with expectations (Lemieux, 2006; Chiswick, Miller, 2010). Wages are higher for white, college-educated married men who are fluent in English. Wages also increase with age, but at a decreasing rate. Wages fall for blacks (relative to the other race category) and separated or divorced people who have not completed college. Single individuals have higher wages than those who are separated or divorced, but number of children does not have a significant effect. Wages rise for those who work more hours per week, but at a diminishing rate. As expected, executives receive higher wages. Relative to working in the manufacturing sector, wages are lower for workers in agriculture, commerce, services, education, and administration and higher for workers in mining and energy, transportation, information/communication, and finance/insurance.

Also for illustration purposes, we produce national-level results for the probability of ownership and for the rental and owned value equations (Table 5). All coefficient estimates are significantly different from zero at the 5% level. The results show that the likelihood of ownership increases for males, whites, and married households, and with age, but at a decreasing rate. The higher is the educational attainment, the higher is the probability of owning. Finally, executives are more likely to own than non-executives, which likely reflects the influence of income on ownership. These results are consistent with those found by Hendershott et al (2009) using Australian data and by Painter et al. (2001) and Jepsen and Jepsen (2009) using U.S. data.

For the rental and homeowner value models, an illustrative set of national-level results (Table 5) suggest that housing is a normal good. That is, factors that increase wages (Table 4) tend to also increase housing expenditures, and vice-versa. When the household head is male, the rental and the owned value are higher than when the household head is female. Expenditures on housing increase with age, but at a diminishing rate, and they increase with educational attainment, number of children, and executive status. Expenditures are highest for married household heads, followed by single and separated or divorced heads. At the national level, the rental and owned values for blacks are smallest, followed by whites and other race. The finding that whites spend less on housing than households heads of other races contrasts with the result that whites have higher wages (Table 4), but is consistent with the study by Ilhandfeldt and Matinez-Vazquez (1986).

Migration choice

The results for the four migration models are presented in Table 6. For Model I, which includes our individual- and area-specific measure of housing costs (*Housingcost*), all of the coefficient estimates are significantly different from zero at the 5% level. The estimated value of the dissimilarity parameter (λ_1 in equations 9 and 10) lies in the unit interval, indicating that our model is consistent with utility maximization for all possible values of the explanatory variables (Train 2003). The results indicate that the likelihood of moving declines for older and married individuals and is lower for blacks and whites relative to other races. As expected, higher wages increase the likelihood that an MA is chosen, all else equal. MAs that are a greater distance from the origin MA, our proxy for higher moving costs, are less likely to be chosen, whereas a greater difference in population between the origin and destination MA increase the likelihood of the latter MA being chosen. We hypothesized, in contrast, that individuals would prefer MAs of similar size.

Most of the coefficients on the area variables in Model I have plausible signs. MAs on the Pacific and South Atlantic coasts are more likely to be chosen than inland MAs, whereas MAs adjacent to the Gulf Coast, the Great Lakes, and the North Atlantic coast are less likely to be selected. Higher lagged employment growth and fewer high ozone days increase the likelihood that an MA will be chosen. MAs with higher January temperatures and lower July temperatures and humidity and less annual precipitation are more desirable to migrants. Contrary to expectations, the coefficient on *Violent crime* is positive and the negative coefficients on the *Mountain* and *Plainhills* variables suggest that varied topography is less desirable to migrants. It is possible that these variables are correlated with other MA attributes, such as the effectiveness of policing in the case of the crime variable.

The coefficient of primary interest is the one on the housing cost variable. The coefficient on *Housingcost* is negative and significantly different from zero, indicating that migrants are more likely to select areas with lower housing costs, all else equal. This is in contrast to the results for Models II-IV. The three alternative MA-level housing cost measures, *Median house value*, *Apartment rent,* and *Land rent,* have positive and significant coefficients. Notably, the coefficients on the other variables are similar in sign and magnitude across the four versions of the model. As a robustness check, we repeat the analysis using a sample of working-age male MA residents with lower educational attainment (1-3 years of college). The results (not reported) are similar to those in Table 6. The estimated coefficient on *Housingcost* is negative and significantly different from zero, whereas two of three alternative housing cost measures (*Median house value*, *Apartment rent*) have positive coefficients. With this sample, the coefficient on *Land rent* has the expected negative sign.

