A Behavioral Analysis of Household's Choices on Housing Consumption and Travel Mode

Dan Li[1]*, T.R. Lakshmanan[2], Chun-Yu Ho[3], W.P. Anderson[4]
Boston University

First Version: November 2004
Second Version: Nov 2005
Current Version: Oct 2007

Abstract

A behavioral model is developed for an integrated analysis of housing and travel mode choices. A household maximizes its utility by trading off non-housing goods consumption, housing consumption and travel time subject to both budget and time constraints. We further link the theoretical model with empirical evidence on the households' housing and mode choices in Boston, 1991. Our empirical results confirm the theoretical implication that both housing and mode choices are determined largely by income.

Keywords: Housing; Travel mode; Income; Boston

JEL Code: R12, R20, C3

1* Corresponding author. Email: danli@bu.edu; Tel: +1 (617) 453-4158; Fax: +1 (617) 353-4449; URL: http://people.bu.edu/danli; Address: Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215.
2 Email: laksh@bu.edu; URL: http://www.bu.edu/transportation/laksh.html; Address: Center for Transportation Studies, Boston University, 675 Commonwealth Avenue, Rm. 439, Boston, MA 02215.
3 Email: chunyu@bu.edu; URL: http://people.bu.edu/chunyu; Address: Department of Economics, Boston University, 270 Bay State Road, Boston, MA 02215.
4 Email: bander@bu.edu; URL: http://www.bu.edu/transportation/anderson.html; Address: Center for Transportation Studies, Boston University, 675 Commonwealth Avenue, Rm. 439, Boston, MA 02215.

"The outcome of the city will depend on the race between the automobile and the elevator, and anyone who bets on the elevator is crazy." Frank Lloyd Wright, Architect.

1. Introduction

Households choose a location that maximizes their utility subject to budget and time constraints. Given housing as normal goods, the wealthier a household become the more spacious and higher-quality house they demand. In a closed-city model, higher demand for housing will bid up the land price throughout the city, which causes the developers to substitute away from land, resulting in higher structural densities. Therefore, cities with limited land, such as Hong Kong and Tokyo, feature by skyscrapers, high density per square kilometers and astonishingly high unit housing prices. An open-city model gives us a different view of urban growth, as excess demand for land will result in "spatial expansion" and the middle and upper classes flee the city center to the suburbs to enjoy spacious houses, and, possible, a better neighborhood. As Mills and Lubuele (1997) point out, central city problems have led people to leave and seek solace in bucolic, socially controlled suburbs with their more attractive social milieu.

Obviously, "flight from plight" can not be the only cause of suburbanization. Without the income growth of households and the technological progress in transportation, it is not feasible for them to migrate from the city center. Margo (1992) shows that one-half of the increased suburbanization between 1950 and 1980 in America can be explained by rising income. Glaeser and Kahn (2004) strongly argue that the urban sprawl is an inexorable product of car-based living, and Kopecky and Suen (2005) show the declining automobile prices account for increased automobile ownership, and thus suburbanization for the period 1910 to 1970 in U.S.

The aim of this paper is to develop a behavioral model to analyze the interdependence among housing choice, household income, the demand for speed and travel time in an open-city circumstance. We build a structural model and employ

Multinomial Logit Model (MNL) and Ordered Logit Model (OLM) to test the implication of the theoretical model by using Boston Household Survey data for 1991. The empirical results confirm the implications of our theoretical model, that people with higher income level tend to live further away from Central Business District (CBD), enjoy more spacious housing and travel with a speedier mode.

Our structural model developed is inspired by the model in Lakshmanan and Hua (1983) -- an open functional model which maximizes personal total net utility (utility minus disutility) derived from the consumption of a composite product, the inherent time required for consumption activities and the searching time. The implication of the model is that, given a fixed substitution effect, high income people tend to spend more time searching for cheap and high quality products in more distant places and they also endogenize the speed choice in order to lower travel cost. Our model is also close to Desalvo and Huq (2005)'s open functional equilibrium urban model including the mode and leisure choices.

We contribute to the literature on urban model by setting up a closed-form structural model which incorporates endogenous choice for speed and housing. To our knowledge, the treatment of speed as endogenous is absent in existing models and our treatment proposes a new aspect of explaining consumer behavior in a spatial-temporal context. We also contribute by linking the theory with the empirical evidence, which has yet not been done in the urban models including both housing and mode choices.

The plan of the paper is as follows. In the next section, a closed structural model is presented. In section 3, 1991 Boston Household Survey data is described. Section 4 discusses estimation methodologies, and empirical results are presented in section 5. The last section provides concluding remarks.

2. Theoretical Model

Households maximize a standard utility function which is given by:

$$U(c,h,l) = \log c + \frac{h^{1-\theta}}{1-\theta} + \frac{l^{1-\gamma}}{1-\gamma} \tag{1}$$

They derive utility from a composite of consumption (c), housing (h), and leisure (l). The marginal rates of substitution between consumption and housing and leisure are controlled by the parameters θ and γ respectively. The total time endowment of the household is divided among work, leisure and travel. We treat working hour t_w as fixed since a job usually requires workers show up regularly and on time no matter how far away they live. The time constraint is:

$$l = (24 - t_w) - 2t = t_d - 2t \tag{2}$$

in which $t_d \equiv 24 - t_w$ is the disposable time left after working and t is the one-way travel time. For simplicity, we only consider home-based work trips and assume people go to work and then come back once. More time spent on travel means less leisure time left.

