Microsimulation of Single-family Residential Land Use for Market Equilibria

By

Bin (Brenda) Zhou
Graduate Student Researcher
The University of Texas at Austin
6.508. Cockrell Jr. Hall
Austin, TX 78712-1076
brendazhou@mail.utexas.edu

and

Kara M. Kockelman
(Corresponding author)
Associate Professor and William J. Murray Jr. Fellow
Department of Civil, Architectural and Environmental Engineering
The University of Texas at Austin
6.9 E. Cockrell Jr. Hall
Austin, TX 78712-1076
kkockelm@mail.utexas.edu
Phone: 512-471-0210
FAX: 512-475-8744

Abstract

This paper investigates single-family residential development for housing market equilibria using microeconomic theory and disaggregate spatial data. A logit model and notions of price competition are used to simulate household location choices in six different scenarios, with either one or multiple employment center(s) and with low, medium and high value-of-travel-time assumptions. Consistent with bid-rent theory, housing market equilibrium for each scenario was reached in an iterative fashion. The spatial allocation of new households in the region of Austin, Texas illustrates the potential shape of things to come, with endogenously determined home prices and demographic distributions.

1. Introduction

As an essential part of urban travel behaviors, prediction of future land use patterns is of great interest to policy makers, developers, planners, transportation engineers and others. Residential land is on the order of 60% of developed land, dominating urban areas. Moreover, the emergence of commercial, industrial, office and civic uses is spatially correlated with residential development (see, e.g., Zhou and Kockelman 2005).

Numerous factors contribute to the complexity of housing location choices (see, e.g., Irwin and Bockstael 2004, Bina et al. 2005). Microeconomic theory tested using disaggregate spatial data offer behavioral foundations and a better understanding of such decisions. These theories of land use can be traced back to Von Thünen's (1826) concept of agricultural rents and travel costs around a market center, followed by Wingo's (1961) and Alonso's (1964) urban examples. These early models treat land as homogeneous and continuous, and recognize only one employment center. Moreover, they neglect taste heterogeneity.

Herbert and Stevens' model (1960) determined residential prices by maximizing aggregate rents subject to constraints on (total) land availability and the number of households to be accommodated. Senior and Wilson (1974) enhanced the Herbert-Stevens model by adding an entropy term to the objective function, reflecting preference dispersion among households. Both models treat spatial elements in an aggregate manner, using an exhaustive zone-based subdivision of the region. Recent, more advanced models (e.g., Anas and Xu 1999, Change and Mackett 2005) depict household distribution via general equilibrium and land use-transportation interactions. However, their complexity has greatly limited their application.

In contrast to these earlier models and methods, this investigation emphasizes parcel-level data (GIS-encoded) and considers taste heterogeneity of individual households via behavioral controls for demographic variables and random utility maximization (RUM). The model applied here relies on bid-rent theory which is both theoretically meaningful and practically feasible. This work examines single-family residential land development based on a microscopic equilibrium of the housing market for recent movers. Each home-seeking household is allocated to the location that offers it the highest utility, and each new home is occupied by the highest bidder. This ensures optimal allocation of land in the sense that each household chooses a home that most satisfies the household, while developers/land owners maximize profits/rents. The spatial distribution of households and the equilibrium home prices are endogenously determined, as the outcome of a housing market mechanism involving land and transport.

2. Data and methods

This section describes the data used to calibrate the location choice model and to reach single-family housing market equilibria. Both procedures were coded in GAUSS matrix programming language (Aptech Systems, 2001 and 2003).

2.1 Location choice model

Bina and Kockelman (2006) undertook a survey of Austin movers in 2005. Sampling half of Travis County's recent[1] home buyers, responses were obtained from over 900 households, or roughly 12% of recent buyers. The data set contains comprehensive information on household demographics, housing characteristics, reasons for relocation and preferences when facing different housing and location-choice scenarios.

Commute distance and cost have a bearing on one's residential location choice (see, e.g., Van Ommeren et al. 1999, Rouwendal and Meijer 2001, Clark et al. 2003, Tillema et al. 2006). The GIS-encoded addresses of homes and workplaces, accompanied by roadway network data, provide a direct measure of commute time[2] for all potential locations. Since value of travel time (*VOTT*) is not directly available in the data set, it was approximated as the average wage[3] of each

---

[1] Here "recent" means within the past 12 months (before the sampling date, and start of the survey).
[2] Commute time was calculated using Caliper's TransCAD software for shortest travel-time path under free-flow conditions.
[3] Part-time employed persons were assumed to work 1,000 hours per year, while full-time employed persons were assumed to work 2,000 hours per year.

household's employed members and was assumed to be equal over the household's employed members. In addition to work access, each potential home's (Euclidean) distance to the nearest of the region's largest 18 shopping centers (*DISTMALL*) helps explain the impact of shopping access on location choice. Furthermore, household annual (pre-tax) income (*HHINC*), home size (*SQFOOT*) and housing prices per interior/built square foot (*UNITP*) help explain the balance of home affordability (where price [equal to *UNITP*×*SQFOOT*] is divided by annual income) and households' preferences for larger home sizes.

