

# A behavioral simulation model for metropolitan policy analysis and planning: residential location and housing market components of UrbanSim

Paul Waddell

Daniel J Evans School of Public Affairs and Department of Urban Design and Planning,  
University of Washington, Seattle, Washington 98195, USA; e-mail: [pwaddell@u.washington.edu](mailto:pwaddell@u.washington.edu)  
Received 28 April 1999; in revised form 16 August 1999

**Abstract.** The growing demand for better coordination of metropolitan land use and transportation planning has led to the need for new analytical tools to examine the potential impacts of land-use and transportation policies and investments. In this paper I report on the residential and market-clearing components of the recently developed UrbanSim land-use model, which is currently being implemented in Hawaii, Oregon, and Utah. The model is based on parcel-level land-use data and simulates the interaction between demand and supply of real estate by using a dynamic behavioral approach that operates on an annual time schedule. The model system is now operational, and in this paper the design and calibration results for the Eugene-Springfield, Oregon metropolitan area are described.

## 1 Policy and planning context

Over the past several years we have observed increased public interest in mitigating the consequences of urban sprawl and making more systematic metropolitan plans that encompass land-use, transportation, and environmental dimensions. The connections between these domains are now readily apparent to the casual observer, and the public increasingly expects policies and plans in these arenas to be made consistent in shaping a sustainable path of urban development. Unfortunately the analytical tools available to planners and policymakers have lagged far behind the public will to develop coherent metropolitan plans and policies, leaving decisionmakers inadequately prepared to anticipate the potential consequences of alternative policy initiatives or infrastructure choices.

In fact, the policy creation process at times leads to rather contradictory, or at least poorly integrated, policy objectives. More often than not, the market effects and constraints of policy initiatives are given insufficient attention or are misunderstood. Overarching these concerns is the increasing fragmentation and polarization of public opinion about the basic goals of urban public policy over issues such as environmental preservation, economic development, quality of life, and social equity. Increasing public scrutiny of the political and technical decisionmaking processes for land use, transportation, and the environment adds further pressure on policymakers to develop open, understandable, systematic, and objective approaches to devising and choosing between major policy strategies.

It is in this politically charged environment that we turn our attention to the evolving role of urban land-use and transportation models. Whereas land-use models, in particular, received substantial criticism in the 1960s for their overly ambitious claims and poor performance by Lee (1973) and others, transportation models have become a standard part of the regional transportation planning process over the past three decades, as documented by Weiner (1997). In 1989 the role of land-use and transportation models in the regional planning process was abruptly, and permanently, changed. The Sierra Club Legal Defense Fund and the Citizens for a Better Environment filed a lawsuit in the Federal District Court of Northern California claiming that the

State of California, the Metropolitan Transportation Commission of San Francisco, and other agencies had violated the Clean Air Act Amendments of 1977 by not adequately addressing clean air standards (Garrett and Wachs, 1996). The focus of the case was on the failure of the planning process to account for the effects of increasing highway capacity on increasing air pollution emissions by reducing transit use, discouraging in-fill and densification, increasing highway speeds, inducing highway travel, and encouraging urban sprawl (Harvey and Deakin, 1991).

In the wake of this lawsuit and the subsequent Clean Air Act Amendments of 1991 and the Intermodal Surface Transportation Efficiency Act of 1990, states and metropolitan planning organizations have been given clear and urgent mandates to integrate land-use, transportation, and environmental planning, and to incorporate maximum public participation in their decisionmaking processes. The role of models as an objective and systematic core of the analytical basis for this activity is no longer in question. The question has shifted to how to improve and integrate land-use, transportation, and environmental models, and how to integrate the technical modeling process with an essential public participation process. With the continuing threat of legal challenge hovering, the question that remains for most planning agencies is when can this be done.

It is within this context that the UrbanSim model has been developed. UrbanSim is a land-use model developed for use by metropolitan planning organizations (MPOs). Its development was shaped by several constraints emerging from the context described above. First, it needed to integrate with existing travel models in use by MPOs in order to be implemented in the short term. Second, it needed to provide a direct capacity to analyze the impact of a range of growth management initiatives ranging from urban growth boundaries to comprehensive land-use plans promoting nodal development patterns and mixed land use. Third, it needed to be developed in a way that would make it adaptable to other metropolitan areas of varying size and complexity, and to future integration with activity-based travel models now under development. And finally, it needed to be developed and tested within one metropolitan area, by use of mostly existing data. Eugene-Springfield was chosen as the site for the initial calibration and testing of the model.

Given this background, the development of the model was framed by theoretical, empirical, and practical constraints and requirements. It needed to be theoretically sound and behaviorally transparent for the purpose of communicating with policy-makers and the public. It needed to integrate with existing travel models, rather than build a unified land-use and travel model. It required an endogenous land-market component to satisfy theoretical and policy analytic requirements such as the effects of urban growth boundaries on housing prices and densities. And finally, it needed explicitly to incorporate public policies embodied in land-use and transportation plans, including factors such as environmental constraints on development.

The organization of the remainder of this paper is as follows. In section 2 the model is placed within the context of existing research in land-use modeling. In section 3 an overview of the key components of the model system is given. In section 4 the specification and estimation results for Eugene-Springfield, Oregon are described, and section 5 concludes the paper with final observations and recommendations for further work.

## **2 Land-use modeling context**

Reviews of the literature in urban land-use modeling are in ample supply, with contributions by Anas (1987), Harris (1985), Kain (1987), Paulley and Webster (1991), Southworth (1995), and Wegener (1994; 1995), among others. In this paper I will not

therefore attempt a systematic review of the literature, but present the UrbanSim model within a context of existing work in land-use modeling.

As pointed out by Anas (1982), the concept of 'bid rent' forms the basis of modern urban economic analysis and provides the foundation for most microeconomic models of urban spatial structure. The concept originated from von Thünen's work on agricultural land use and became the centerpiece of Alonso's (1964) models of residential location and urban housing markets. In this bidding approach, in the microeconomic tradition, it is assumed that perfectly informed and efficient consumers make bids on all properties. Property owners, also fully informed, auction the property to the highest bidder, with steady-state long-run equilibrium reached when this bid-auctioning process succeeds in allocating each property to the highest bidder. All consumers are expected to change residence costlessly and instantaneously when any change in conditions occurs that temporarily disrupts equilibrium, by repeating this bid auctioning process. The monocentric model and its extensions by Alonso (1964), Mills (1967), and Muth (1969) adopted the constraint that all employment was located in the CBD, and focused on the problem of predicting residential location as a function of transportation and housing costs.