Define Y as the income of the representative household, which consists of wage income and non-wage income where w is the wage rate and z is the non-wage income.

$$Y = wt_w + z \tag{3}$$

Since housing stock is a normal good, when people get richer, they demand a higher-quality and more spacious house. House price gradient theory, which says that real estate price falls with distances from the CBD, applies here. Therefore, housing price $q(s)$ is a decreasing function in distance s.

$$q = \frac{f}{s^{\alpha}} \tag{4}$$

where f is the representative price/square foot of houses in the CBD and $q_s \equiv \partial q/\partial s < 0$.

Locomotion is characterized by three attributes: travel distance, speed and travel time, which are related by the identity: distance = travel time * speed.

$$s = tv \tag{5}$$

where v is the speed demand or estimated speed of the household. Here, we do not treat speed as exogenous. From walking to omnibus (a horse-drawn vehicle) to street car to automobile, every improvement in speed arises from technological progress in transportation and the provision of public transportation infrastructure, which are exogenous to individual household decisions. However, given these exogenous factors, together with the different monetary costs and travel speed associated with different modes, households will choose the most suitable speed to maximize their utility according to their own constraints. From this point of view, speed is endogenous --- it is the consumer's choice rather than given by exogenous force. Moreover, even though each individual's speed choice has no direct effect on the speed supply, the collective consumer demand for speed does impact the supply of speed. To a certain extent, the technological changes and public infrastructure supply in transportation reflect people's increasing demand for speed.

In this paper, a mode is defined in terms of its speed v, which is an endogenous variable. We use the unit monetary cost of travel advocated by Lakshmanan and Hua (1983). The monetary cost per mile $r(s,v)$ is a function of distance and speed which decreases in distance and increases in speed. It includes both fixed cost and variable cost. The unit fixed cost is decreasing in s because the fixed investment cost per mile declines in the distance traveled. And the variable cost is increasing in v because the higher speed cause higher fuel consumption per mile.

$$r = \frac{k}{s^{\eta}} + v^{\beta} \tag{6}$$

where k is the fixed cost such as the payment for a car, monthly parking cost, car insurance and so on. The assumption implies $r_s < 0$ and $r_v > 0$. Apparently, the choice for speed has two effects on travel cost: on one hand, higher speed has a positive effect on traveler's monetary cost; on the other hand, given a fixed distance, higher speed reduces total travel time. Combining equations (3), (4), (5) and (6), we obtain modified budget constraint (7). Total income is equal to the sum of non-housing consumption expenditure, housing expenditure and monetary travel cost where the price index for non-housing consumption is denoted as p .

$$\begin{aligned} Y &= wt_w + z \\ &= pc + qh + 2rs \\ &= pc + \frac{f}{(tv)^{\alpha}} h + 2[k(tv)^{1-\eta} + tv^{1+\beta}] \end{aligned} \tag{7}$$

The household maximization problem is reduced to the following

$$Max_{c,h,t,v} \log c + \frac{h^{1-\theta}}{1-\theta} + \frac{(t_d - 2t)^{1-\gamma}}{1-\gamma} \tag{8}$$

$$s.t.\ \ Y = pc + \frac{f}{(tv)^{\alpha}} h + 2[k(tv)^{1-\eta} + tv^{1+\beta}] \tag{9}$$

where t enters the utility function as a disutility which means households sacrifice leisure for travel. The willingness to pay for avoiding more travel time is the same as the willingness to pay for more leisure. All parameters α, β, γ, θ and η are assumed to be positive. Since the objective is strictly concave, the optimal solution is in the interior and characterized by the following Kuhn-Tucker conditions:

$$\frac{MU_c}{P_c} = \frac{MU_h}{P_h} = \frac{MU_t}{P_t} \tag{10}$$

The optimality conditions indicate that the marginal utilities of all activities (consumption of non-housing goods, consumption of housing, consumption of leisure) are equalized and there is no reallocation of choices to increase utility. Following (10), we derive the following optimal conditions.

Non-housing and Housing Consumption tradeoff

$$\frac{1/c}{p} = \frac{1/h^{\theta}}{f/(tv)^{\alpha}} \quad \Rightarrow \quad fh^{\theta} = pc(tv)^{\alpha} \tag{11}$$

The optimal allocation between non-housing and housing consumptions should equalize the marginal utility per unit price derived from these two activities. Agents cannot obtain higher utility by reallocating the consumption level of non-housing and housing goods.

Non-housing Consumption and Travel Time tradeoff

$$\frac{1/c}{p} = \frac{2\gamma/(t_d - 2t)}{\alpha fh/t^{\alpha+1}v^{\alpha} - 2[(1-\eta)kt^{-\eta}v^{1-\eta} + v^{1+\beta}]}$$

$$\Rightarrow \quad pc = \frac{(t_d - 2t)^{\gamma}}{2}\{\frac{\alpha f}{t^{\alpha+1}v^{\alpha}}h - 2[(1-\eta)kt^{-\eta}v^{1-\eta} + v^{1+\beta}]\} \tag{12}$$

The marginal utility from non-housing consumption is equal to the marginal utility obtained from more leisure time saved from less travel. Namely, the optimality requires that the amount of utility that the agent gives up per unit of non-housing consumption should be equal to the amount of utility obtained from reduced travel time and monetary travel cost.