Residential location choice was modeled via a multinomial logit framework. The random utility was specified as follows:

$$U_{hi} = \beta_1 + \beta_2 \frac{UNITP_i \times SQFOOT_i}{HHINC_h} + \beta_3 \left(\frac{UNITP_i \times SQFOOT_i}{HHINC_h} \right)^2 + \beta_4 \frac{SQFOOT_i}{1{,}000}$$
$$+ \beta_5 \sum_{n=1}^{N_h} \left(VOTT_h \times TT_{hin} \right) + \beta_6 \sum_{n=1}^{N_h} \left(DISTWORK_{hin} \right) + \beta_7 DISTMALL_i + \varepsilon_{hi}$$

where U_{hi} is the random utility of household *h* for choosing home *i*, the β's are parameters to be estimated, *DISTMALL, SQFOOT*, and *UNITP* are as defined above, N_h is the number of workers in household *h*, $VOTT_h$ is the household's approximate value of travel time, TT_{hin} is the network commute time for worker *n* in household *h* when residing in home *i*, $DISTWORK_{hin}$ is the corresponding Euclidean distance. The random component (ε_{hi}) is assumed to be independent identically distributed (IID) Gumbel, across households *h* and their alternatives *i*.

For model calibration, each household's choice set consisted of 20 alternatives: nineteen randomly drawn from the pool of all homes purchased by respondents in the recent mover survey, plus the chosen option. These model results are shown in Table 1. The model indicates a concave relationship between strength of preference (systematic utility) and the ratio of home price to annual income. The parameter values on the ratio and its squared term suggest that more expensive homes are preferred when the ratio is less 1.7, becoming less attractive as the ratio of price to income exceeds this threshold.

Larger homes, of course, are more desired, with *SQFOOT* increasing the likelihood of a home's selection, everything else constant. The negative signs associated with commute costs and Euclidean distances to workers' workplace(s) support the notion that households favor homes closer to their employed workers' jobs. Major mall access, however, was not favored; perhaps the potentially high volumes of traffic and congestion in the vicinity of major shopping centers offset any possible access gains. Other forms of shopping access may be desired, but require geocoding of far more, smaller shopping

2.2 Single-family housing market equilibrium

Microsimulation of single-family residential land development for housing market equilibrium was applied to the City of Austin and its 2-mile extraterritorial jurisdiction, assuming a 25% growth in household numbers[4]. Both the supply of and demand for homes were modeled explicitly.

On the supply side, undeveloped sites with potential for residential development were located using a year 2000 land use parcel map, obtained from the City of Austin's Neighborhood

[4] The study area accommodated about 304,800 households in the year 2000. With the projected 25% growth, the number of newly-added households is around 76,000 in the whole area.

Planning and Zoning Department. Undeveloped parcels over 3,000 square feet in size (in the year 2000) were considered available for single-family residential development. Due to computational memory constraints (on a standard office PC, with 1GB of RAM), a 10% random sample was drawn from all 16,750 developable parcels. Figure 1 depicts the study area, the undeveloped parcels and the 10% sample. The distribution of existing single-family residential parcel sizes in Austin resembles a Chi-square distribution and large, undeveloped parcels were assumed to subdivide according to this distribution. Of course, not all subdivided parcels will be occupied by newly-added households; only the chosen sites were assumed to be developed into single-family residential land after the housing market reaches equilibrium. To simulate home size, a floor-area ratio (FAR) of 0.25 was used[5]. The newly generated single-family residential sites, defined by home size, parcel-specific unit price per interior square foot, and distances to employment sites and shopping centers, were allocated to individual households based on rent-maximizing and utility-maximizing principles.