Substantial research activity evolved in the direction of predicting housing prices and the willingness-to-pay of consumers for the underlying attributes of housing. This body of work, which draws on Lancaster's (1966) theory of consumer behavior, views housing as a bundle of services, and households as utility-maximizing consumers based on some function of these underlying attributes of housing, including locational characteristics. Rosen (1974) developed the hedonic theory of housing markets, in which households choose housing so as to maximize a utility function subject to a budget constraint. Extensions of this body of theory and empirical estimation of hedonic price functions have been widely published, beginning with contributions by Strazheim (1974), Wheaton (1977), and Galster (1977).

A separate stream of research activity emerged in parallel, focusing not on prices but on household residential choice. In particular, the work of McFadden (1978; 1984), Quigley (1976; 1983), and others on the use of random utility theory to develop multinomial logit models of residential location opened a significant direction for research in this area. This approach was used by Lerman (1977) to assess the importance of accessibility and mode on residential location, and by Williams (1979) to examine the effects of neighborhood on location choice.

Some research has emerged that crosses these two streams, notably work by Ellickson (1981) that develops a logit model of the property auction process by using the bid rent function rather than the utility function. Essentially this approach focuses on the landowners's problem of selling to the highest bidder, which is the consumer making the highest bid. It differs from the majority of logit models of residential choice, which focus on the consumer's problem of choosing among properties based on maximizing their utility function. These approaches represent the two sides of the auction: the buyer's perspective, and the seller's. Martinez (1992) extended Ellickson's work by developing a 'bid-choice' model that dealt with both sides of the auction simultaneously, through a formulation in which the higher level of the model represented the consumer's choice among properties and the lower level represented the landowner's choice among bidders. Under equilibrium assumptions Martinez showed the consistency of these approaches.

At the core of the model developed by Martinez is a formulation based on consumer surplus, defined as the willingness-to-pay for an alternative less the market price of that alternative. It has a simple and intuitive interpretation: consumers are happiest with an alternative that maximizes the difference between what they are willing to pay and what they must pay based on the market price.

Martinez (1992) derives a multinomial logit model predicting the probability that a consumer  $h$  will choose lot  $i$ :

$$P_{i|h} = \frac{\exp[\mu(\Theta_{hi} - p_i)]}{\sum_j \exp[\mu(\Theta_{hj} - p_j)]}, \quad (1)$$

where

$\Theta_{hi}$  is the willingness of consumer  $h$  to pay for lot  $i$ ,

$p_i$  is the market price of lot  $i$ ,

$\mu$  is a positive scale parameter.

The probability of choosing alternative  $i$  then is a function of the relative consumer surplus ( $S$ ) of the alternatives:

$$S_{hi} = \Theta_{hi} - p_i. \quad (2)$$

Martinez (1992) adopts an equilibrium formulation in which the market price is endogenous and determined by the highest bid for each site among all consumers. This interpretation is founded on the view of land as a quasi-unique commodity in fixed supply, so that demand dictates price. It does not, apparently, represent the supply of buildings, with either short-term or long-term adjustment in supply interacting with demand to influence prices.

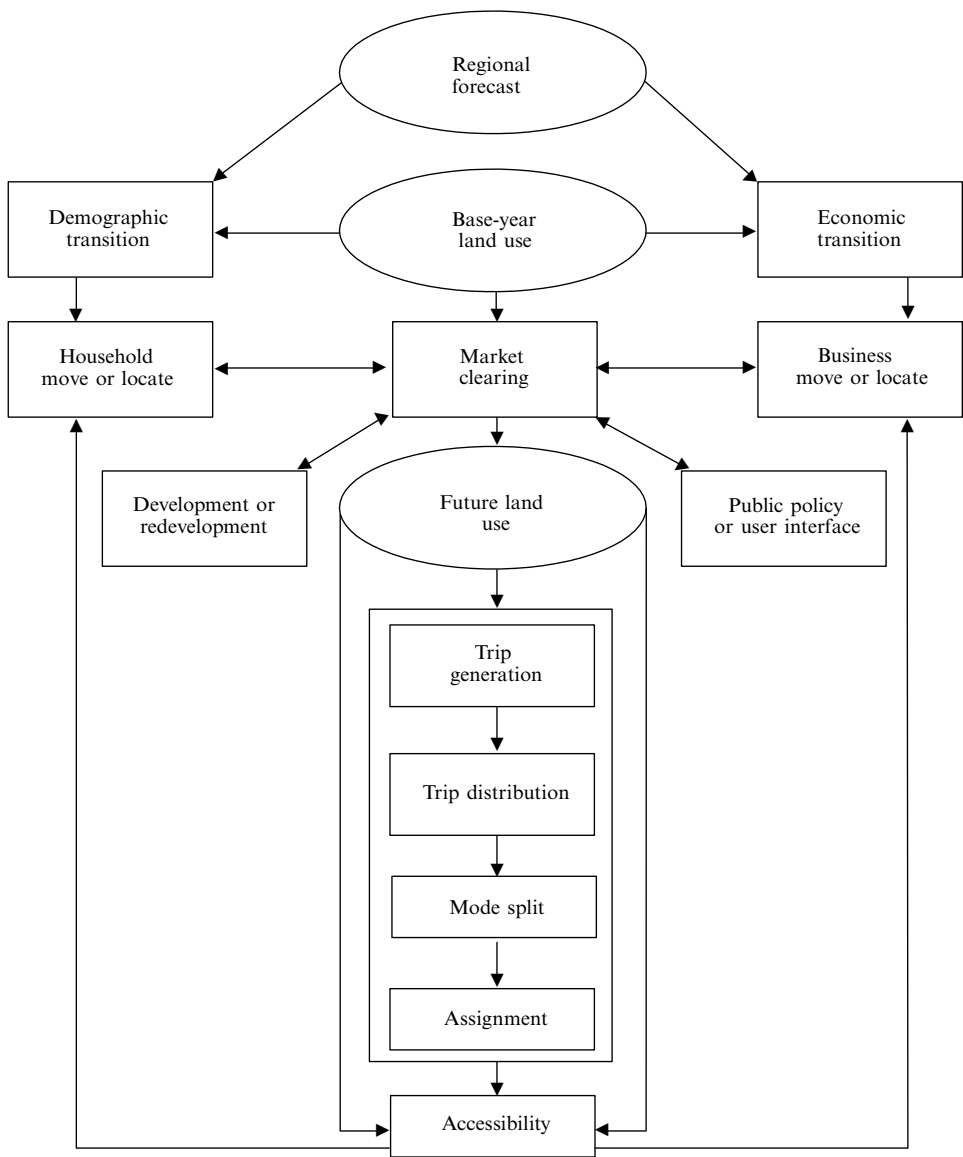
A third relevant line of research in residential location, originating in geography and sociology, is based on residential mobility. These models include work that focuses on the household characteristics and on dissatisfaction, or push factors, inducing mobility. Research in this vein includes that of Wolpert (1965), work by Brown and Moore (1970) on separating the decision to move and the decision to search, and by Speare et al (1975). Economists formalized these models as disequilibrium models of housing expenditure, for example by Hanushek and Quigley (1978). Onaka (1983) has formulated a variation of the housing disequilibrium model by using hedonic theory. More recent work has linked the mobility and location choice approaches. Van Lierop and Rima (1982), Onaka and Clark (1983), Clark and Onaka (1985), and Kain and Apgar (1985) have developed two-stage models of the decision to move and the choice of location.