Optimal Speed Choice (OS) condition:

$$\frac{\alpha f}{t^{\alpha}v^{\alpha+1}}h = 2[(1-\eta)kt^{1-\eta}v^{-\eta} + (1+\beta)tv^{\beta}]$$

$$\Rightarrow \quad fh = \frac{t^{\alpha+1}v^{\alpha}}{\alpha}\{2[(1-\eta)kt^{-\eta}v^{1-\eta} + (1+\beta)v^{1+\beta}]\} \tag{13}$$

The optimal speed choice condition indicates that the marginal gain from lower housing price due to living further away from CBD is equal to the change in travel monetary cost and time cost. The individual's choice for optimal mode is a trade-off between the housing price and the total travel cost.

The previous three optimality conditions (Equations 11-13) and the budget constraint (Equation 9) determine the value of the decision variables c, h, t and v. If we assume consumption as a constant proportion of household income, i.e. pc = ϕ Y, we obtain the following reduced form equations for housing choice h and mode choice v.

$$\log h = -\frac{1}{\theta}\log\frac{f}{\phi} + \frac{1}{\theta}\log Y + \frac{\alpha}{\theta}\log s \tag{14}$$

$$\log v = \frac{1}{1+\beta}\log(\frac{\phi}{\beta}) + \frac{1}{1+\beta}\log Y - \frac{\gamma}{1+\beta}\log(t_d - 2t) \tag{15}$$

The system of equations indicates that households will choose larger houses h, speedier transportation mode v, when they become wealthier Y and the housing stock h is positively correlated with the distance s. In addition, the housing choice depends on the price of housing stock. However, our data is only one year household survey for Boston in 1991, so there is no cross-section and time-series variation in the representative housing price in the CBD. Therefore, the residual is absorbed into the constant term in equation (14). The difficulties, encountered here in directly estimating the parameters in equation (14) and (15), consist of both identification problem (β can not be identified) and data constraint. The data constraints include that there is no direct data on the distance; the household income is available in six categories rather than continuous; and speed is not observable.

However, a plausible way is to test the predictions of our model indirectly is via the following propositions:

Proposition 1: Ceteris paribus, a high-income-level household chooses more spacious housing;

Proposition 2: Ceteris paribus, a high-income individual tends to use speedier mode.

3. Data

This study uses cross-sectional data. The data source is the 1991 Boston Household Survey conducted by the Boston Metropolitan Planning Organization (MPO). This household survey consists of three main files: a household file, a personal file and a trip file. They provide information on household characteristics (household income, household size, household location, number of children over five years old, number of vehicles, parking type, dwelling type and etc.), personal characteristics (gender, birth year, employment status, location of job, ability for driving and so on), and daily trip (purpose of trips, origins and destinations of trips, travel time, travel mode). There are 3,854 household records, 9,281 personal records and 39,373 trip records.

For the structural model, the main assumptions are that there is a mono-centric city where most economic activities are located in CBD and the trip purpose for visiting CBD is for work. As in any large metropolitan areas, Boston does have a distinct CBD, called Downtown Boston, which is devoted primarily to banking, finance and commerce. Although the Boston core provides just fewer than 30% of the region's total employment, it still has the largest concentration of employment in its MPO region. Hence, despite the fact that the assumption of mono-centric city is not literally valid, it should nevertheless be a useful working assumption. Historically, downtown transportation studies have focused on the area known as Boston Proper, which is bordered by Massachusetts Avenue, Charles River, Boston Harbor and Fort Point Channel (The census tracts most closely approximating downtown area are 105 - 305 and 701 - 712).

The purposes of trips made by each person are many and various. However, trips to work are mandatory and essential for every household and hence, when they make choices on mode and housing, the convenience and ease of making working trips will be given priority consideration. Focusing on Home-based-work (HBW) trips is a plausible, but simplified, way to study the household behavior in choosing mode and housing. Therefore, we restrict the data sample to those people who make HBW trips to the downtown area and our sample is reduced to include around 600 trip records. The variables, that we are interested, are summarized in Table 1.

4. Estimation Methodology

In this paper, we use Multinomial Logit Model (MNL) and Ordered Logit Model (OLM) to test the propositions of housing choice and mode choice, respectively. The MNL assumes that a household has unobservable, latent preferences or utilities for different housing choice (or transport mode) and they choose the housing (or mode) providing the highest utility (Ben-Akiva and Lerman, 1985). The different choices available to the household can not be ranked in MNL model – that is, the variable is measured on a nominal scale rather than on an ordinal scale. The OLM is an extension of the better-known MNL used when the choice variable takes on three or more possible values which are subject to some logical ordering.

The factors, which contribute to the housing choice decision, are diverse. Given the information on the housing type, there is no one dominating factor that can be ranked according to a logical ordering. There are four dwelling types in our data set: Apartment, Duplex, Single Family and Townhouse/Condo, which can only be measured on a nominal scale. Hence, we can not place them in a credible ordering according to any specific characteristics. For instance, since our dataset does not reveal any information on the area space (or some other logically ordered characteristics) of the housing type, we can not tell that Duplex is larger than Townhouse or Single Family is smaller than Duplex since the size for these four dwelling types are not fixed according to some normal rules when they were built.