On the demand side, the 7,600 future households consist of five types, categorized by annual income levels (based on standard Census class weighted average): \$11,000, \$28,000, \$42,000, \$72,000, and \$170,000. The new households were assumed to be demographically distributed according to the 2002 American Community Survey (ACS)[6]. Corresponding to a 10% random sampling of undeveloped parcels and a 25% population growth assumption, the numbers of households to be allocated (for each of the five types) are 1500, 1200, 1200, 2300, and 1400, respectively. Three VOTT scenarios were designed to examine the impact of how VOTT may affect spatial allocation of residences. The low, medium and high VOTTs for each of the five household types were assumed to be as follows: (1) Low VOTTs: \$1.40/hour, \$3.50, \$5.30, \$9.00 and \$10.60; (2) medium VOTTs: \$2.80/hour, \$7.00, \$10.50, \$18.00 and \$21.30; and (3) high VOTTs: \$5.50, \$14.00, \$21.00, \$36.00 and \$42.50, respectively[7]. These households compete for homes that offer them the highest utilities. Due to this competition, home prices are bid up, until the market reaches equilibrium.

Essentially, individual households are assumed to evaluate all new (single-family) residential parcels, as a function of their price, size, and site accessibility (in terms of travel costs and/or distances to employment centers and shopping malls). When a home is selected as the "best choice" by more than one household, the competition/supply-demand imbalance should increase the unit price. Following such price increases, the previous best choice becomes unaffordable or at least less preferable due to the price increase, and other, relatively more preferred homes may emerge. Via this implicit price mechanism households withdraw from competition over home sites that are experiencing high demand. Ultimately, the model presumes that land developers sell the home/home site to the highest bidder, at the market equilibrium's highest price.

2.3 Equilibration results

The market equilibrium for new home buyers (considering 10% of the presently undeveloped land in Austin) was reached in an iterative fashion. The starting home value was

[5] As an extension to this work, this global variable is being made more site-specific and random.

[6] This puts 19.1% in the first, lowest income bracket, 16.0% in the second, 15.5% in the third, 30.5% in the fourth, and 18.9% in the fifth, highest bracket.

[7] The low, medium and high VOTTs are taken to be 25%, 50% and 100% of employed members' wage (assuming one full-time employed person in the first 4 types of households and two full-time employed persons in the last type of household).

assumed to be low, at just $100 per interior/built square foot (or $25 per square foot of parcel land). Each household was assumed to consider 20 randomly selected alternative homes/home sites with specific sizes and accessibilities. IID Gumbel error terms were associated with each competing household and its set of considered alternatives. Knowing price and size, households were assumed to choose those offering the highest utilities, as defined by the location choice model. Prices rose in steps of $\$1/ft^2$ when a home was desired by more than one household. When each household finally was aligned with a single, utility-maximizing home site, each occupied house allocated to the household that tenders the highest bid. At this stage, the housing market (for new buyers/movers) is said to have reached equilibrium. In this way, Austin's single-family residential development was simulated for each of six scenarios: the three sets of VOTTs for a study area having either a single employment center (the central business district [CBD]) or multiple employment centers (each housing at least 500 jobs[8]). Figure 2 illustrates the locations of these employment centers, the CBD, and the locations of the 18 shopping centers as well. The new households' working members were assumed to be allocated job sites according to the scenario (i.e., either all worked at the CBD or at sites nearest to their chosen homes).

In each simulation, the average equilibrium unit price for each (large/subdivided) parcel was computed by averaging the unit prices of the occupied pieces subdivided from the parcel, and average occupant income was calculated as the average annual income of households who choose to reside on the parcel. Exhibit 3 plots the average equilibrium unit price against the distance to the CBD or to the nearest employment center, depending on the scenario setup. As expected, the resulting plots illuminate how undeveloped parcels located close to employment sites enjoy higher average equilibrium unit prices. When VOTTs are low, there is no clear relationship between the average equilibrium unit price and the distance or travel time to employment sites. As VOTTs increase, the average equilibrium unit prices near employment sites rise and the average equilibrium unit prices far away from employment sites decline. This tendency is more significant for single-employment-site (i.e., monocentric job) scenarios than the corresponding multiple-employment-sites scenarios. Moreover, for the six scenarios, Moran's I statistics (calculated based on an inverse Euclidean-distance matrix [see, e.g., Lee and Wong 2000]) indicate that average equilibrium unit prices for residentially developed parcels have positive spatial autocorrelation over the entire region, confirming the visual information conveyed by the plots. Using Moran's statistics, one also observes a clustering of households of similar income, as expected.