### 3 UrbanSim overview

The UrbanSim model is based on the view of urban development over time and space as the composite outcome of the interactions of individual choices and actions taken by households, businesses, developers, and governments. The structure of this model includes components modeling the behavior of these decisionmaking units interfaced through the land market. As Wegener (1995) has pointed out, different urban processes take place on different time scales, with household and business location occurring over time periods of one year or less, construction of buildings occurring over slightly longer periods, and infrastructure planning and construction over much longer time frames. This model attempts to reflect this insight in its design by modeling urban development in a series of one-year steps.

The flowchart in figure 1 presents a graphical view of the key components of the model system, each of which is described briefly in this section, followed by a more detailed treatment of the mobility and location choice components that lie at the core of the model system. The model components represent the behavior of household, businesses, developers, and governments, all interfaced through the land market.

Exogenous inputs to the model include base-year land use, population and employment, regional economic forecasts, transportation system plans, land-use plans, and



**Figure 1.** UrbanSim Model structure.

land-development policies such as density constraints, environmental constraints, and development-impact fees. The user interacts with the model to create scenarios that combine alternative packages of assumptions and exogenous inputs. The model is then executed using a given scenario, and the results of one or more scenarios can be examined and compared in a GIS.

Two modules, demographic change and economic transition, predict changes in the distribution of households and businesses by type (for example, age, income, businesses by industry) at the regional level, consistent with exogenous aggregate forecasts of population and employment. The model endogenously predicts the location of businesses and households; the location, type, and quantity of new construction and redevelopment by developers; and the prices of land and buildings.

In the household mobility and location module, the model simulates household decisions about whether to move or remain in the current residence, and, if they choose to move, their selection of a housing type and zone. In the business mobility and location module, businesses make similar choices regarding mobility, building type, and location choice. Household and business characteristics influence choices, as do locational attributes such as accessibility and prices.

In the land-development component, the model simulates developer choices to convert vacant or developed land to urban uses, including the type of improvements and density, based on their profitability expectations and subject to constraints imposed by governmental policies such as zoning and infrastructure availability. These profitability expectations are influenced by prior prices and revealed demand in the location and building-type preferences of businesses and households.

The model simulates land-market clearing by adjusting prices to reconcile the competing demands for locations and structures among households and businesses against the supply of space in each zone. The ratio of demand to supply in each zone for each type of space (housing and commercial structures by type) induces proportional price adjustments for these structures. The adjusted prices produce new market signals to demanders in the subsequent year, thereby influencing preferences for zones and building types.

These interactions of households, businesses, developers, and governments produce outcomes representing the distribution of population and employment, as well as the prices, uses, and density of land development. These results are written out for any desired year for which the travel models will be run. The data are fed into the traditional four-step travel models to produce new travel times, costs, and patterns by mode. The analysis then uses these travel times to compute new accessibility indices in subsequent years, until the travel models are run for the next target year.

In the following sections a more detailed description of the residential and business location components of the model, and estimation results for Eugene-Springfield are provided. There is, however, insufficient space for a full elaboration of each of the model components and the results of their estimation in this paper. For a more complete description of the balance of the model, see Waddell (1998a; 1998b; 1998c).

### **3.1 Mobility and location choice**

The location model we propose relates in key respects to the bid component of the model developed by Martinez (1992), based on consumer surplus [see equation (1)]. It diverges from that model in significant ways, but most notably by dropping the equilibrium assumption. I have assumed that market prices are exogenous to the location choice faced by the individual consumer, and instead resolve price adjustments through an aggregate market-clearing process external to the choices made by individual consumers. Time is structured in annual steps, within each of which mobility and location choices are accounted for, as well as space construction decisions. Between years I will account for price adjustment and new construction and redevelopment. Rather than impose the equilibrium assumption that households make costless and instantaneous adjustments in residential location, I follow in the tradition of research that identifies a two-stage process of residential location, consisting of the decision to move, and the choice of location and dwelling type.

#### **3.1.1 Aggregation of alternatives**

As we are not dealing with elemental housing units or lots as the level of choice, but instead aggregating alternatives into zones (traffic analysis zones) and land-use types, we need to adjust the choice equation to account for the aggregate nature of the choice set. I will do this by including the size of the choice set represented by each of the

aggregate choices. Equation (1) can then be rewritten to include the size term, and equation (2) can be substituted into equation (1):

$$P_{i|h} = \frac{\exp[\mu^*(S_{hi} + \mu' \ln S_i)]}{\sum_j \exp[\mu^*(S_{hj} + \mu' \ln S_j)]}, \quad (3)$$

where  $\mu^*$  is a positive scale parameter,  $\mu' = \mu^*/\mu$ , and  $0 \leq \mu' \leq 1$  (Ben-Akiva and Lerman, 1987, page 259). This is the form of the location-choice model estimated.

Note that the model has a simple specification and direct interpretation. The consumer surplus is the difference, in dollars, between what a consumer is willing to bid for an alternative and the market price of the alternative. The log of the size term accounts for the aggregation of the elemental alternatives into zones of arbitrary size.

### 3.1.2 The bid function

Whereas the urban economic literature generally refers to a willingness-to-pay function that defines a hypothetical maximum a consumer would be willing to pay for an alternative, I suggest that the use of empirical data to estimate this function will reveal a function more aptly described as a bid function. What is actually observed when examining market outcomes of consumer choices is the prices they paid for the alternatives chosen. They might have been willing to pay more than they were required to according to the market price for the alternative but they would undoubtedly be reluctant to reveal their true maximum willingness-to-pay for it. What we observe, then, in examining actual market outcomes are the successful bids made by particular consumers, which match the market price for the alternative. This is consistent with the interpretation of consumers as price takers in the market.

This interpretation is summarized for the chosen alternative as the bid matching the market price and bounded by the maximum willingness-to-pay:

$$B_{hi} = P_i \leq \Theta_{hi}. \quad (4)$$

In order to compare alternatives for predicting the probability that a consumer will select a particular one, however, we need to estimate a bid function that encapsulates consumer preferences for the attributes of the alternatives. The bid function must predict what consumers would have been willing to bid for each of the alternatives in the choice set based on their attributes.