However, for the mode choice, the most important factor is speed, which can be ranked from the lowest to the highest velocity after controlling for exogenous factors. In our dataset, there are 10 mode choices, which can be grouped into four categories according to their speed in ascending order: Non-motor, Transit, Bus and Car. Therefore, OLM is preferred to estimate mode choice.

The basic idea behind the Logit model is that consumers are trying to maximize their utility by making proper choices controlling for their characteristics. An individual n 's utility function is

$$U_n = V_n + \varepsilon_n \qquad (16)$$

where U_n = "total utility " that a consumer n derives from making a choice

V_n = systematic or "observed" utility

ε_n = random component

The V_n is defined as a linear function of attributes of the consumer:

$$V_n = \beta X_n \qquad (17)$$

where β and X_n are, respectively, a vector of parameters and a vector of independent variables. The random component ε_n is the part of utility that is unknown to the researcher but known to the consumers themselves, such as tastes and preferences.

4.1 Multinomial Logit Model for Housing Choice

As we mentioned earlier, the housing types can not be ordered, hence, the MNL is a proper methodology. MNL model is consistent with global utility-maximization. It is based on the principle that households associate a utility value with each housing type i and choose the one with maximum utility:

$$Choice = i \Leftrightarrow U_{ni} \geq Max(U_{n1}, ...U_{n,k\neq i}, ..U_{nI}) \quad \text{where } i = 0,1,2 \text{ and } 3 \qquad (18)$$

In the MNL model the random terms are assumed to be identically and independently Gumbel distributed: $F(\varepsilon_n) = e^{-e^{-\varepsilon_n}}$. [5] Given this assumption, the probability for

---

[5] This assumption makes the model easy to estimate. However, it also leads to the independence from irrelevant alternatives (IIA) property, which limits the applicability of the MNL model. The IIA property assumes that the relative probability of two existing outcomes is unaffected by the addition of a third outcome. For example, suppose that an individual's choice is initially between two different outcomes and that she is evenly split between the two. Now, suppose we add a third alternative that is nearly identical to the second. We would then expect the probability of choosing the second outcome to be split in half and the probability of choosing the first outcome to be unaffected. Unfortunately, the IIA property does not account for this, but rather splits the probabilities equally among all three alternatives in order to keep the relative probabilities of the first two options equal. Hence, in cases

choosing a specific type of housing can be defined as:

$$\Pr(Y_n = i) = P_{ni} = \frac{\exp(\beta_i' x_n)}{\sum_{i=1}^{I} \exp(\beta_i' x_n)} \quad \text{for i = 0, 1,2 and 3} \tag{19}$$

where P_{ni} is the probability that the dependent variable, which is an indicator variable that can take values from 0 to the number of possible outcomes (I), takes value i at *n*-th observation.

In the case of housing choice, we denote Apartment (i=0), Duplex (i=1), Single Family (i=2) and Townhouse/Condo (i=3). Identification imposes the normalization $\beta_0 = 0$, which implies that $\ln(P_{ni} / P_{n0}) = \beta X_n$. Then, the point estimates of a MNL tell us, for each choice i, the effect on the probability of the outcome i $\neq 0$, relative to the baseline outcome (i = 0), caused by a unit change in the explanatory variables. In a multinomial framework, this does not assure that the absolute probability of outcome i will increase or decrease, but that it will be more or less likely relative to the baseline outcome. In our case, the probability of membership in other housing categories is compared to the probability of membership in the reference category---Apartment.

The coefficient estimates in MNL do not represent the marginal effects of the independent variables as in other normal regression models do and are therefore difficult to interpret (Greene, 2003). The MNL coefficients in the housing choice represent the effect that the independent variables have on the probability of non-Apartment type housing dwelling relative to the base category of Apartment. To directly compare the marginal effects on the four different types of housing stock induced by any change in the independent variables, we have to convert the coefficient estimates into the desired marginal effects for each of the independent variable. The marginal effect δ_{ic} of the change of an independent variable x_c on

---

where two alternatives are close substitutes MNL may be inappropriate as it relies on the IIA property. Our estimation is highly likely to be free from this concern since these four accommodation types are usually not close substitutes.

the adoption probability of the type i housing is given by

$$\delta_{ic} = \frac{\partial P_i}{\partial x_c} = P_i \, [\beta_i - \sum_{i=0}^{3} P_i \, \beta_i \,] = P_i \, [\beta_i - \overline{\beta}] \tag{20}$$

P_i is the probability of choosing housing type i and δ_{ic} measures the impact of a variation of an exogenous variable x_c on the probability of choosing housing type i. As we notice, every sub-vector of β enters every marginal effect, both through the probabilities itself and through the weighted average that appears in δ_{ic}. Greene (2003) stresses that for any particular x_c, δ_{ic} does not need to have the same sign as β_{ic}.

4.2 Ordered Logit Estimation for Speed Choice

There are ten modes in the dataset, which are grouped into the four categories from lowest speed mode to the highest speed mode: Non-motor (walk and bicycling), Transit (rapid transit and commuter rail), Bus (school bus, MBTA bus, Non-MBTA bus) and Car (car/van/pickup, taxi, and ride). We index the speed for taking the Non-motor as 0, Transit 1, Bus 2 and Car 3.