6. Conclusions

The purpose of this paper has been to incorporate housing prices in the household migration decision in a way that reflects individual household consumption determinants along with individual determinants of wage income and mobility and along with various area amenities. Our approach specifies a utility function that is additively separable in housing, other goods, and area amenities. Our approach to incorporating housing also distinguishes owners and renters and consistently produces estimates that imply that higher housing costs, other things equal, lead to a lower probability of area selection. More traditional area-level, as opposed to household-level, housing cost measures typically produce a counterintuitive direct relationship between housing costs and probability of area selection. Use of such area-level measures in our analysis confirm these counterintuitive results; whereas use of our household-level housing cost measure reflects an inverse relationship between housing cost and area selection, *ceteris paribus*.

In addition to developing a method to incorporate housing costs in a manner that produces results that are in line with basic economic principles, our results are consistent with housing being a normal good, wage income being produced in accordance with Mincerian human capital principles, and amenity effects that are qualitatively equivalent to previously reported results in almost every dimension. We conclude that the theoretical and methodological approaches that we develop and implement empirically with U.S. Census microdata provide a means to obtain theoretically expected results for the effects of housing costs, wages, and amenities on household migration behavior.

References

Bayer, P., Keohane, N., and C. Timmins. 2009. Migration and Hedonic Valuation: The Case of Air Quality. *Journal of Environmental Economics and Management* 58(1): 1-14.

Berger, M.C., and G.C. Blomquist. 1992. Mobility and Destination in Migration Decisions: The Role of Earnings, Quality of Life, and Housing Prices. *Journal of Housing Economics* 2: 37-59.

Bishop, K.C. 2008. A Dynamic Model of Location Choice and Hedonic Valuation. Unpublished paper, Olin Business School, Washington University in St. Louis.

Chan, S. 2001. Spatial Lock-in: Do Falling House Prices Constrain Residential Mobility? *Journal of Urban Economics* 49: 567-586.

Chen, Y., and S.S. Rosenthal. 2008. Local Amenities and Life Cycle Migration: Do People Move for Jobs or Fun? *Journal of Urban Economics* 65(3): 519-37.

Cheshire, P.C., and S. Magrini. 2006. Population Growth in European Cities: Weather Matters – But Only Nationally. *Regional Studies* 40(1): 23-37.

Clark, D.E., and W.J. Hunter. 1992. The Impact of Economic Opportunity, Amenities and Fiscal Factors on Age-Specific Migration Rates. *Journal of Regional Science* 32(3): 349-65.

Clark, D.E., and C.A. Murphy. 1996. Countywide Employment and Population Growth: An Analysis of the 1980s. *Journal of Regional Science* 36(2): 235-256.

Dahl, M.S., and O. Sorenson. 2010. The Migration of Technical Workers. *Journal of Urban Economics* 67: 33-45.

Davies, P.S., Greenwood, M.J., and H. Li. 2001. A Conditional Logit Approach to U.S. State-to-State Migration. *Journal of Regional Science* 41(2): 337-360.

Detang-Dessendre, C., Goffette-Nagot, F., and V. Piguet. 2008. Life Cycle and Migration to Urban and Rural Areas: Estimation of a Mixed Logit Model on French Data. *Journal of Regional Science* 48(4): 789-824.

Duffy-Deno, K.T. 1998. The Effect of Federal Wilderness on County Growth in the Intermountain Western United States. *Journal of Regional Science* 38(1): 109-36.

Engelhardt, G.V. 2003. Nominal Loss Aversion, Housing Equity Constraints, and Household Mobility: Evidence from the United States. *Journal of Urban Economics* 53: 171-195.

Eichman, H., Hunt, G.L., Kerkvliet, J., and A.J. Plantinga. 2010. Local Employment Growth, Migration, and Public Land Policy: Evidence from the Northwest Forest Plan. *Journal of Agricultural and Resource Economics* 35(2):316-333.

Ferreira, F., Gyourko, J., and J. Tracy. 2010. Housing Busts and Household Mobility. *Journal of Urban Economics* 68: 34-45.

Gottlieb, P.D., and G. Joseph. 2006. College-to-Work Migration of Technology Graduates and Holders of Doctorates within the United States. *Journal of Regional Science* 46(4): 627-659.

Hendershott P.H., Ong, R., Wood, G.A., and P. Flatau. 2009. Marital History and Home Ownership: Evidence from Australia. *Journal of Housing Economics* 18: 13-24.

Hunt, G.L. 2000. Alternative Nested Logit Model Structures and the Special Case of the Partial Degeneracy. *Journal of Regional Science* 40(1):89-113.

Hunt, G.L., and R.E. Mueller. 2004. North American Migration: Returns to Skill, Border Effects, and Mobility Costs. *Review of Economics and Statistics* 86(4): 988-1007.

Ihlanfeldt, K.R., and J. Martinez-Vazquez. 1986. Alternative Value Estimates of Owner-Occupied Housing: Evidence on Sample Selection Bias and Systematic Errors. *Journal of Urban Economics* 20 : 356-369.

