Incorporating Principles of Bounded Rationality into Models of Pedestrian Shopping Behavior

Theory and Example

Wei Zhu and Harry Timmermans
Urban Planning Group, Eindhoven University of Technology
Vrt 08.16, Postbus 513, Den Dolech 2
5600 MB Eindhoven
The Netherlands
w.zhu@tue.nl

Key words: Bounded rationality, decision heuristics, heterogeneity, pedestrian, store patronage decision

Abstract: The modeling and simulation of pedestrian behavior has been dominantly relied on rational choice models in which pedestrians are assumed to be capable of processing a large number of choice alternatives and trade off attribute utilities. In reality, however, pedestrian behaivor shows the evidence of bounded rationality. They simplify the decision problem by considering a limited number of factors as well alternatives, using heuristics to arrive at satisfactory as opposed to optimal choices. Incorporating principles of bounded rationality in pedestrian modeling will benefit the understanding of individual decision processes and planning practice. This paper proposes an approach that models the cognitive processes such as filtering factors, constructing preference structures, deriving heterogeneous decision heuristics, and selecting decision strategies. The approach is also exemplified through estimating the model on pedestrian store patronage behavior data, collected in a shopping center in Shanghai, China. The results show the estimated probabilities of usage of decision strategies and the sequences of factor search. Compared with the conventional multinomial logit models, the results indicate the statistical advantages of the new approach.

1. INTRODUCTION

The majority of models developed to predict or simulate pedestrian shopping behavior are based on the principle of utility-maximizing behavior (e.g., Borgers and Timmermans, 1986; Saito and Ishibashi, 1992; Antonini and Bierlaire, 2007). To the extent we do not view this as simply a convenient modeling principle but add value to the theoretical underpinnings of utility-maximizing models, it is assumed that pedestrians when making decisions take into account a potentially large set of influential factor, discriminate between choice options in a very detailed, continuous manner and arrive at a decision by combining the utilities they attach to each factor according to some typically weighted additive combination rule. Considering the complexity of the decision making process, this theoretical framework does not seem very appealing in the sense that the validity of these assumptions is highly questionable. It seems more likely that pedestrians simplify the decision problem by considering a limited number of factors, compare options in a more global way and use heuristics to arrive at satisfactory as opposed to optimal choices. That is, they likely show evidence of bounded rationality (Simon, 1959) as opposed to full rationality.

Although the concept of bounded rationality has been around for many years, and have attracted some research (e.g., Payne, et al., 1988; Gigerenzer, et al., 1999), fully operational models are scarce and we are not aware of any model based on principles of bounded rationality developed to predict and simulate pedestrian shopping behavior.

In this paper, we present the development of a model of pedestrian behavior based on principles of bounded rationality. In addition to the fundamentals of the model, we present an application to store choice behavior, using data collected in East Nanjing Road, Shanghai, China. The performance of these models based on principles of bounded rationality will be compared with the results of a multinomial logit model, currently the dominant model in this area of research.

The paper is organized as follows. We will first present the development and specification of the various models of bounded rationality. Next, general principles will be applied to the problem of store choice behavior. We will continue with a description of the data collection, followed with a discussion of the main results. The paper is completed with a summary and discussion.

2. MODEL DEVELOPMENT

Let $X = \{x_j, j = 1,...,J\}$ represent the set of attributes or factors influencing the decision of interest. Assume that individuals do not necessarily take all these factors into account, but rather solve a decision problem by mentally (re)constructing the problem and selecting a subset of these factors. This *filtering* process is not invariant, but will depend on the decision problem, and more importantly on the activation level of the individual. Let δ_j represent an activation threshold for factor x_j. These thresholds act as filters. Thus, by consciously or unconsciously applying these thresholds, a subset of activated factors will enter the decision making process. Only if all the thresholds are equal to zero (defining that all factor stimulation can be transformed into positive real numbers and larger values represent stronger stimulation), all factors will be considered. Mathematically, this can be expressed as:

$$s_j = \begin{cases} 0 & \text{if } x_j < \delta_j \\ 1 & \text{if } x_j \geq \delta_j \end{cases} \tag{1}$$

where s_j is the state of the factor in mental representation. Consequently, the set of factors considered, $X' = \{x_j \mid s_j = 1\}$.