The OLM assumes “local” instead of “global” utility maximization. Local utility maximization implies a choice situation in which each adjacent binary decision consists of whether to accept the current value or “take one more”. For instance, a consumer achieves her local utility maximization by taking Bus and she will stop rather than taking the alternative speedier mode--Car. The model also defines a set of “cut points” associated with each of the possible outcomes. In our case, a decision maker faces four ordered speed choices—j = 0, 1, 2, 3. Define a cut point λ_1 such that consumer n will take speed 0 if U_n is less than λ_1, or in probabilistic terms

$$P_{n0} = P_n \, (j = 0) = \Pr(\beta X_n + \varepsilon_n \le \lambda_1 \,) \tag{21A}$$

where P_{n0} is the probability that consumer n takes speed 0. The probability that the traveler takes speed 1 is defined as the probability that U_n is greater than λ_1 but less than a second cut point λ_2 :

$$P_{n1} = P_n(j=1) = \Pr(\lambda_1 < \beta X_n + \varepsilon_n \le \lambda_2) \tag{21B}$$

or, more generally,

$$P_{nj} = \Pr(\lambda_j < \beta X_n + \varepsilon_n \le \lambda_{j+1}) \quad \text{for} \quad j = 1, \dots J-1 \tag{21C}$$

$$P_{nJ} = P_n(j=J) = 1 - \Pr(\beta X_n + \varepsilon_n \le \lambda_J) \tag{21D}$$

where λ_J is the assumed highest cut point of utility. The random components ε_n are assumed to follow a standard logistic distribution $F(\varepsilon_n) = 1/[1+\exp(-\varepsilon_n)]$ and consequently, an explicit form for the probabilities can be written:

$$P_{n0} = 1/[1+\exp(\beta X_n - \lambda_1)] \tag{22A}$$

$$P_{nj} = 1/[1+\exp(\beta X_n - \lambda_{j+1})] - 1/[1+\exp(\beta X_n - \lambda_j)] \quad \text{for } j = 1, \dots J-1 \tag{22B}$$

$$P_{nJ} = 1 - 1/[1+\exp(\beta X_n - \lambda_J)] \tag{22C}$$

Estimates of β and the cut points λ s can be obtained by using the maximum likelihood method based on a set of observations making 0, 1, 2 or 3 speed choice for which the independent attributes data in X_n are available.

There are three advantages of using OLM over MNL for testing the implication of the mode choice. First, it offers a way to exploit the ordering speed information of the different modes which is the focus of our theoretical model. Second, estimation of cut points is of interest in part because it generally increases the efficiency of estimated choice probabilities relative to estimators that ignore the information contained in the threshold structure. Finally, the property that choice probabilities are necessarily between in 0 and 1 means that in prediction of the mode, the estimated value for speed choice will not be negative or infinite, which might happen in the MNL model.

As in the case of the MNL coefficients, the OLM coefficients represent the effect that the independent variables have on the likelihood of choosing one mode type versus the other mode types, but not the marginal effects. For the four probabilities of choosing different modes, the marginal effects of changes in the regressors are (Green, 1993)

$$\delta_0 = \partial P_0 / \partial x = -f(\beta' x)\beta \tag{23A}$$

$$\delta_1 = \partial P_1 / \partial x = [f(-\beta' x) - f(\lambda_1 - \beta' x)]\beta \tag{23B}$$

$$\delta_2 = \partial P_2 / \partial x = [f(\lambda_2 - \beta' x) - f(\lambda_1 - \beta' x)]\beta \tag{23C}$$

$$\delta_3 = \partial P_3 / \partial x = f(\lambda_3 - \beta' x)\beta \tag{23D}$$

where f is the derivative of the standard logistic distribution F .

5. Empirical Results

5.1 Housing Choice

The model's dependent variable is the choice probability of a dwelling type. And the explanatory variables include household income, household size and number of vehicles in the household. Household income is expected to be positively correlated with the housing stock choice in our theoretical model since higher household income relax the budget constraint and, ceteris paribus, make more spacious and higher quality house affordable. A large household size will, on one hand, exert pressure on household to choose comparatively more spacious house and on the other hand, make the household budget for housing stock tighter since bigger household size usually implies higher dependence ratio. Bread earners in the household can travel further to work thanks to the possession of vehicles, and in turn, it will induce the household to choose housing location further from CBD where the housing price is lower.

The coefficient estimates of the MNL model for housing choices are presented in Table 2. The coefficients show the effect of the explanatory variables on the marginal utility of the housing type under consideration relative to the reference---Apartment.

The statistical significance of a coefficient indicates the extent to which the corresponding explanatory variable affects the marginal utility of the relevant housing type relative to the Apartment. Estimates with the negative sign imply the preference for the Apartment and positive sign, against. To assess the simultaneous effect of the explanatory variables on the probabilities of the four housing types, one should turn to the marginal effects, which are presented in Table 3. The estimated parameters show the effect of the explanatory variables on the probability of choosing the housing type under consideration, given the value of all independent variables at their mean.

The positive and highly significant coefficients on household income in Table 2 suggest the aversion toward the Apartment over the other dwelling types. This implies that, ceteris paribus, the wealthier a household becomes the less likely it is to choose Apartment. The marginal effects bring out some interesting phenomenon (Table 3). Income is positively and significantly associated with Single Family and Townhouse/Condo, negatively and significantly with Apartment and Duplex. Given all explanatory variables at the mean level, 1% increase in income will increase the probability of choosing Single Family by 24.3% and the probability for Townhouse/Condo 9.8%, while, and decrease the probability for choosing Apartment by 33.4%. There is no effect on the probability of choosing Duplex caused by the change in income. From these estimates, we infer that Single Family is the highest--income-elastic housing type among the four and, obviously, the Apartment is an inferior good. When people are getting richer, they have a strong desire for moving into Single Family and moving away from Apartment. Single Family is usually larger than Apartment in Great Boston Area, and hence the **Proposition 1** is confirmed.