3. Conclusions

This paper developed a model for distributing new households and tracking home price fluctuations, based on microeconomic theories and microsimulation. Disaggregate spatial data facilitatedmodel calibration and application for Austin, Texas, a medium-sized urban region. The results are reasonable and tangible. Perhaps most importantly, they suggest that microsimulation of an entire region's land market is viable. The model used here can be improved through more realistic developer tendencies of parcels (rather than, for example, a single-valued FAR or solely single-family residential parcels) and consideration of additional policy tools (such as roadway pricing and land regulation effects). Such approaches herald a new wave of land use modeling opportunities.

---

[8] There were 114 such employment centers located within the study area in the year 2000.

References

Alonso, W. (1964) *Location and Land Use.* Harvard University Press, Cambridge, Massachusetts.

Anas, A., Xu, R. (1999) Congestion, land use, and job dispersion: a general equilibrium model. *Journal of Urban Economics* 45, 451-473.

Berechman, J., Small, K. (1988) Modeling land use and transportation: an interpretive review for growth areas. *Environment and Planning A* 20, 1285-1309

Bina, M., Warburg, V., Kockelman, K. (2006) Location Choice vis-à-vis Transportation: The Case of Apartment Dwellers. Forthcoming in *Transportation Research Record*.

Chang, J., Mackett, R, (2005) A bi-level model of the relationship between transport and residential location. *Transportation Research Part B 40*, 123-146.

Clark, W. A. V., Huang, Y. Withers, S. (2003) Does commuting distance matter? Commuting tolerance and residential change. *Regional Science and Urban Economics* 33, 199-221.

de la Barra, T. (1989) *Integrated Land Use and Transport Modelling: Decision Chains and Hierarchies*. Cambridge University Press, New York.

Herbert, J., Stevens, B. (1960) A model of the distribution of residential activity in urban areas. *Journal of Regional Science* 2, 21-36

Irwin, E., Bockstael, N. (2004) Land use externalities, open space preservation, and urban sprawl. *Regional Science and Urban Economics* 34, 705-725.

Lee, J., Wong, D. (2000) *Statistical Analysis with Arcview GIS*. John Wiley & Sons, Inc, New York.

Rouwendal, J., Meijer, E. (2001) Preferences for housing, jobs, and commuting: a mixed logit analysis. *Journal of Regional Science* 41, 475-505.

Senior, M., Wilson, A. (1974) Explorations and syntheses of linear programming and spatial interation models of residential location. *Geographical Analysis* 6, 209-238.

Tillema, T., Ettema, D., van Wee, B. (2006) Road pricing and (re)location decisions of households. Paper presented at the *85th Annual Meeting of the Transportation Research Board*, Washington D.C.

Wingo, L. (1961) *Transportation and Urban Land Use*. Johns Hopkins University Press, Baltimore, Maryland.

Van Ommeren, J., Rietveld, P., Nijkamp, P. (1999) Job moving, residential moving, and commuting: a search perspective. *Journal of Urban Economics* 46, 230-253.

Zhou, B., Kockelman, K. (2005) Neighborhood impacts on land use change: a multinomial logit model of spatial relationships. Presented at the 52nd Annual North American Meeting of the Regional Science Association International, Las Vegas, Nevada.

Table 1. Location Choice Model Results

Explanatory Variables	**Coefficients**	**t-statistics**
Constant	-2.59	-15.5
Home price divided by Household Income	0.171	1.71
(Home Price divided by Household Income)^2	-0.0509	-4.04
Total interior quare footage (ft^2)	0.262	4.46
Euclidean commute distance (miles)	-0.0643	-7.86
Commute cost (dollars)	-0.0208	-4.66
Euclidean distance to the nearest shopping mall (miles)	0.121	6.28
Log Likelihood Values		
Market Shares	-2293.0	
Convergence	-2437.8	
LRI	0.0594	
Number of Observations	614[9]	

[9] While the original survey contains 965 records, the number of observations available for analysis here is just 614, due to missing data on workplace location (and/or selected home attributes, such as home price).