Once the irrelevant factors have been filtered, bounded rationality suggests that individuals tend not to discriminate between all possible values of factors. Rather, they will categorize the continuous factors into discrete classes or states, or re-categorize discrete factors. Assume that in case of continuous factors, this process of factor representation involves the application of a monotonically increasing set of threshold values, which discretize the continuous factors into an ordered set of discrete classes. Let $\Delta_j = \{\delta_{j1} < \delta_{jn} < \delta_{jN}, n = 1,...,N\}$ be a set of successively increasing activation thresholds for x_j, corresponding to stricter judgment standards. (Note that *N* can be factor-dependent, so it should be N_j. For representation simplicity, the subscript is ignored.) A factor may then meet one or more of these increasingly stricter activation thresholds and hence becomes more informative. The relevant equations then become,

$$s_{jn} = \begin{cases} 0 & \text{if } x_j < \delta_{jn} \\ 1 & \text{if } x_j \geq \delta_{jn} \end{cases}$$
$$X' = \{x_j \mid s_{jn} = 1\} \tag{2}$$

Thus, filtering and factor representation transform categorical and continuous external factors into a set of activated and non-activated internal (mental) factor states.

Individuals will judge these states by (1) attaching values, (2) assigning relative importance weights, (3) integrating these values for individual states in some way to arrive at an overall judgment, and (4) evaluating the overall judgment against some overall threshold value in light of underlying goals. Attaching judgment values to states implies that the state is judged and valued in light of the decision goal. Weights indicate the relative importance of states in the decision problem. Because these values and weights are all unknown parameters in this approach, they are combined into a single value, w_{jn}, which can be interpreted as a part-worth utility. Let $u_{jn} = w_{jn} s_{jn}$ denote the value judgment of state n of factor x_j. All states that are incorporated in the decision making process need to be combined according to some integration rule to arrive at an overall value judgment for each choice alternative. Various rules can be used. Thus, if an additive integration rule is assumed, the overall value judgment of choice alternative i equals:

$$v_i = \sum_j \sum_n u_{ijn} \tag{3}$$

In the final step, assume that the overall values are also categorized and mapped by checking them against a set of successively increasing overall thresholds $\Lambda = \{\lambda_1 < \lambda_m < \lambda_M, m = 1,...,M\}$, resulting in the overall states, p_{im}. This can be expressed as:

$$p_{im} = \begin{cases} 0 & \text{if } v_i < \lambda_m \\ 1 & \text{if } v_i \geq \lambda_m \end{cases} \tag{4}$$

In case this mapping only involves two preference orders (reject or accept), only one λ is needed and $p_{i1} = 0$ defines rejecting the alternative, whereas $p_{i1} = 1$ implies accepting it. For representation simplicity, the following model formulations assume only two preference levels exist.

Define a state value set for each factor,

$$V_j = \{v_{j1} = 0, v_{j2} = w_{j1}, v_{j3} = w_{j1} + w_{j2}, ..., v_{jN+1} = \sum_{n=1}^{N} w_{jn}\} \tag{5}$$

which includes all possible value judgments related to the factor. Let $\overline{v}_k$ represent any factorial combination from value judgments in the sets and construct the value set of all possible overall value judgments,

$$\overline{V} = \{\overline{v}_1 < \overline{v}_k < \overline{v}_K; k = 1,2,...,K; K = \prod_j (N_j + 1)\} \tag{6}$$

Checking these overall value judgments against the overall threshold λ, results in a unique pattern of relationships with some value judgments above the threshold, and some below the threshold. Thus, the set of overall value judgments $\overline{V}$ can be divided into a subset $\overline{V}_0$ of rejected overall value judgments and a set $\overline{V}_1$ of accepted ones. This pattern can be viewed as a discrete preference structure, Φ, that is used to classify overall value judgments of alternatives into an ordered set of preferences (in this case reject or accept). Mathematically,

$$\Phi = \begin{Bmatrix} \bar{v}_k \in \bar{V}_0 \mid \bar{v}_k < \lambda \\ \bar{v}_k \in \bar{V}_1 \mid \bar{v}_k \geq \lambda \end{Bmatrix} \tag{7}$$

We assume that in every choice context, individuals will consciously or unconsciously define a set of threshold values and apply choice heuristics which are logically consistent with the preference structure. Because for different individuals or in different contexts preference structures may differ in terms of the pattern of the sets of accepted and rejected values, this implies that our cognitive process model automatically generates *heterogeneous choice heuristics*. On extreme is the strictest preference structure in the sense that no single value (judgment) combination survives the overall threshold,

$$\Phi = \left\{ \bar{v}_k \in \bar{V}_0 \mid \bar{v}_k < \lambda \right\} \tag{8}$$