The coefficients on household size (Table 2) are positive and significant, which indicates having more family members induces a household move away from Apartment. Turning to the marginal effects in Table 3, the probability of choosing Single Family is increased by 22.1% when there is 1.0% increase in the household size. The household size has a highly negative impact on probability of choosing

Apartment. As a household becomes bigger, they tend to move into Single Family which is roomier than Apartment in Great Boston Area.

The "number of vehicles" is also a significant determinant of the preference for a given housing type (Table 2). The coefficients are significant and positive, which shows that, when compared with the reference, the ownership of more vehicles will increase the probability of a household choosing a dwelling type other than Apartment. Looking into the marginal effects (Table 3), we can see that the "number of vehicles" is the second largest force driving a household away from Apartment. For 1% increase in the number of vehicles, the probability of choosing Apartment decreases by 24.0%.

The number of vehicles is highly positively correlated with the probability of choosing Single Family. The explanation is that vehicles are essential for those who live in the Single Family, a spacious housing type that is usually located further from CBD and cheaper unit price, to make trips to work, school and other important activities. Apartments are usually located close to the CBD and the parking space is very limited. The car ownership is not favored by the limited parking space. Moreover, car might not be necessary because of the short distance to working place.[6]

The coefficient estimates for the Duplex is significant as shown in Table 2, while, the marginal effect becomes insignificant in Table 3. It might be due to the small number of observations for Duplex ---less than 6% of our sample (see Table 6). The predicted probability at its mean is very imprecise and thus makes the estimated marginal effects insignificant.

5.2 Speed Choice

OLM coefficient estimates and marginal effects of speed choice are presented in Tables 4 and 5, respectively. The dependent variables are a series of ordered indices

---

[6] We note that more vehicles in the household are due to preference. However, as shown in Table 1, average number of vehicles is less than average number of household member. Vehicle is more likely to be necessity for driving to work rather than luxury goods in our sample.

for speed from the lowest to the highest. The rank for speed is arranged according to common knowledge of the speed for the observed modes. The independent variables include personal characteristics (individual income,[7] gender and age) and household characteristics (household size and number of children under 5). The reason we use the personal income instead of household income is that we believe that mode choice is more personally determined rather than collectively determined within household. Sex and age are believed to affect consumers' taste for mode. Older people are more likely to choose a Car and females are less likely to drive. Household characteristics influence the decision making process since the family welfare is an essential part of an individual family member's utility. For example, a family may have a child less than 5 who needs a parent to take her to a kindergarten; the parent is more likely to consider driving a Car.

The coefficients for all factors except for female dummy are highly significant and positive, which indicates any positive change in the significant explanatory variables will induce the agent to switch to a higher speed. The insignificance of the female dummy implies that there is no evidence that gender plays a role in determining the choices of travel mode in our sample. The significance of the estimates for the cut points, λs , show that the assumed speed ordering and consequently the OLM is an appropriate specification for this exercise.

Turning to the marginal effects estimated in Table 5 (estimates are evaluated when all independent variables are at their mean value), the marginal effects on the female dummy are insignificant in every category. For the lowest speed category---Non-motor, the increase in personal income will make it less likely for people to choose walking or bicycling. 1% increase in income will increase the

[7] Because the dataset only records the income range for every household rather than for each person, we use the information on total household income, number of jobs taken up by family members per household and the type of jobs (full time or part time) to calculate the personal income. Assume the earning from part time job is half of the full time job within each household. We divide total household income by the number of jobs took by the workers in the household and then distribute the portion of total income to each member according to the type of job processed by workers in each household.

probability of taking a Car-the speediest travel mode by 14.7%, which confirms the **Proposition 2** of our structural model.

Apart from income, other factors are also shown to have positive impact on the choice of car, which indicate people will be more likely to drive to work when they get older, and have larger household size, as well as more children under 5. To be précis, the results show that, holding other things constant, 1% increase in age, household size and number of children under five will increase the probability of having a car by 0.69%, 11.4% and 11.8%, respectively.

The non-significant estimated marginal effects for the category of Bus (Table 5) might be caused by the similar reason as for Duplex -- a small number of observations (see Table 6). Marginal effects for the category of Transit are slightly negative. It shows that the change in explanatory variables impacts the probability of choosing the Transit negatively, but by much less than their effects on Non-motor.

5.3 Assessment of Prediction Ability

This section is conducted to assess the ability of our methodology to predict the housing choice and speed choice probabilities. The assessment model presented in this section is similar to that in Agyemang-Duah, Anderson, and Hall (1995). Define A_k as the aggregate probability that a household choose housing type k or the probability that a person choose a certain speed level k, calculated as a relative frequency:

$$A_k = \sum_{n \in N} \frac{P_{nk}}{N} \tag{24}$$

where P_{nk} is the probability that the household (or individual) n choose housing type k (or a certain speed level k). N is the total number of observations. A_k is calculated both for housing choice and speed choice. The calculation is done first in our data sample, and then on the fitted choice probabilities on the basis of the estimated model. The observed and fitted probabilities are presented in Table 7 for the purpose of comparison. The evidence shows that both models perform well in terms of their predicting ability.