Jepsen, C., and L.K. Jepsen. 2009. Does Home Ownership Vary by Sexual Orientation? *Regional Science and Urban Economics* 39: 307-315.

Kennan, J., and J.R. Walker. 2005. The Effect of Expected Income on Individual Migration Decisions. Unpublished paper, Department of Economics, University of Wisconsin-Madison.

Lewis, D.J., Hunt, G.L., and A.J. Plantinga. 2002. Public Conservation Land and Employment Growth in the Northern Forest Region. *Land Economics* 78(2):245-59.

Lubowski, R.N. 2002. *Determinants of Land-Use Transitions in the United States: Econometric Analysis of Changes among the Major Land-Use Categories*. Ph.D. Dissertation, Harvard University, Cambridge, MA.

McGranahan, D.A. 1999. Natural Amenities Drive Population Change. Food and Rural Economics Division, Economic Research Service, U.S. Department of Agriculture. Report 781, 1-24.

Mueser, P.R., and P.E. Graves. 1995. Examining the Role of Economic Opportunity and Amenities in Explaining Population Redistribution. *Journal of Urban Economics* 37: 176-200.

Painter, G., Gabriel, S., and D. Myers. 2001. Race, Immigrant Status, and Housing Tenure Choice. *Journal of Urban Economics* 49: 150-167.

Poston, D.L., Zhang, L., Gotcher, D.J., and Y. Gu. 2009. The Effect of Climate on Migration: United States, 1995-2000. *Social Science Research* 38: 743-753.

Rappaport, J. 2007. Moving to Nice Weather. *Regional Science and Urban Economics* 37: 375-98.

Roback, J. 1982. Wage, Rents, and the Quality of Life. *Journal of Political Economy* 90(6): 1257-1278.

Seig, H., Smith, V.K., Banzhaf, H.S., and R. Walsh. 2004. Estimating the General Equilibrium Benefits of Large Changes in Spatially Delineated Public Goods. *International Economic Review* 45(4): 1047-1077.

So, K.M., Orazem, P.F., and D.M. Otto. 2001. The Effects of Housing Prices, Wages and Commuting Time on Joint Residential and Job Location Choices. *American Journal of Agricultural Economics* 83: 1036-1048.

Train, K.E. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Table 1. Individual-level Variables

Variable	Description
Male	Indicator variable for male
Female	Indicator variable for female
Age	Age in years
White	Indicator variable for white race
Black	Indicator variable for black race
Other	Indicator variable for other race
Married	Indicator variable for married
Separated/divorced	Indicator variable for separated or divorced
Single	Indicator variable for single
Children	Number of children
Household	Indicator variable for head of household
English	Indicator variable for English fluency
No English	Indicator variable for lack of English fluency
Less than high school	Educational attainment is less than high school
High school	Educational attainment is high school
Some college	Educational attainment is 1-3 years of college
College or more	Educational attainment is 4 years college or more
Wage	Log of annual salary wages divided by number of weeks worked
Usual work hours	Typical number of hours worked per week
Executive	Indicator variable for executive position
Not executive	Indicator variable for non-executive position
Owner	Indicator variable for home ownership
Renter	Indicator variable for home rental
Ownervalue	Value of an owned home
Annualrent	Annual rent
Manufacturing	Indicator variable for employment in manufacturing sector
Agriculture	Indicator variable for employment in agriculture sector
Mining and energy	Indicator variable for employment in mining and energy sector
Construction	Indicator variable for employment in construction sector
Commerce	Indicator variable for employment in commerce sector
Transportation	Indicator variable for employment in transportation sector
Information/communication	Indicator variable for employment in information or communication sector
Finance/insurance	Indicator variable for employment in finance or insurance sector
Services to enterprises	Indicator variable for employment in services to enterprises sector
Education	Indicator variable for employment in education sector
Services to individuals	Indicator variable for employment in services to individuals sector
Administration	Indicator variable for employment in administration sector

Note: all variables are measured in 2000 and taken from the PUMS 5% sample.