That means that regardless of the states of the factors, the choice alternative under consideration will be rejected. In this case, no choice heuristics are implied (or the heuristic of "no action" since the individual has no need to consider any information). Relaxing λ a little leads to the preference structure where only the value combination of factor states with the highest threshold values is accepted,

$$\Phi = \begin{Bmatrix} \bar{v}_k \in \bar{V}_0 \mid \bar{v}_k < \lambda \\ \bar{v}_k \in \bar{V}_1 \mid \bar{v}_k \geq \lambda, \bar{v}_k = \sum_j v_{jN+1} \end{Bmatrix} \tag{9}$$

This preference structure implies conjunctive heuristics in the sense that an alternative will be accepted only when all factors are in their highest states. During the decision process, any single factor being unsatisfactory will cause the decision process to stop, regardless of the states of the other factors.

At the opposite end is the most relaxed preference structure, represented the case that all factor combinations will be accepted.

$$\Phi = \left\{ \bar{v}_k \in \bar{V}_1 \mid \bar{v}_k \geq \lambda \right\} \tag{10}$$

This preference structure implies the other "no action" heuristic since factors being in whatever state will lead to the alternative being accepted. A little less tolerance for λ may result in a preference structure where all but the value combinations of non-activated factor states are accepted,

$$\Phi = \begin{Bmatrix} \bar{v}_k \in \bar{V}_0 \mid \bar{v}_k < \lambda, \bar{v}_k = \sum_j v_{j1} \\ \bar{v}_k \in \bar{V}_1 \mid \bar{v}_k \geq \lambda \end{Bmatrix} \tag{11}$$

Disjunctive heuristics can be inferred from this preference structure since any factor state (except the non-activated state) being satisfactory will cause the decision process to stop and accept the choice alternative, regardless of the state of the other factors.

Within the spectrum, various other preference structures and heuristics can be identified. For example, the lexicographic heuristic is implied in a preference structure,

$$P = \begin{Bmatrix} \bar{v}_k \in \bar{V}_0 \mid \bar{v}_k < \lambda, \sum_k \sum_{t=1}^{n} s_{jt|k} = 0 \\ \bar{v}_k \in \bar{V}_1 \mid \bar{v}_k \geq \lambda, \prod_k \prod_{t=n'}^{N} s_{jt|k} = 1 \end{Bmatrix} \qquad n < n' \tag{12}$$

According to this preference structure, there exists at least one factor j. When some states of this factor are not activated, the decision process will stop and reject the alternative. When some states are activated, the decision process will stop at accepting the alternative. In-between are those states whose status cannot determine accepting or rejecting the alternative and further consideration on other factors is needed.

For above we can see, preference structures are directly related to choice heuristics. Assume that different individuals or the individual in different contexts may apply different preference structures and corresponding choice heuristics to solve problems. That is, people have a context-dependent repertoire of preference structures and corresponding heuristics. Although we should always try to specify the context as much as possible, there will always remain some stochastic element from the viewpoint of the analyst. Such randomness can be mathematically included into the overall threshold, so that we get $\lambda \sim f$, where f is a probability density function. Because $\bar{V}$ is a discrete set, between consecutive pairs of $\bar{v}_k$, there is a range of λ, satisfying $\bar{v}_{k-1} < \lambda \leq \bar{v}_k$. It represents the range of an invariant preference structure. The probability of this preference structure Φ_k being applied, p_k, equals the probability of λ being in this range, given f is a continuous distribution:

$$p_k = \int_{\bar{v}_{k-1}}^{\bar{v}_k} f \; dt \tag{13}$$

We may equivalently view this as the probability of applying choice heuristics implied by the preference structure. Thus, any single decision may be a two-step process, choosing an appropriate preference structure and applying this structure to the choice task, forming preferences among alternatives and making the choice. Because the preference structure actually applied by the decision maker is usually unknown, the final probability of an alternative being satisfactory can be modeled as the expected result of choice outcomes aggregated across all possible choice outcomes under these latent preference structures, or mathematically:

$$p_i = \sum_{k=1}^{K+1} p_k p_{i|k} \quad (14)$$

where $p_{i|k}$ is the probability that alternative i is satisfactory when preference structure k is applied. It has the same specification as p_{im} in Equation 4. However, because within an invariant range the value of λ does not affect choice outcomes, λ does not need to be identified. Instead, $\bar{V}$ is enough as a critical value for the overall threshold. Although the process of selecting A preference structure itself may be susceptible to bounded rationality, here only the outcome of this process is modeled. Assuming the distribution of preference structures can be represented by a multinomial logit distribution, the individual will select the preference structure which generates the highest utility. The probability of a preference structure being applied can then be modeled as

$$p_k = \frac{\exp(u_k)}{\sum_{k=1}^{K+1} \exp(u_k)} \quad (15)$$

where u_k is the observable utility that the individual derives from preference structure k . In fact, Equation 15 represents the probabilities of the heuristics implied by a preference structure being selected. However, because for certain preference structures applying different heuristics does not affect the choice outcome, Equation 15 can also be formulated as the aggregation of probabilities that heuristics within implied the preference structures are chosen,

$$p_k = \sum_{h=1}^{J!} p_{kh} = \frac{\sum_{h=1}^{J!} \exp(u_{kh})}{\sum_{k'=1}^{K+1} \sum_{h'=1}^{J!} \exp(u_{k'h'})} \quad (16)$$

where u_{kh} is the utility of heuristic h implied by preference structure k . We assume that this utility is composed of three factors: mental effort, risk attitude and expected outcome.

It is the obvious that the more factors considered and the more alternatives to be evaluated to make a decision, the more mental effort has to be invested. In this paper, the influence of the number of alternatives will not be modeled. It is treated as part of the context of the decision problem. The emphasis is on the factors involved, and moreover, on the sequence of factor consideration. This emphasis is motivated by the fact that the influence of factors on decisions differs. Important factors may be decisive for a decision and if they are considered earlier, they may obliterate the need to consider other factors and cost less mental effort. However, different types of information may also cause the processing effort to differ between factors. For example, judging the shape of an object may be more difficult than judging color. A complicating factor is that individuals cannot be sure about the amount of mental effort that is involved. They can only subjectively estimate it based on their beliefs p_{jn} that the activated factors occupy states that make any further consideration of subsequent factors useless.

To illustrate, let three factors x_1, x_2, and x_3 have, A, B, and C states respectively ($a = 1,...,A; b = 1,...,B; c = 1,...,C$). Assume that the heuristic under consideration implies the search sequence $x_1 \to x_2 \to x_3$. Let e_1, e_2, and e_3 denote the amount of mental effort inflicted when considering factors x_1, x_2, and x_3 respectively, and let p_a, p_b, and p_c represent the individual's beliefs that factors are in the states with value judgments v_a, v_b, and v_c respectively, such that $\sum_a p_a = 1, \sum_b p_b = 1, \sum_c p_c = 1$. The expected amount of mental effort is then defined as,

$$e_h = e_1 + \sum_a (p_a e_2 Y_a + \sum_b p_a p_b e_3 Y_{ab}) \tag{17}$$

$$Y_a = \begin{cases} 0 & \text{if } v_{abc} < \lambda \vee v_{abc} \geq \lambda \quad \forall b, \forall c \\ 1 & \text{otherwise} \end{cases} \tag{18}$$

$$Y_{ab} = \begin{cases} 0, & \text{if } v_{abc} < \lambda \vee v_{abc} \geq \lambda \quad \forall c \\ 1, & \text{otherwise} \end{cases} \tag{19}$$

Equation 17 reflects the fact that e_1 is inevitably fully inflicted since x_1 is considered first. For each possible state of x_1, expected efforts are derived from two terms. First, the effort of considering x_2 is weighted by the probability of x_1 being in a particular state and Y_a, an identity function defined by Equation 18. Y_a represents a judgment process, according to which an individual checks whether all subsequent value combinations $v_{abc} \mid a$, are inactivated against λ, or all value combinations are activated. If all value combinations are inactivated, the corresponding factor is not considered and no additional mental effort is involved. If all factors are activated, it means that that the same decision or preference applies to all instances of that factor and hence considering the factor will not have any effect on the preference ordering or decision. In these cases, $Y_a = 0$. In contrast, when $Y_a = 1$, x_2 needs to be considered. According to the same logic, the second term relates to considering x_3 when at a state of x_2. e_3 is weighted by $p_a p_b$, the joint probability of being in the previous two factor states, and Y_{ab} is another identity function judging whether the simultaneous conditions $v_{abc} \mid a, b$ against λ are satisfied or not. By this definition, due to the fact that efforts for considering factors may differ and different factor values may cause earlier or later termination of the decision process the expected efforts related to some consideration sequences may differ as well when the expected overall values are homogeneous against the overall threshold.