The bid function proposed here differs from the typical hedonic regression in that consumers are stratified into market segments, and a separate bid function for each is estimated. This allows the identification of bid price functions specific to subgroups of consumers, whose tastes vary significantly across incomes and stages of life cycle. A hedonic regression cannot be interpreted as a demand function because it represents the envelope of demand and supply curves of all consumers and suppliers in the market.

I will estimate bid functions for households stratified by income level and by the presence of children, a key life-cycle characteristic. Once the bid functions have been estimated, the bid equation is used to generate bids for each of the alternatives in the choice set, to estimate the consumer surplus for each alternative, and ultimately to predict the location-choice probability.

## 3.2 Accessibility: the land-use – transportation link

Before developing the specification of the model in further detail, I will address briefly the linkage of land use and transportation through accessibility, in order to bring this into the discussion of model specification that follows. As this model is not of the monocentric or spatial interaction genre, in which the choice of workplace is exogenous

and residential locations are chosen on the basis principally of commute to the city center or to a predetermined workplace, we must deal with access in a more general framework. Accessibility is considered a normal good, like other positive attributes of housing, on which consumers place a positive economic value. We would therefore expect that consumers value access to workplaces and shopping opportunities, among the many other attributes they consider in their housing preferences. Nor would all households respond to accessibility in the same way. For instance, retired persons would be less influenced by accessibility to job opportunities than would working-age households.

We operationalize the concept of accessibility for a given location as the distribution of opportunities weighted by the utility of travel to those destinations. The utility of travel is operationalized as the composite utility across all modes of travel for each zone pair, obtained as the log-sum of the mode choice for each origin–destination pair (Williams, 1977). The resulting access measure for each location, is then:

$$\text{Access}_i = \sum_j A_j \exp(\beta L_{ij}), \quad (5)$$

where

$A_j$  is the quantity of activity in location  $j$ ,

$L_{ij}$  is the composite utility, or log-sum (for one-car households), from location  $i$  to  $j$ ,

$\beta$  is a scaling parameter.

The scaling parameter can be used to amplify or attenuate the contribution of the activity in more accessible, higher utility destinations, thus varying the scale of the access measurement.

### 3.3 Market clearing

The land-market clearing is a component of the model system that reconciles demand for housing with the available housing supply in every year. The housing stock is assumed fixed within a single year, although it will adjust in the longer term. This model component assigns moving households to their highest utility alternative that is available and adjusts housing prices according to the ratio of demand to supply in each submarket. As prices enter the location-choice utility functions for households, an adjustment in prices will alter their location preferences, causing higher price alternatives to become more likely to be chosen by occupants that have lower price elasticity of demand, all else being equal. Similarly, any adjustment in housing prices alters the preferences of developers to build new construction by type of housing, and the density of the construction.

Once households have evaluated all the available alternatives, and expressed their preferences (through a probability prediction from the location-choice models), the simulation attempts to place households into vacant housing in proportion to their predicted probabilities. Housing that is occupied during this operation is removed from the remainder of the allocation process, and households that are unable to locate into their highest utility building are forced to accept their next highest utility alternative. This process iterates until all households are located in houses<sup>(1)</sup>.

The market-clearing mechanism, then, is not strictly through a full equilibrium price adjustment, in which perfect information exists, and transaction costs are zero, so that prices on all buildings at each location adjust to the equilibrium solution that clears the market. Instead the solution is based on an expectation of incomplete information and nontrivial transactions and search costs, so that movers obtain the

<sup>(1)</sup> At present there is no accommodation in the design for a homelessness outcome, although this refinement would be feasible.

highest satisfactory location that is available, and prices respond at the end of the year to the balance of demand and supply at each location.

Once the market assignment is completed, the information generated by the market simulation about the relative demand and supply of each building type at each location is used to update prices. I make the following assumptions.

- (1) Households and businesses respond to prices of housing and nonresidential space at the start of the current year, if they are moving in that year.
- (2) Developers respond to current market prices, after they have been adjusted to account for demand and supply in the current year.
- (3) Households, businesses, and developers are all price takers, and market adjustments are made by the market in response to aggregate demand and supply relationships.
- (4) Location preferences are capitalized into land values. Building value reflects building replacement costs only and can include variations in development costs owing to terrain, environmental constraints, or development policy. The market price per housing unit or per square foot of nonresidential space is the sum of land and building value.
- (5) There is a normal vacancy rate, above which the market triggers an upward adjustment in land prices and below which it triggers a decline (for an elaboration of this approach to real-estate market price adjustment, see DiPasquale and Wheaton, 1996).

Based on these assumptions, which are consistent with urban economic theory, the price adjustment mechanism causes an adjustment in the location preferences of businesses and households in the following year. Developers respond to current prices by maximizing the profitability of construction, subject to the constraints of land supply and development policy. The supply of housing and commercial space consumed in any iteration comes from existing vacant structures plus any new construction and redevelopment of structures that occurred in the most recent period. New construction in a forecast interval can include committed, proposed, and potential development projects identified by the user as exogenous policy input.

The form of the price adjustment is

$$P_{lbt} = P_{lbt-1} \left[ \frac{(1 + \alpha_b - V_{lbt}) + \lambda(1 + \alpha_b - V_{bt})}{1 + \lambda} \right]^\beta, \quad (6)$$

where

$P_{lbt}$  is the land price of building type  $b$  in location  $l$  in year  $t$ ,

$P_{lbt-1}$  is the previous year closing land price for the same building and location,

$V_{lbt}$  is the current vacancy rate for space in the building type  $b$  and location  $l$ ,

$\alpha_b$  is the normal vacancy rate for building type  $b$ ,

$\beta$  is a scaling parameter for the price adjustment,

$\lambda$  is a parameter for weighting the regional and zonal influence.

The price adjustment is based on the relationship between the current vacancy rate and a normal or structural vacancy rate at which prices would remain relatively stable. The local submarket and the aggregate metropolitan market conditions both contribute to the resulting price shifts and may reinforce or attenuate each other.

#### 4 The data

The area for this study encompasses the transportation planning area of the Lane Council of Governments (LCOG) in Oregon. The planning area lies towards the southern end of the Willamette Valley, on the Interstate 5 corridor. Like other metropolitan areas in Oregon, Eugene-Springfield is experiencing rapid growth and attempting to manage this growth effectively within the constraints of state and local policy initiatives such as the Urban Growth Boundary. The planning area consists of 271 traffic analysis zones, which form the locational choices for the model.

**Table 1.** Household classification categories.

| Income (\$)   | Age of head | Household size | Children | Mobility (years) |
|---------------|-------------|----------------|----------|------------------|
| < 10 000      | < 29        | 1              | 0        | ≤ 5              |
| 10 000–24 999 | 30–49       | 2              | ≥ 1      | > 5              |
| 25 000–49 999 | 50–64       | 3              |          |                  |
| ≥ 50 000      | ≥ 65        | ≥ 4            |          |                  |

Households are classified by using five characteristics—household income, age of household head, household size, presence of children, and mobility—with categories as in table 1.