In our database, 44.4% household live in Single Family and 44.6% households choose Car as their travel mode, whereas, 34.1% live in Apartment and 37.6% walk or bicycle to CBD. In Great Boston Area, Apartments are located on average closer to the CBD, so some workers in those households can just walk to work. However, for the people live in Single Family, which are located on average further from CBD, the most preferred choice for them is to drive. Our fitted values are highly consistent with the data and strongly confirm the implication of our theoretical model: household with higher income level tends to live further, choose spacious dwelling type, and use speedy travel mode.

6. Conclusion

In this paper, a structural household behavior model is developed for integrated analysis of housing choice and travel mode choice. We derive the optimality conditions for non-housing consumption, housing consumption, leisure and speed demand. We employ Multinomial Logit Model (MNL) and Ordered Logit Model (OLM) to test the implication of the structural model by using 1991 Boston Household Survey data. Empirical results support the prediction of our model that people with higher income tend to live further from the CBD and travel with a speedier mode.

According to the Demographics of Commuting in Great Boston 1990 (by MPO), an increase in the number of commuting trips from distant locations combined with a decrease in trips from close-in locations yields an increase in the length of average trips. Census data on work trip travel time support this conclusion, showing that the average work trip to the urban core increased from 29.1 minutes in 1980 to 31.1 minutes in 1990. For commuters who drove alone, the increase was even more significant, from 26.9 minutes to 30.4 minutes. The evidences do not support one of the most commonly cited benefits of locating offices in the suburbs which is reducing the length of commuting trips, since most employees were presumably living in the suburbs. The rapid growth of employment at recently developed suburban job centers indicates that the relocation of jobs does not necessarily shorten commutes. However, it implies that the impact of suburbanization of jobs is reallocating suburban workers to live in towns further away from the core than an urban worker could and still have a tolerable commute. The further people live from downtown, the lower price they pay for the housing. As shown in this paper, the income elasticity of single family housing is high (the probability of choosing Single Family is increased by 24.3% with 1% increasing household income), which provide people more incentive to move further away to enjoy larger houses and use speedier travel mode. Hence, even though our model deals with the mono-centric city, the prediction can also be extended to be applied to the suburbanization process of jobs and multi-centric city model.

Acknowledgements

We would like to thank Victor Aguirregabiria, Marc Rysman and all the participants in Western Regional Science Association 45th Annual Meeting, Feb 2006 for their helpful comments and suggestions. We are grateful to Boston Metropolitan Planning Organization for providing us the data.

References

Agyemang-Duah, K, Anderson, WP, Hall FL (1995) Trip generation for shopping travel. Transportation Research Record 1493: 12-20.

Ben-Akiva M, Lerman SR (1985) Discrete choice analysis: Theory and application to travel demand. MIT Press, Cambridge, MA.

Boston Metropolitan Planning Organization (1990) The demographics of commuting in Great Boston 1990. Boston, MA.

Boston Metropolitan Planning Organization (1991) 1991 Boston household survey. Boston, MA

DeSalvo JS, Huq M (2005) Mode choice, commuting cost, and urban household behavior, Journal of Regional Science 45 (3): 493-517.

Glaeser E, Kahn M (2004) Sprawl and urban growth. Handbook of Growth and Urban Economics, edited by Henderson V and Thisse J, Volume 4. North Holland Press.

Greene HW (2003) Econometric analysis. Macmillan Publishing Company, New York.

Kopecky K, Suen MH (2005) Suburbanization and the automobile, University of Rochester, mimeo.

Lakshmanan TR, Hua C (1983) A temporal spatial theory of consumer behavior, Regional Science and Urban Economics 13: 341-361.

Margo RA (1992) Explaining the post-war suburbanization of population in the United States: The role of income, Journal of Urban Economics 31: 301-310.

Mills ES, Lubuele LS (1997) Inner cities, Journal of Economic Literature 35: 727-756.

Table 1: Statistics Description of Our Dataset

Variables	Obs	Mean	Stand-Dev.	Min	Max
Housing Type[1]	539	1.45	1.15	0	3
Speed[2]	598	1.59	1.37	0	3
Household Income(Log)	550	0.63	0.54	-1.79	1.79
Household Size	599	2.54	1.26	1	7
No. of Children under5	599	0.15	0.42	0	2
No. of Vehicles	599	1.50	1.06	0	8
Age	595	37.91	12.37	19	82
Female[3]	599	0.51	0.50	0	1

Note: [1] There are four housing types. We denote Apartment = 0, Duplex = 1, Single Family = 2 and Townhouse/Condo=3.

[2] There are four categories of speed ordered from the lowest to the highest: Non-motor = 0, Transit =1, Bus=2 and Car =3.

[3] Gender enters our estimation function as a dummy. We denote female=1 and male =0.

Sources: 1991 Boston household survey conducted by Boston Metropolitan Planning Organization (MPO).