By making decisions, individuals take risks. By selecting a very high or very low λ within the overall value space, the individual will have a very high probability to reject or accept an alternative since most of the information will fall into the overall value set that is lower or higher than λ. Decisions may be very simple and cost little cognitive effort, however, at higher expected opportunity costs. Such opportunity costs are tightly related to the expected regret resulting from a potential false rejection or false acceptation. In contrast, selecting a mild λ will control the expected regret to a minimum by staying uncertain about the outcome at every stage of information search and looking for more information in order to get a comprehensive view of the choice alternatives. No doubt that the drawback of doing so is that it is very effortful. Therefore, more outcome variety means lower decision risk. The degree of risk attitude is determined by the location of λ and the probability beliefs. To model this property, Shannon's Information Entropy is applied because it has been developed specifically for measuring information uncertainty. Let $\overline{p}_k$, corresponding to $\overline{v}_k$, be the factorial joint product of the probability beliefs on factor states. The probability of a positive, r_{kh}^{+}, respectively negative, r_{kh}^{-}, outcome equals:

$$\begin{aligned} r_{kh}^{+} &= \sum_k \overline{p}_k Y(\overline{v}_k \geq \lambda) \\ r_{kh}^{-} &= 1 - r_{kh}^{+} \end{aligned} \tag{20}$$

where $Y(\zeta)$ is an identity function being 1 when ζ is true and 0 when ζ is false. It follows that the risk attitude for heuristics implied by the same preference structure is the same because different information search sequences do not change the choice outcome. Therefore, subscript h may be excluded. Then, the risk attitude of a preference structure is,

$$r_k = -r_k^{+} \log_2(r_k^{+}) - r_k^{-} \log_2(r_k^{-}) \tag{21}$$

The property of Information Entropy measure is that the value is at its maximum when the occurrences of alternatives are equally possible and is at its minimum when the occurrences are absolutely certain. In this case, the maximum value is 0.5, which represents the lowest risk attitude, and the minimum value is 0, which represents the highest risk attitude.

Individuals may have expectations about the outcome of a decision which is directly related to the context of the decision problem because people may have preference for outcomes. This makes the selection of judgment standard not a neutral process but it likely involves value biases. A decision standard which leads to more probable occurrences of preferred outcomes is more likely to be selected. Assuming that the expected outcome can be represented in a form consistent with the expected utility theory, we have:

$$o_k = o_k^+ r_k^+ + o_k^- r_k^- \qquad (22)$$

where o_k^+ is the value of the satisfactory outcome and o_k^- is the value of the unsatisfactory outcome. The sequence of factor search does not have an influence. In total, the utility of a heuristic is defined as the linear combination of the three elements,

$$u_{kh} = \beta^e e_{kh} + \beta^r r_k + \beta^o o_k \qquad (23)$$

β^e, the parameter of effort, is assumed to be positive because e_{kh} is set negative to represent a kind of costs. β^r is also assumed to be positive because people are assumed to prefer low decision risks. The sign of β^o needs to be empirically determined.

3. APPLICATION TO STORE PATRONAGE DECISIONS

Let $i = 1,...,I$ represent the stores that could be patronized in the order of ascending distance to the pedestrian's current location, d_i. Store I is chosen because it is satisfactory to the pedestrian and stores evaluated earlier are unsatisfactory. Expressed in probabilistic terms:

$$\begin{aligned} &p_I = p_I^S \prod_{i=1}^{I-1} p_i^U \quad d_1 < ... < d_{I-1} < d_I \\ &p_i^S + p_i^U = 1 \end{aligned} \qquad (24)$$