The data produce 111 unique nonzero household types from the cross-classification of these categories. For estimation of the probability that a household moved within the past five years, all household types were used. For estimation of location choice, income and presence of children were used to stratify households.

The development of the 1994 base-year household data at the zone level was developed by using a ‘synthetic baseline population’ method developed by Beckman et al (1995) as part of the TRANSIMS project. This approach uses iterative proportional fitting to generate a joint distribution of household characteristics with the required geographic detail, by linking the US Census Public Use Microdata Sample (5%) with marginal distributions of household characteristics at the census-block group level from Summary Tape File 3A. Iterative proportional fitting is applied to the household weights in the PUMS sample to estimate the joint distributions at a census-block group level that are consistent with the geographically detailed marginal distributions and with the aggregate joint distribution from the 5% PUMS household sample.

These joint distributions are converted from census-block group to the zones in the model by using parcel data. As the census-block group and traffic zone can both be identified for every ownership parcel, the parcels with housing were used to develop an allocation from block group to traffic zone. In the allocation procedure it is assumed that the households in a sample within a block group have a probability of location within a traffic zone that is proportional to the quantity of housing of the type occupied by the household. This process is implemented by using a sampling procedure that also scales the household weights from 1990 to the model base year of 1994, on the assumption that the distribution of household types remained consistent over that time.

The housing inventory and prices originate from the Lane County parcel database. Assessed values are used as surrogates for market prices in the base year, because sales data would be insufficient to populate values for each zone and housing type. In Oregon, properties are assessed at full market value, though errors in the assessed values clearly add noise to the estimation of the bid functions. The advantage of this procedure is that it uses available data sources so it can be readily replicated. It also provides a large sample size for estimation of the residential mobility rates and location and housing-type choice components of the model. The alternative of using household travel surveys to estimate the location-choice model would be preferred in metropolitan areas having surveys with sufficient sample size. Similarly, an alternative to assessed values would be preferred for the market price estimates, but the results discussed in the following section suggest that, in spite of the inevitable data measurement errors from the data described here, the results of estimation appear reasonable and robust.

---

## 5 Model specification and estimation results

The prototype model has now been specified and calibrated for Eugene-Springfield, Oregon. In the interest of space, I will focus on the residential location-choice component of the model. The specification of the bid functions and results from the estimation for households are reviewed below.

### 5.1 Household bid function

The variables included in the household bid function, and their theoretical justification, are discussed below. The variables are drawn from the literature in urban economics, urban geography, and urban sociology. An initial feature of the model specification is the incorporation of the classical urban economic trade-off between transportation and land cost. This has been generalized to account not only for travel time to the classical monocentric center, the CBD, but also for access to employment opportunities and to shopping. These accessibilities to work and shopping are measured by weighting the opportunities at each destination zone with a composite utility of travel across all modes to the destination, based on the log-sum from the mode choice travel model. The scale parameter in equation (5) is set to 1 for employment access, and to 4 for shopping access, to reflect more the localized influence of shopping destinations and to reduce the correlation between shopping employment and total employment in the access measures.

These accessibility measures should negate the traditional pull of the CBD and, for some population segments, potentially reverse it. In addition to these accessibility variables, I have included in the model a net building density to measure the input-substitution effect of land and capital. To the extent that land near high-accessibility locations is bid up in price, we should expect that builders would substitute capital for land and build at higher densities. Consumers for whom land is a more important amenity will choose larger lot housing with less accessibility, and the converse should hold for households that value accessibility more than land, such as higher income childless households.

The age of housing is considered for two reasons. First, we should expect that housing depreciates with age, because the expected life of a building is finite and a consistent stream of maintenance investments are required to slow the deterioration of the structure once it is built. Second, owing to changing architectural styles, amenities, and tastes, we should expect that the wealthiest households prefer newer housing, all else being equal.

A related hypothesis from urban economics is that, because housing is considered a normal good, it therefore has a positive income elasticity of demand. This implies that, as incomes rise, households will spend a portion of the gains in income to purchase housing that is more expensive, and which provides more amenities (structural and neighborhood) than their prior dwelling. A similar hypothesis is articulated in urban sociology (for example, see Massey and Mullen, 1984), in which upward social mobility is associated with spatial proximity to higher status households. Both of these hypotheses would predict that households of any given income level would prefer, all else being equal, to locate in neighborhoods that have a higher income mix.

The age hypothesis and the two income-related hypotheses are consistent with the housing filtering model, which explains the dynamic of new housing construction for wealthy households that sets in motion a chain of vacancies. The vacancy chain causes households to move into higher status neighborhoods than the ones they leave, and housing units to be occupied successively by lower status occupants. At the end of the vacancy chain, in the least desirable housing stock and the least desirable neighborhoods, there is insufficient demand to sustain the housing stock, and vacancies go

unsatisfied, leading ultimately to housing abandonment. I have included in the model an age-depreciation variable, along with a set of variables of neighborhood income composition, to test collectively the housing filtering and related hypotheses.

Housing type is included in the model as a set of dummy variables for alternative housing types, with single-family housing excluded as a base of comparison. Residential housing with two to four units (duplex, triplex, and quadplex), and multifamily housing, are expected to be discounted significantly in bid prices because they are likely to have smaller living spaces and fewer amenities than single-family housing.

Given the stratification of households into consumer market segments based on income and the presence of children, I have tested for compositional effects not only of income, but also of the proportion of households with children in a neighborhood. It is expected that households with children would be willing to pay more for a house in a neighborhood with more children, all else being equal, than a similar-income household without children.

Among the amenities that households prefer are attributes of the land-use mix within the neighborhood. It is likely that residential land use, as a proxy for land uses that are compatible with residential use, positively influences housing bids. On the other hand, industrial land use, as a proxy for less desirable land-use characteristics, would lower bids.