Figure 1. Map of Study Area, Showing All Undeveloped Parcels and 10% Sample

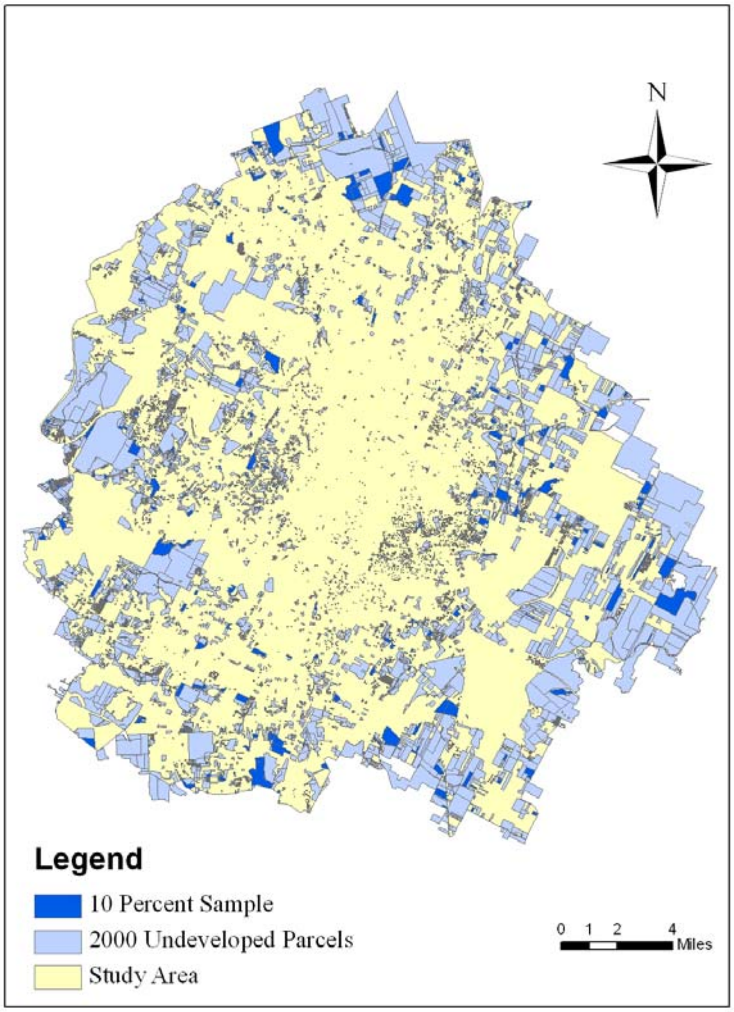

Figure 2. Locations of Austin's Employment Centers, Central Business District and Shopping Centers

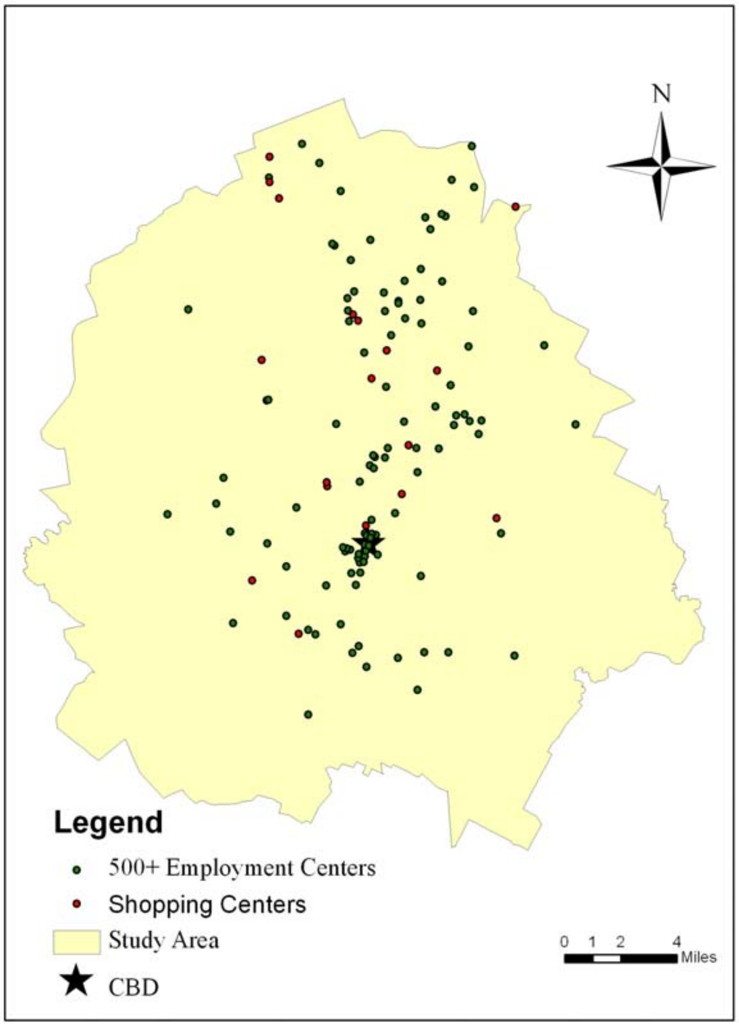

Figure 3. Equilibration Results