where p_i^S is the probability of the store being considered satisfactory and p_i^U the probability of being considered unsatisfactory. It is assumed that four factors influence the store patronage decision: (1) c_i, the number of times that the pedestrian has visited the store during this shopping trip. It is hypothesized that the probability to patronize the store will decrease if the store has been patronized; (2) q_i, the retail floorspace of the store representing store attractiveness; (3) s_{ij}, a dummy variable representing the type of the store. The following store types were identified and each store is labeled a retail type, including, art ($j=1$, Arts), book & media ($j=2$, Book), children ($j=3$, Chil), clothes ($j=4$, Clth), department ($j=5$, Dept), equipment ($j=6$, Equi), drink and food ($j=7$, Fddr), fast food $j=8$, Fdfa), formal meal ($j=9$, Fdfo), food retailing ($j=10$, Fdre), jewelry ($j=11$, Jewe), optical ($j=12$, Opti), pharmaceutical ($j=13$, Phar), shoe ($j=14$, Shoe), sports ($j=15$, Spor), tobacco ($j=16$, Toba), tourism ($j=17$, Tour), and others ($j=18$, Oths); (4) store dominance, m_i, defined as the ratio of the store floorspace to the total floorspace of the stores within 100 m radius of the store, to represent the uniqueness or competitiveness of a store relative to its retail environment. The value is between 0 – 1.

For the continuous and ordinal variables, c, q, and m, thresholds are used for representing factors. For discrete variable, s, each store type is represented as an interest category. This set of categories reflects that the pedestrian may cognize different types of stores into limited degrees of interests, with $Z=\{z_1,...,z_K, K \le J\}$. Then each category is assigned a value, β_k^z. The pedestrian will patronize the store, if he/she finds that,

$$\sum_x \mathbf{W}^x \Psi(x_i \ge \mathbf{\Delta}^x) + \sum_{k=1}^{K} \beta_k^z z_{ik} \ge \lambda \quad x = c, q, m \tag{25}$$

Here $\mathbf{W}^x = [w_1^x,...,w_j^x,...,w_J^x]$ is a *J*-element row vector of factor state values, $\mathbf{\Delta}^x = [\delta_1^x,...,\delta_j^x,...,\delta_J^x]^{\mathrm{T}}$ is a column vector of factor threshold values, and $\Psi(\psi)$ is an element-wise identity function being 1 for the true relationships ψ, being 0 for the false relationships. As has been explained in the previous section, λ is assumed to be a multinomial logit distribution from which heterogeneous decision strategies originate. To estimate this distribution, the utility of each decision heuristic is calculated. The estimations of $\mathbf{W}^x$ and $\mathbf{\Delta}^x$ provide the cognitive structure, from which the stopping conditions for each heuristic can be inferred.

For comparison, the probability of store satisfactoriness under the MNL framework can be specified as,

$$p_i = \exp(v_i^S) / (\exp(v_i^S) + \exp(v_i^U))$$
$$v_i^S = \sum_x \beta^x x_i + \sum_{j=1}^{17} \beta_j^s s_{ij} \quad x = c, q, m \qquad (26)$$
$$v_i^U = \beta^U$$

Note that the "others" store type is set as the base type, so 17 type parameters are involved. The utility of not visiting this store, v_i^U, is just represented by a parameter, β^U. Another version of the MNL model, with the continuous variables being naturally logged for representing marginal decreasing utility change, will also be estimated.

4. DATA

The data for estimating the models were collected in 2007 in East Nanjing Road (ENR), Shanghai. The street is about 1,600 meters long, and 1,000 meters of this is pedestrianized (*Figure 1*).

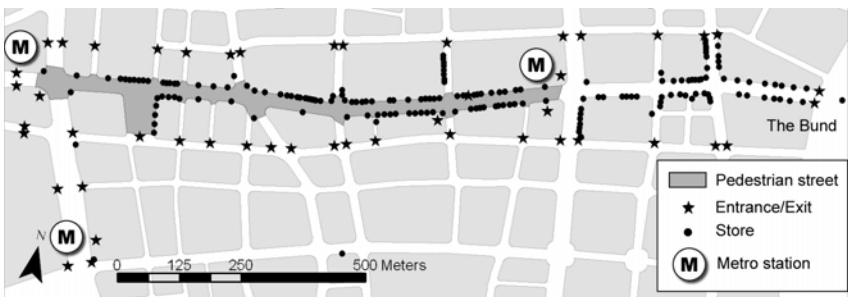

Figure 1 The survey area of ENR

Most of the shops are located along both sides of the street, shaping a linear shopping space. Twenty students from the Department of Urban Planning, Tongji University, administrated the survey during two days, May 19 (Saturday) and May 22 (Tuesday). Each day from 12:00 to 20:00, they randomly invited pedestrians to answer the questions and recorded the respondent's activity diary up to the moment of the interview. Each record of a pedestrian diary starts from recording the arrival time of the pedestrian and the entrance, the location where the shopping begins. Then in sequence, the respondent reported the stores he/she had visited, the estimated duration in the stores, the goods bought, and the expenditure, up to the current situation. Because that not all the respondents had completed the shopping trip when they were being interviewed, the model estimation used the subsample of the data including the respondents with complete diaries, with the sample size 236. Each visit in a store is treated as an independent store patronage decision case. That results in 47,111 cases.