Finally, we test for market supply effects on bids by including the size of the housing stock in the neighborhood. If we were truly observing the maximum willingness-to-pay, then a supply effect should be present. Because we expect to observe a bid that matches the market price, and is below the willingness-to-pay, then higher market supply might well be expected to lower the market price, and therefore the bid. This market supply effect is different from the other effects included in the model because it does not represent an attribute that influences demand. We therefore use the variable in the estimation step to remove the potential bias on other variables that might be encountered by excluding it. In the subsequent step in which we use the fitted bid-price function to predict bids, however, this term is dropped. This will generate bids that respond to underlying consumer preferences and leave the market price adjustment external to the formation of bids. The variable definitions are presented in table 2, and

**Table 2.** Household bid-price variables.

| Variable            | Definitions                                                                                |
|---------------------|--------------------------------------------------------------------------------------------|
| Residential 2–4     | Dummy variable for housing type residential, with 2–4 units (duplex, triplex, or quadplex) |
| Multi-family        | Dummy variable for housing type multifamily, with 5 or more units                          |
| InAccEmployment     | Log of accessibility to total employment, with an exponent on the log-sum of 1             |
| InAcc4Retail        | Log of accessibility to retail employment, with an exponent on the log-sum of 4            |
| Density(units/acre) | The net density in units per acre of a particular housing type in a zone                   |
| LnUnits             | Log of the number of housing units of a particular type in the zone                        |
| LnAge               | Log of the average age of the buildings of a type in a zone                                |
| PctIncome1          | Percentage of households in a zone in the lowest income group                              |
| PctIncome2          | Percentage of households in a zone in the second lowest income group                       |
| PctIncome4          | Percentage of households in a zone in the highest income group                             |
| PctChild            | Percentage of the households in a zone that have one or more children                      |
| PctLandIndustrial   | Percentage of the developed land in the zone that is in industrial use                     |
| PctLandResidential  | Percentage of the developed land in a zone that is in residual use                         |
| TimeToCBD           | Travel time to the CBD, in minutes                                                         |

**Table 3.** Household bid-price functions.

|                                | Income group 1       |                      | Income group 2       |                      | Income group 3       |                      | Income group 4       |                      |
|--------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
|                                | child                | no child             | child                | no child             | child                | no child             | child                | no child             |
| Intercept                      | 0.8661<br>(0.759)    | 5.0712<br>(5.170)    | 0.8730<br>(0.680)    | 3.3327<br>(3.227)    | 4.9405<br>(7.243)    | 5.5321<br>(8.915)    | 7.2061<br>(3.197)    | 4.1270<br>(1.163)    |
| Residential 2–4                | –0.4297<br>(–15.306) | –0.4815<br>(–19.994) | –0.4507<br>(–15.398) | –0.5162<br>(–22.857) | –0.4179<br>(–9.869)  | –0.5848<br>(–30.252) |                      | –0.7776<br>(–6.979)  |
| Multi-family                   | –1.0447<br>(–27.161) | –1.0189<br>(–41.734) | –1.1022<br>(–20.692) | –1.0711<br>(–33.888) | –1.1099<br>(–9.894)  | –1.3607<br>(–35.988) | –0.9856<br>(–4.170)  | –1.8252<br>(–20.402) |
| LnAccEmployment                | 0.8089<br>(8.081)    | 0.4326<br>(5.045)    | 0.7912<br>(6.845)    | 0.6155<br>(6.694)    | 0.4593<br>(7.167)    | 0.4198<br>(7.059)    | 0.3329<br>(1.619)    | 0.6557<br>(2.020)    |
| LnAcc4Retail                   |                      | 0.0307<br>(3.427)    | 0.0077<br>(0.931)    | 0.0075<br>(0.954)    | 0.0347<br>(3.983)    | 0.0087<br>(0.999)    | 0.0955<br>(5.799)    | 0.0403<br>(1.537)    |
| Density(units/acre)            | –0.0063<br>(–5.488)  | –0.0053<br>(–9.786)  | –0.0091<br>(–4.530)  | –0.0061<br>(–6.963)  | –0.0310<br>(–7.605)  | –0.0018<br>(–1.557)  | –0.0128<br>(–2.700)  | –0.0020<br>(–0.592)  |
| LnUnits                        | –0.0016<br>(–0.131)  | –0.0270<br>(–2.850)  | –0.0056<br>(–0.532)  | –0.0271<br>(–3.104)  | –0.0011<br>(–0.094)  |                      | –0.1389<br>(–6.510)  | –0.0351<br>(–1.202)  |
| LnAge                          | –0.0929<br>(–3.775)  | –0.0789<br>(–4.484)  | –0.0791<br>(–3.229)  | –0.1416<br>(–7.672)  | –0.1789<br>(–7.008)  | –0.1770<br>(–8.486)  | –0.3714<br>(–9.261)  | –0.5391<br>(–9.385)  |
| PctIncome1                     | –0.0131<br>(–12.603) | –0.0096<br>(–10.906) | –0.0142<br>(–15.290) | –0.0102<br>(–11.828) | –0.0123<br>(–12.604) | –0.0107<br>(–11.475) | –0.0103<br>(–4.937)  | –0.0064<br>(–1.893)  |
| PctIncome2                     | –0.0145<br>(–12.006) | –0.0161<br>(–14.971) | –0.0157<br>(–14.684) | –0.0140<br>(–14.921) | –0.0166<br>(–14.184) | –0.0137<br>(–13.167) | –0.0178<br>(–10.116) | –0.0084<br>(–3.419)  |
| PctIncome4                     | 0.0045<br>(1.642)    | 0.0098<br>(4.505)    | 0.0075<br>(3.600)    | 0.0157<br>(8.996)    | 0.0030<br>(1.738)    | 0.0080<br>(4.924)    | 0.0019<br>(0.861)    | 0.0094<br>(2.803)    |
| PctChild                       | 0.0032<br>(3.445)    | 0.0026<br>(3.211)    | 0.0059<br>(7.029)    | 0.0027<br>(3.376)    | 0.0014<br>(1.501)    | –0.0020<br>(–2.050)  | –0.0061<br>(–3.863)  | –0.0144<br>(–5.798)  |
| PctLandIndustrial              | –0.0057<br>(–6.241)  | –0.0108<br>(–15.680) | –0.0051<br>(–6.077)  | –0.0097<br>(–14.347) | –0.0029<br>(–2.852)  | –0.0057<br>(–5.292)  |                      | –0.0188<br>(–4.819)  |
| PctLandResidential             | 0.0011<br>(1.732)    |                      | 0.0027<br>(4.944)    |                      | 0.0053<br>(9.612)    | 0.0031<br>(5.870)    | 0.0099<br>(12.323)   | 0.0039<br>(2.971)    |
| TimeToCBD                      | 0.0204<br>(2.642)    | 0.0126<br>(2.006)    | 0.0187<br>(2.552)    | 0.0129<br>(2.033)    |                      |                      | 0.0160<br>(1.377)    | 0.0181<br>(0.964)    |
| <i>N</i>                       | 1352                 | 2432                 | 1681                 | 2878                 | 1413                 | 2085                 | 438                  | 508                  |
| Adjusted <i>R</i> <sup>2</sup> | 0.76                 | 0.79                 | 0.73                 | 0.73                 | 0.69                 | 0.74                 | 0.76                 | 0.73                 |

Note: figures in parentheses denote *t*-statistics

the results of estimation of the bid-price functions of households stratified by market segments by income level and age of children in table 3.

These results provide strong support for the core hypotheses described in this section. Taken together, the model explains approximately three fourths of the variation in the bid prices of these household market segments. In addition, the segmentation of the market by income and the presence of children is borne out by informal comparison of coefficients across market segments and is substantiated by more rigorous Chow tests.