Table 2 : Multinomial Logit Model Estimates for Housing Choice

House Choice	Coefficient	Standard error	Z-value
Duplex			
Household income (log)	1.0910	0.4360	2.5023**
Household size	0.6354	0.2354	2.6992**
No. of Vehicles	1.2123	0.2811	4.3127**
Constant	-5.3123	0.7753	-6.8519**
Single family			
Household income (log)	1.8287	0.3112	5.8763**
Household size	1.2001	0.1695	7.0802**
No. of Vehicles	1.6575	0.2104	7.8779**
Constant	-6.4650	0.6116	-10.5706**
Townhouse/Condo			
Household income (log)	1.5511	0.3135	4.9477**
Household size	0.3997	0.1689	2.3665*
No. of Vehicles	0.4728	0.1872	2.5256**
Constant	-2.8291	0.4648	-6.0867**

Summary statistics

Number of Observations	499
Chi-Square	319.3
Degree of freedom	9
Prob>chi-square	0.0000
Log likelihood (β)	-451.18
Pseudo R-square	0.2614

Notes: The comparison dwelling type is Apartment.

*Significant at 5%

** Significant at 1%

Sources: 1991 Boston household survey conducted by Boston Metropolitan Planning Organization (MPO).

Table 3: Marginal Effects after MNL Estimates for Housing Choice

House Choice	dP/dx	Standard error	Z-value
Apartment			
Household income (log)	-0.3336	0.0538	-6.2007**
Household size	-0.1716	0.0281	-6.1068**
No. of Vehicles	-0.2396	0.0339	-7.0678**
Duplex			
Household income (log)	-0.0071	0.0314	-0.2261
Household size	0.0030	0.0164	0.1829
No. of Vehicles	0.0329	0.0120	2.7417
Single family			
Household income (log)	0.2428	0.0566	4.2898**
Household size	0.2207	0.0304	7.2599**
No. of Vehicles	0.3014	0.0397	7.5919**
Townhouse/Condo			
Household income (log)	0.0979	0.0501	1.9541*
Household size	-0.0521	0.0255	-2.0431*
No. of Vehicles	-0.0947	0.0323	-2.9319**

Note: The marginal effects are estimated at the mean for every independent variable.
*Significant at 5%
** Significant at 1%

Sources: 1991 Boston household survey conducted by Boston Metropolitan Planning Organization (MPO).

Table 4 : Ordered Logit Estimation for Speed Choice

Variable Name	Coefficient	Standard error	z-values
Cut point specific to			
Speed=1 for Transit (λ_1)	2.1250	0.3892	5.4599**
Speed=2 for Bus (λ_2)	2.5827	0.3944	6.5484**
Speed=3 for Car (λ_3)	2.8985	0.3989	7.2662**
Household income (log value)	0.5946	0.1712	3.4731**
Household size	0.4627	0.0790	5.8570**
No. of Children Under 5	0.4769	0.2448	1.9481*
Female	-0.0034	0.1756	-0.0194
Age	0.0278	0.0067	4.1493**
Summary statistics			
Number of Observations	546		
Chi-Square	82.67		
Degree of freedom	5		
Prob>chi-square	0.0000		
Log likelihood (β)	-580.72		
Pseudo R-square	0.0664		

Notes: Z-values= Coefficient/Standard error

* Significant at 5%

** Significant at 1%

Sources: 1991 Boston household survey conducted by Boston Metropolitan Planning Organization (MPO).

Table 5: Marginal Effects after OLM Estimates for Mode Choice

Mode Choice	dP/dx	Standard error	Z-value
Non-motor 0			
Household income (log)	-0.1384	0.0397	-3.4861**
Household size	-0.1077	0.0183	-5.8852**
No. of Children Under 5	-0.1110	0.0569	-1.9508*
Female	0.0007	0.0409	0.0171
Age	-0.0065	0.0016	-4.0625**
Transit 1			
Household income (log)	-0.0100	0.0043	-2.3256*
Household size	-0.0078	0.0028	-2.7857**
No. of Children Under 5	-0.0080	0.0049	-1.6327*
Female	0.0001	0.0030	0.0333
Age	-0.0005	0.0002	-2.5000**
Bus 2			
Household income (log)	0.0018	0.0021	0.8571
Household size	0.0014	0.0016	0.8750
No. of Children Under 5	0.0015	0.0018	0.8333
Female	0.0000	0.0005	0.0000
Age	0.0000	0.0001	0.0000
Car 3			
Household income (log)	0.1466	0.0423	3.4657**
Household size	0.1141	0.0196	5.8214**
No. of Children Under 5	0.1176	0.0604	1.9470*
Female	-0.0008	0.0433	-0.0185
Age	0.0069	0.0017	4.0588**

Note: The marginal effects are estimated at the mean for every independent variable.
The number associated with every mode is the speed ordered from the slowest (Non-motor) to the speediest (Car).
*Significant at 5%
** Significant at 1%

Sources: 1991 Boston household survey conducted by Boston Metropolitan Planning Organization (MPO).

Table 6: Observed and Fitted Aggregate Probabilities

Housing Choice		Observed	Predicted by MNL
Duplex		0.0594	0.0596
Single Family		0.4044	0.3870
Townhouse/Condo		0.1948	0.1982
Apartment		0.3414	0.3552
Mode Choice	**Order**	**Observed**	**Predicted by OLM**
Non-motor	0	0.3756	0.3884
Transit	1	0.1018	0.0990
Bus	2	0.0751	0.0687
Car	3	0.4457	0.4438

Sources: 1991 Boston household survey conducted by Boston Metropolitan Planning Organization (MPO).