5. RESULTS

The model parameters are estimated using log-likelihood statistic (LL). Consistent Akaike Information Creterion (CAIC) is used for model selection in the sense the models with the least CAIC will be presented. The calibration results in Table 1 and 2 show that the number of visits in the store is insignificant for all models, as it turns out that the occurrence of the visited store being in the store consideration sets in the following decisions is little. The MNL model with normal variables performs well with MLR equal to 0.78. The model improves considerably when the floorspace variable is logged, suggesting the decreasing trend of utility increment relative to increasing store size. The factor of store dominance is not significant. Tourism site is the most attractive type as its parameter is the highest. The second most attractive store type is department store, followed by food retailing store and book store. The stores for formal meals seem to be the least attractive to the pedestrians.

The proposed model has the best LL as well CAIC. Floorspace is represented into four states, [< 50 m^2, 50 – 420 m^2, 420 – 24,000 m^2, ≥ 24,000 m^2). The influence of store dominance is significant. Its threshold value is near 1, suggesting that detached store location to other stores can enhance the attractiveness of the store, probably because pedestrians can fully concentrate their attention on the store. However, since only one store suffices this condition, there is the possibility that this variable represents alternative-specific tastes. As for the store type, four interest categories are estimated, from the most interesting (4) to the least interesting (1). The most interesting category only includes tourism sites. The second most interesting category includes department stores and food retailing stores, mainly those selling local special food. Figure 2 shows the estimated probabilities of the preference structures being applied. As the there are 3 states for floorspace, 2 states for dominance and 4 states for store type, the total number of preference structures is $33 = 3 * 2 * 4 + 1$, with the preference strictness increasing from left to right. It indicates that in most decisions the pedestrians could have rejected the stores being evaluated without considering any information. This is understandable since there are so many stores which are assumed to be potentially considered by the pedestrian before the satisfactory store is found. β^o being negative suggests that strict heuristics are preferred in general, as can be seen from the slightly rising lines from the left. However, a decreasing trend is found after preference structure 25 as the result of high decision risk. The dash lines in the figure indicate the probabilities of each factor being searched first, aggregated from the probabilities of every heuristic starting from this factor. Store type seems to be the most important factor for store evaluation since it almost always has the highest probabilities to be search first in the preference structures implying information search behavior. This is consistent with the common sense that people patronize stores for realizing their needs. The probabilities of floorspace and store dominance being search first are very similar. In total, the aggregated probabilities of first-to-search factors are, *s* – 41%, *q* – 17%, and *m* – 14%.

Table 1. Results of the MNL models

Normal variables		Logged variables	
Parameter	Estimate	Parameter	Estimate
β^q	2.480e-5	β^q	0.193
[β^m][(1)]	0.000	[β^m]	0.000
[β^s_{Arts}]	0.000	[β^s_{Arts}]	0.000
β^s_{Book}	1.151	β^s_{Book}	0.817
β^s_{Chil}	1.268	[β^s_{Chil}]	0.000
β^s_{Clth}	0.575	β^s_{Clth}	0.410
β^s_{Dept}	2.217	β^s_{Dept}	1.823
[β^s_{Equi}]	0.000	[β^s_{Equi}]	0.000
[β^s_{Fddr}]	0.000	[β^s_{Fddr}]	0.000
β^s_{Fdfa}	0.660	β^s_{Fdfa}	0.546
[β^s_{Fdfo}]	0.000	β^s_{Fdfo}	-1.154
β^s_{Fdre}	0.954	β^s_{Fdre}	0.895
[β^s_{Jewe}]	0.000	[β^s_{Jewe}]	0.000
β^s_{Opti}	-0.788	β^s_{Opti}	-0.670
[β^s_{Phar}]	0.000	[β^s_{Phar}]	0.000
[β^s_{Shoe}]	0.000	[β^s_{Shoe}]	0.000
[β^s_{Spor}]	0.000	[β^s_{Spor}]	0.000
[β^s_{Toba}]	0.000	[β^s_{Toba}]	0.000
β^s_{Tour}	3.328	β^s_{Tour}	2.935
β^U	3.977	β^U	4.892
NC	47,111	NC	47,111
NP	10	NP	10
LL	-7,138	LL	-7,095
CAIC	14,394	CAIC	14,308