In almost all cases, the results had the expected signs and were significant at the 5% level. Compared with single-family housing, we find consistent discounting of bids for duplex–quadplex and multifamily housing. Density, the inverse of lot size,

is consistently negative and significant in all but two cases. Access to employment is consistently significant and positive, as is access to shopping.<sup>(2)</sup> Age depreciation was reflected in lower bids, with a trend that higher income households discounted for age more heavily than lower income households. A consistent pattern also emerged from the income composition of the neighborhood, with all household types discounting for higher percentages of the two lowest income levels, and increasing bids for neighborhoods with greater concentration in the highest income level. This pattern of coefficients supports the housing filtering hypothesis, and the underlying conversion of upward social mobility into residential location choices in more affluent neighborhoods. The wealthiest households will tend to outbid for the newest housing with the highest concentration of wealth, and lower income households will move up through the vacancy chain as affluent households move into new housing.

After accounting for the access to employment and shopping, proximity to the CBD is discounted in bids for most household types. This discounting perhaps reflects negative externalities associated with proximity to the center, such as localized congestion, noise, school quality, or crime, or other factors not included in the model.

The relative preference of households with children for neighborhoods with higher proportions of households that have children appears to be generally supported, but with an interesting trend across income levels. Lower income households tend to bid higher for neighborhoods with higher proportions of children, whereas the highest income households tend to discount their bids somewhat at a higher concentration of households with children.

The relative bids for this attribute within each income level suggest consistency in households with children bidding higher than those without children, but the income trend remains. One potential explanation for this trend is that neighborhoods with a high proportion of households with children are a fairly middle-class phenomenon, which represents an attraction for lower income households, and a deterrent for the highest income households.

The land-use mix in a neighborhood was found to be significant and reasonably consistent, with a higher percentage of the developed land in a zone in residential use prompting higher bids, and more industrial land associated with discounting. The preference for residential land, and its associated characteristics, increased with income. The distaste for industrial land appeared to be more significant for households with children than for those without.

There were some exceptions to the consistency of the results. The market supply effect, measured by the log of housing units, was negative in all but one case, but varied in significance across household groups. This variable was dropped from the one case in which it was positive. Similarly, exceptional signs were found on residential land for childless low-income and low-middle-income households, which might reflect their preference for multifamily housing, but for policy analytical purposes a negative coefficient on residential land is problematic and was dropped. Travel time to the CBD was dropped for income group 3, in which it was insignificant, and for childless households, where it was negative.

## 5.2 Logit estimation of residential location choice and housing type

The fitted bid functions described in the preceding section were used to generate estimates of the consumer surplus of each alternative and were included in a much simpler logit specification that included only the consumer surplus estimate and the

<sup>(2)</sup> Access to shopping was measured with a value of 4 on the exponent of the log-sum composite utility of travel, compared with an exponent of 1 for employment access. This reflects more localized accessibility to shopping opportunities.

**Table 4.** Logit estimation results for residential consumer surplus.

| Variable                             | Coefficient           | Z     |
|--------------------------------------|-----------------------|-------|
| Consumer Surplus                     | $4.73 \times 10^{-5}$ | 6.748 |
| LnSize                               | $2.30 \times 10^{-2}$ | 4.156 |
| Log-likelihood estimated:            | -70 047               |       |
| Log-likelihood with no coefficients: | -81 592               |       |

log of the size variable. As the estimation of consumer surplus for all households makes them directly comparable, and it is expected that households of all types compete in the housing market according to their consumer surplus, all households are pooled for the logit estimation of location choice. The results are shown in table 4.

These results provide confirmation of the basic structure of the model, with expected signs and significance on the two terms. The greater the consumer preference, the more likely a consumer is to choose the option. The larger the number of underlying alternatives within an aggregate choice (zone), the more likely a consumer is to choose it, with the consumer surplus of the alternatives being held constant.

**6 Conclusions and recommendations**

I have presented only the residential location and market-clearing components of the UrbanSim model here owing to space limitations. The full model system has been calibrated and implemented in software. Other papers will follow that describe the remaining components of the model and their implementation results in more detail. The partial results reported here provide a valuable test of the work to date, and a foundation for the further development of this modeling strategy. A historical validation exercise is now underway to learn more about the ability of the full model system to capture key influences on location choices and prices over the period 1980–94.

Future directions for development of the model include its application in other metropolitan areas of varying sizes and conditions, and a series of specific extensions of the model and its software implementation. One priority for future research is to increase the resolution of location choices by using a grid representation of land that allows better measurement of localized spatial context. This development will be important to improving the sensitivity of the model to urban design scale characteristics that may influence nonmotorized and transit mode share and residential location choices.

A second priority for development is a microsimulation implementation of the location choice and market-clearing component of the model (the land-development component is already based on microsimulation at the parcel level). Although the current model does not attempt to link the residence and workplace of the individual workers within the household it may be worthwhile, or necessary, to attempt this complication within a microsimulation implementation. A microsimulation model with linkages between job and residence changes, especially in households with multiple workers, presents a significant challenge for specification and calibration. The microsimulation approach and grid representation of location set the stage for the development of more direct linkages to environmental models to integrate better the analysis of urban development with the assessment of the environmental impacts of development alternatives. Alberti and Waddell (1999a; 1999b) have developed a general framework for linking this model to a human stressor model to integrate urban and ecosystem modeling and strategic analysis.

**Acknowledgements.** The development of the UrbanSim model has been supported by the Oahu Metropolitan Planning Organization, the Oregon Department of Transportation, the State of Utah Governor's Office, the National Cooperative Highway Research Board [NCHRP 8-32(3): Integration of Land Use Planning and Multimodal Transportation Planning], The University of Washington PRISM Project, and the National Science Foundation (NSF CMS-9818378: Reusable Modeling Components for Land Use, Transportation and Land Cover). This model development effort has been informed by feedback from peer review committees and collaborators in each of these projects. In particular, the comments of Rick Donnelly, John Douglas Hunt, Frank Koppelman, Keith Lawton, Jim Ryan, David Simmonds, Michael Wegener, and other have contributed to improvements in the model design.