Table 2. Metropolitan Area Measures

Variable	Description	Source
Population density	Metropolitan area population in 1996 divided by land area	State and Metropolitan Area Data Book 1997-1998
Distance	Radial distance between centroids of metropolitan areas	Authors' calculation
Population spread	Difference in population density between pairs of metropolitan areas	Authors' calculation
Gulf Coast	Indicator variable for border with Gulf Coast	Authors' calculation
Great Lakes	Indicator variable for border with Great Lakes	Authors' calculation
North Atlantic	Indicator variable for border with North Atlantic	Authors' calculation
Pacific	Indicator variable for border with Pacific Ocean	Authors' calculation
South Atlantic	Indicator variable for border with South Atlantic	Authors' calculation
Violent	Violent crimes per 100,000 population, 1995	State and Metropolitan Area Data Book 1997-1998
Poverty rate	Percent of population living in poverty, 1993	State and Metropolitan Area Data Book 1997-1998
Employment growth	Average annual growth rate in total employment, 1990-1995	State and Metropolitan Area Data Book 1997-1998
January temperature	Mean temperature for January, 1941-1970	McGranahan (1999)
July temperature	Mean temperature for July, 1941-1970	McGranahan (1999)
July humidity	Mean relative humidity for July, 1941-1970	McGranahan (1999)
Plainshills	Proportion of land area classified as plains with hills or mountains	McGranahan (1999)
Mountain	Proportion of land area classified as hills and mountains	McGranahan (1999)
Topography	Dominant land surface form	McGranahan (1999)
Rainfall	Annual precipitation in inches	Lawrence Berkeley Laboratory (http://eande.lbl.gov/IEP/high-radon/data/lbnl-met.html)
Ozone	Maximum number of days any monitor exceeds ozone standard, 1995	U.S. Environmental Protection Agency (http://www.epa.gov/aqspubl1/annual_summary.html)
Mean housing value	Mean value of owner-occupied housing units, 1990	County and City Data Book 1994
Housing value change	Percent change in median value of owner-occupied housing units, 1990-2000	County and City Data Books, various dates
Land rent	Land rent per acre for all urbanized land, 1995	Lubowski (2002)
Apartment rent	40th percentile rental rate for 2 bedroom apartments, 1995	U.S. Department of Housing and Urban Development (http://www.huduser.org/datasets/fmr.html)

Table 3. Summary Statistics for Variables in the Migration Models

Variables	Individual		Area		Individual/area	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
Age	42.38	10.15				
Married	0.71	0.45				
Black	0.06	0.24				
White	0.84	0.37				
Wage					978.75	277.52
Distance					1727.51	1124.11
Population spread					-2000406	2851714
Mountain			0.34	0.44		
Plainhills			0.08	0.24		
Gulf Coast			0.05	0.23		
Great Lakes			0.05	0.23		
North Atlantic			0.04	0.20		
Pacific			0.05	0.23		
South Pacific			0.04	0.21		
Employment growth			0.02	0.01		
Violent crime			60015	30829		
January temperature			36.52	12.68		
July temperature			75.96	5.45		
July humidity			58.54	14.60		
Precipitation			36.47	13.53		
Ozone			1.75	6.71		
Housingcost					6480.85	4385.43
Median house value			80278	46276		
Apartment rent			515.85	128.49		
Land rent			82706	82798		
Number of observations	24604		291		2460400	

Notes: Because we randomly sample 100 MAs for each individual, the number of observations for the individual/area variables is 2460400=24604×100.

Table 4. National-level wage equation

Variable	Parameter	Standard error	*t*-statistic
Intercept	1.4846	0.001920	771.3
Male	0.0815	0.000249	327.8
Age	0.0167	0.000084	198.1
Age squared	-0.0002	0.000001	-163.6
White	0.0083	0.000353	23.5
Black	-0.0124	0.000456	-27.2
Married	0.0399	0.000313	127.5
Separated/divorced	-0.0008	0.000369	-2.2
Children	0.0001	0.000100	0.8
Household	0.0562	0.000240	233.8
English	0.0551	0.000432	127.4
Less than high school	-0.0600	0.000398	-150.9
Some college	0.0531	0.000265	200.1
College or more	0.1621	0.000311	521.3
Usual work hours	0.0234	0.000035	660.3
Usual work hours squared	-0.0002	0.000000	-442.3
Executive	0.1106	0.000264	419.3
Agriculture	-0.1508	0.001030	-145.7
Mining and energy	0.0544	0.000858	63.4
Construction	0.0005	0.000474	1.1
Commerce	-0.0562	0.000374	-150.3
Transportation	0.0135	0.000531	25.5
Information/communication	0.0137	0.000631	21.8
Finance/insurance	0.0029	0.000470	6.1
Services to enterprises	-0.0250	0.000433	-57.8
Education	-0.0760	0.000369	-206.1
Services to individuals	-0.1168	0.000411	-284.0
Administration	-0.0116	0.000484	-23.9

Dependent variable = Wage

The omitted categories are female, other, single, no English, high school, not executive, manufacturing