[(1)] Parameters in [ ] are not counted as free parameter

Table 2 Results of the proposed model

Parameter	Estimate	Parameter	Estimate	Parameter	Estimate
[δ_1^q]	50 m^2	[z_{Clth}]	2	[z_{Spor}]	2
[δ_2^q]	420 m^2	[z_{Dept}]	3	[z_{Toba}]	1
[δ_3^q]	24,000 m^2	[z_{Equi}]	2	[z_{Tour}]	4
w_1^q	4.293	[z_{Fddr}]	2	[(z_{Oths})]	1
w_2^q	4.775	[z_{Fdfa}]	2	[β_1^z]	0.000
w_2^q	1.405	[z_{Fdfo}]	1	(β_2^z) (1)	1.000
[δ^m]	0.999	[z_{Fdre}]	3	β_3^z	3.895
w^m	9.370	[z_{Jewe}]	1	β_4^z	9.770
[z_{Arts}]	1	[z_{Opti}]	1	β^e	-2.893
[z_{Book}]	2	[z_{Phar}]	1	β^r	6.450
[z_{Chil}]	2	[z_{Shoe}]	1	β^o	-2.894
NC	47,111				
NP	10				
LL	-7,012				
CAIC	14,141				

(1) Parameters in ( ) are set in estimation

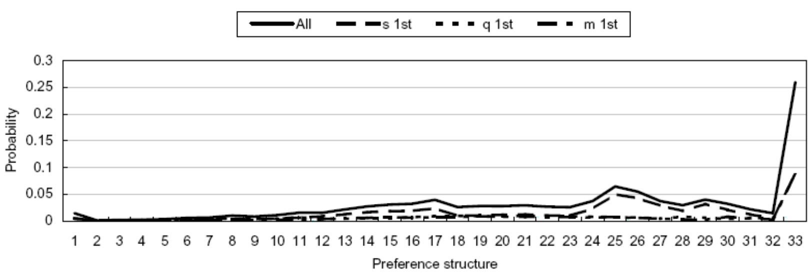

Figure 2 Distribution of preference structure

6. CONCLUSIONS AND DISCUSSION

In this paper, we have developed a model of pedestrian store choice/patronage behavior, based on principles of bounded rationality, and applied it to data collected in East Nanjing Road, Shanghai. The performance of the model was compared against the performance of two multinomial logit models, based on the assumption of utility-maximizing bahavior. Results indicate that the proposed model performs better than the MNL models both in terms of LL and CAIC. This suggests that pedestrians could use simplifying decision strategies rather than the principle of utility-maximization in their store choice decision making. Moreover, it shows that strict and simple decision heuristics are preferred by pedestrians. However, strict but too risky heuristics, such as pure conjunctive rules, are rarely applied. The exception is the unconditional rejection heuristic, just because it costs little mental effort.

Thus, this study has provided empirical evidence that pedestrian decisions may be based on the principle of bounded rationality. Developing and elaborating these models thus has some potential advantages in terms of the validity of the models. Whether, these models also lead to significantly better predictions is a subject of future research.

7. REFERENCES

Antonini, G., Bierlaire, M., 2007, "A Discrete Choice Framework for Acceleration and Direction Change Behaviors in Walking Pedestrians", in *Pedestrian and Evacuation Dynamics 2005*, Waldau N, Gattermann P, Knoflacher H, Schrechenberg M (Ed.), 155-166

Borgers, A., Timmermans, H. J. P., 1986, "A Model of Pedestrian Route Choice and Demand for Retail Facilities within Inner-city Shopping Areas", *Geographical Analysis*, 18(2), 115-128

Gigerenzer, G., Todd, P. M., and ABC Research Group, 1999, *Simple Heuristics that Make us Smart*, Oxford University Press, Oxford, New York.

Payne, J. W., Bettman, J. R., Johnson E. J., 1988, "Adaptive Strategy Selection in Decision Making", *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14, 534-552

Saito, S., and Ishibashi, K., 1992, "A Markov Chain Model with Covariates to Forecast Consumer's Shopping Trip Chain within a Central Commercial District", *Presented at Fourth World Congress of Regional Science Association International*, RSAI, Leeds, United Kingdom

Simon, H. A., 1959, "Theories of Decision-Making in Economics and Behavioral Science", *The American Economic Review*, 49 (3), 253-283