## References

- Alberti M, Waddell P, 1999a, "An integrated urban development and ecological model. Part I: Rationale for a synthesis", in *Computers in Urban Planning and Urban Management* (Franco Angeli, Milan); full versions of paper are on CD-ROM available with the book
- Alberti M, Waddell P, 1999b, "An integrated urban development and ecological model. Part II: The strategy", in *Computers in Urban Planning and Urban Management* (Franco Angeli, Milan); full versions of paper are on CD-ROM available with the book
- Alonso W, 1964 *Location and Land Use* (Harvard University Press, Cambridge, MA)
- Anas A, 1982 *Residential Location Markets and Urban Transportation: Economic Theory, Econometrics, and Policy Analysis with Discrete Choice Models* (Academic Press, New York)
- Anas A, 1987 *Modeling in Regional and Urban Economics* (Harwood Academic, New York)
- Beckman R J, Baggerly K A, McKay M D, 1995, "Creating synthetic baseline populations" *Transportation Research, Part A: Policy and Practice*
- Ben-Akiva M, Lerman S, 1987 *Discrete Choice Analysis: Theory and Application to Travel Demand* (MIT Press, Cambridge, MA)
- Brown L A, Moore E G, 1970, "The intraurban migration process: a perspective" *Geografiska Annaler, Series B* 1–13
- Clark W A V, Onaka J, 1985, "An empirical test of a joint model of residential mobility and housing choice" *Environment and Planning A* 17 915–930
- Ellickson B, 1981, "An alternative test of the hedonic theory of housing markets" *Journal of Urban Economics* 9 56–79
- Galster G C, 1977, "A bid-rent analysis of housing market discrimination" *American Economic Review* 67(2) 144–155
- Garrett M, Wachs M, 1996 *Transportation Planning on Trial: The Clean Air Act and Travel Forecasting* (Sage, Thousand Oaks, CA)
- Hanushek E, Quigley J, 1978, "An explicit model of residential mobility" *Land Economics* 54 411–429
- Harris B, 1985, "Urban simulation models in regional science" *Journal of Regional Science* 25 545–567
- Harvey G, Deakin E, 1991 *Toward Improved Regional Transportation Modeling Practice* Clean Air Project, National Association of Regional Councils, 1700 K Street, N.W., Suite 1300, Washington, DC 20006
- Kain J F, 1987, "Computer simulation models of urban location", in *Handbook of Regional and Urban Economics* volume 2, Ed. E S Mills (Elsevier, Amsterdam) pp 847–875
- Kain J F, Apgar W C, 1987 *Housing and Neighborhood Dynamics: A Simulation Study* (Harvard University Press, Cambridge, MA)
- Lancaster K J, 1966, "A new approach to consumer theory" *Journal of Political Economy* 74 132–157
- Lee D B, 1973, "Requiem for large-scale models" *Journal of the American Institute of Planners* 39 163–178
- Lee D B, 1994, "Retrospective on large-scale urban models" *Journal of the American Planning Association* 60 35–40
- Lerman S R, 1977, "Location, housing, automobile ownership and mode to work, a joint choice model" *Transportation Research Record* number 610, 6–11
- McFadden D, 1978, "Modelling the choice of residential location in spatial interaction theory and planning models", in *Spatial Interaction Theory and Planning Models* Eds A Karlqvist, L Lundqvist, F Snickars, J W Weibull (North-Holland, Amsterdam) pp 75–96
- McFadden D, 1984, "Econometric analysis of qualitative response models", in *Handbook of Econometrics* volume 2, Eds Z Griliches, M Intriligator (North-Holland, Amsterdam) pp 1395–1457

- 
- Martinez F, 1992, "The bid-choice land-use model: an integrated economic framework" *Environment and Planning A* **24** 871–885
- Massey D, Mullen B, 1984, "Processes of Hispanic and Black spatial assimilation" *American Journal of Sociology* **89** 836–873
- Mills E S, 1967, "An aggregative model of resource allocation in a metropolitan area" *American Economic Review* **57** 197–210
- Muth R F, 1969 *Cities and Housing* (University of Chicago Press, Chicago, IL)
- Onaka J, 1983, "A multi-attribute housing consumption disequilibrium model of residential mobility" *Environment and Planning A* **15** 751–765
- Onaka J, Clark W A V, 1983, "A disaggregate model of residential mobility and housing choice" *Geographical Analysis* **15** 287–304
- Paulley N J, Webster F V, 1991, "Overview of and international study to compare models and evaluate land-use and transport policies" *Transport Reviews* **11** 197–222
- Quigley J M, 1976, "Housing demand in the short-run: an analysis of polytomous choice" *Explorations in Economic Research* **3** 76–102
- Quigley J M, 1983, "Estimates of a more general model of housing choice", in *The Urban Economy and Housing* Ed. R E Grieson (Lexington Books, Lexington, MA) pp 125–141
- Rosen S, 1974, "Hedonic prices and implicit markets: product differentiation in pure competition" *Journal of Political Economy* **82** 34–55
- Southworth F, 1995, "A technical review of urban land use–transportation models as tools for evaluating vehicle reduction strategies", report to US Department of Energy; copy available from Oak Ridge National Laboratory, PO Box 2008, Oak Ridge, TN 37831
- Speare A, Goldstein S, Frey W, 1975 *Residential Mobility and Migration and Metropolitan Change* (Ballinger, Cambridge, MA)
- Strazheim M R, 1974 *An Econometric Analysis of the Urban Housing Market* (National Bureau of Economic Research, New York)
- Van Lierop W F J, Rima A, 1982, "Towards an operational disaggregate model of choice for the housing market", second interim report, The Free University, Amsterdam
- Waddell P A, 1998a, "Development and calibration of the prototype metropolitan land use model", final report to the Oregon Department of Transportation; copy available from the author
- Waddell P A, 1998b, "UrbanSim: the Oregon prototype metropolitan land use model", in *Transportation, Land Use and Air Quality: Making the Connection: Proceedings of a Speciality Conference* Eds S Easa, D Samdahl (American Society of Civil Engineers, Reston, VA) pp 549–558
- Waddell P A, 1998c, "Exploiting parcel level GIS in land use models", in *Transportation, Land Use and Air Quality: Making the Connection: Proceedings of a Speciality Conference* Eds S Easa, D Samdahl (American Society of Civil Engineers, Reston, VA) pp 632–641
- Wegener M, 1994, "Operational urban models: state of the art" *Journal of the American Planning Association* **60**(1) 17–30
- Wegener M, 1995, "Current and future land use models", publication DOT-T-96-09, US Department of Transportation, Federal Highway Administration, 400 Seventh Street, S.W., Washington, DC 20590
- Wheaton W C, 1977, "A bid rent approach to housing demand" *Journal of Urban Economics* **4** 200–217
- Weiner E, 1997 *Urban Transportation Planning in the United States: An Historical Overview* publication DOT-T-97-24, US Department of Transportation, Washington, DC
- Williams H C L W, 1977, "On the formation of travel demand models and economic evaluation measures of user benefit" *Environment and Planning A* **9** 285–344
- Williams R C, 1979, "A logit model of demand for neighborhood", in *The Economics of Neighborhood* Ed. D Segal (Academic Press, New York) pp 17–42
- Wolpert J, 1965, "Behavioral aspects of the decision to migrate" *Papers and Proceedings of the Regional Science Association* **15** 159–169