No. observations = 5379510

Adj. R-squared = 0.40

Table 5. National-level ownership, rental value, and owned value equations

	Parameter estimates		
Variable	Ownership	Rental value	Owned value
Intercept	-3.92	5.921	10331
	0.003	0.007	0.008
Male	0.127	0.027	0.038
	0.0004	0.001	0.001
Age	0.099	0.013	0.037
	0.0001	0.0004	0.0004
Age squared	-0.0007	-0.0002	-0.0003
	0.000002	0.000005	0.000004
White	0.405	-0.047	-0.074
	0.0006	0.001	0.002
Black	-0.098	-0.159	-0.338
	0.0007	0.002	0.002
Married	0.853	0.116	0.206
	0.0005	0.001	0.002
Separted/divorced	0.119	-0.016	-0.046
	0.0005	0.001	0.002
Children	0.075	0.018	0.033
	0.0002	0.0004	0.0004
English	0.559	-0.071	-0.064
	0.0007	0.002	0.002
Less than high school	-0.289	-0.149	-0.272
	0.0006	0.001	0.002
Some college	0.082	0.141	0.222
	0.0004	0.001	0.001
College or more	0.151	0.322	0.556
	0.0005	0.001	0.001
Executive	0.143	0.148	0.169
	0.0004	0.001	0.001
Dependent variable	Owner	ln(Annualrent)	ln(Ownervalue)
No. observations	3769967	1123432	2646535
Adj. R-square	NA	0.17	0.22

Notes: The omitted categories are female, other, single, no English, high school, not executive. Standard errors are given below parameter estimates.

Table 6. Estimation Results for the Nested Logit Models of Migration Choice

	Model I		Model II		Model III		Model IV	
Variable	Parameter	z-statistic	Parameter	z-statistic	Parameter	z-statistic	Parameter	z-statistic
Move/stay decision								
Intercept	1.292	-858.7	1.186	296.5	1.304	344.42	1.257	266.9
Age	-0.071	-106.2	-0.067	-828.9	-0.071	-859.07	-0.072	-694.8
Married	-0.168	-129.6	-0.189	-121.0	-0.167	-105.68	-0.112	-57.1
Black	-0.299	-107.7	-0.345	-150.3	-0.299	-129.79	-0.294	-102.2
White	-0.394	340.9	-0.411	-114.1	-0.394	-107.56	-0.533	-117.1
MA choice								
Wage	0.0004	231.4	0.0003	157.4	0.0003	176.54	0.0004	167.4
Distance	-0.0002	-342.0	-0.0002	-210.8	-0.0002	-333.88	-0.0002	-279.7
Population spread	0.209	355.2	0.214	249.5	0.201	344.5	0.212	291.5
Mountain	-0.018	-42.1	-0.047	-93.3	-0.021	-50.41	-0.029	-51.9
Plainhills	-0.024	-33.4	-0.033	-44.1	-0.038	-54.92	-0.021	-23.8
Gulf Coast	-0.019	-24.3	-0.022	-26.6	-0.017	-21.98	-0.012	-11.8
Great Lakes	-0.136	-197.5	-0.146	-160.2	-0.122	-184.16	-0.130	-152.8
North Atlantic	-0.025	-46.6	-0.081	-114.6	-0.061	-108.8	-0.037	-55.4
Pacific	0.030	37.3	0.008	8.8	0.021	26.53	0.022	21.7
South Atlantic	0.015	21.6	0.010	13.2	0.002	2.81	0.029	31.4
Employment growth	3.124	214.7	4.921	189.3	3.508	240.01	3.447	186.6
Violent crime	0.0000001	13.8	-0.0000004	-68.7	0.0000001	18.74	0.0000001	12.2
January temperature	0.004	173.0	0.003	121.7	0.003	139.29	0.003	91.1
July temperature	-0.009	-167.7	-0.009	-130.5	-0.007	-137.85	-0.006	-87.2
July humidity	-0.001	-73.3	-0.001	-80.3	-0.001	-73.43	-0.001	-64.7
Precipitation	-0.001	-36.0	0.001	49.3	-0.0002	-15.32	0.00001	0.5
Ozone	-0.002	-131.0	-0.001	-80.0	-0.002	-121.99	-0.002	-108.3
Housingcost	-0.0000007	-16.0						
Median house value			0.000001	140.0				
Apartment rent					0.0002	137.6		
Land rent							0.0000003	72.0
lambda_1	0.193	0.0005*	0.212	0.0009*	0.189	0.0005*	0.197	0.0007*
Number of individuals	13456202		13456202		13456202		8895995	

* Standard error

Note: The models are estimated with the sample of 24,604 working-age, college-educated male MA residents, who represent approximately 13.5 million individuals. Due to missing values of the land rent variable, this number is smaller for Model IV.