Environment and Planning B: Planning and Design 2010, volume 37, pages 59 – 74

doi:10.1068/b35039

Cognitive process model of individual choice behaviour incorporating principles of bounded rationality and heterogeneous decision heuristics

Wei Zhu¶, Harry Timmermans
Urban Planning Group, Vertigo 08.18, Eindhoven University of Technology, Den Dolech 2, 5600 MB Eindhoven, The Netherlands; e-mail: zhu@mpib-berlin.mpg.de, h.j.p.timmermans@bwk.tue.nl
Received 1 April 2008; in revised form 30 December 2008; published online 26 October 2009

Abstract. Although the principle of bounded rationality seems more realistic for formulating formal models of individual choice behaviour than traditional decision-outcome-based discrete choice models, existing studies have some limitations: (1) applications focus on the noncompensatory nature of the models and largely ignore the factor-selection process; (2) heterogeneity of heuristics in reaching a decision is insufficiently studied; (3) the choice of decision strategy is rarely modeled formally. A modeling approach that simultaneously deals with these issues is suggested. Factor thresholds are used as the mechanism for factor selection and representation, resulting in a set of activated and nonactivated factor states. Under the assumption of stochastic contextual effects, the model automatically generates heterogeneous decision heuristics, including the conjunctive, disjunctive, and lexicographic rule. Mental effort and risk perception are assumed to influence the evaluation and choice of heuristics. The concept of preference tolerance is used to predict the probability of selecting a particular heuristic under different decision contexts. The modeling approach is illustrated using the go-home decision of pedestrians in a shopping street in China as an example.

1 Introduction

Modeling pedestrian behaviour in urban environments has received much attention since the late 1990s. The main incentive stimulating this trend is probably the advancement of computer and programming technologies which allow researchers to model and simulate complex behaviour from a bottom-up perspective through agent-based modeling (see, for example, Dijkstra and Timmermans, 2002; Haklay et al, 2001; Kerridge et al, 2001). In addition, other studies examine local movement dynamics of pedestrians such as queue following and obstacle evading. A typical example is the social force model (see, for example, Helbing and Molnár, 1995; Helbing et al, 2001) in which pedestrian movement is modeled as the result of competing forces from the environment surrounding the pedestrian, based on principles similar to Newtonian mechanics. However, work on calibrating such models is very rare (see, for example, Hoogendoorn et al, 2007); the parameters of these models are often set arbitrarily or sometimes measured directly from observable walking properties such as speed and spacing (see, for example, Willis et al, 2004). Calibration is usually implemented at more aggregate levels to examine the extent to which emergent phenomena produced through simulation, such as specific shape of flow, density, queue, and clog, are consistent with actual observations.

We consider that, apart from viewing and modeling pedestrian movement analogous to the behaviour of particles in physical fields, aspects such as cognition, decision, and psychological activities are equally important for pedestrian modeling and useful in practice. Moreover, calibrating pedestrian models at the level of decision making is

¶ Also: Centre for Adaptive Behavior and Cognition, Max-Planck-Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany.

considered to be equally important for validating underlying assumptions. Almost all behaviour can be understood as the result of choice decisions. For example, pedestrians decide to leave a place or stay, or go in one direction rather than another.

Models of pedestrian choice behaviour (and other types of activity choices) have mostly been based on theories of rational choice. These theories suggest that individuals invariably take into account the set of factors influencing their decision, attach some value judgment (utility, attitude, satisfaction) to each choice alternative, and choose the alternative that maximises their value judgment. Most operational models use an additive function to represent the value-judgment process, implying that a compensatory decision-making process is assumed in the sense that a low judgment on some factor may be at least partially compensated by higher judgments of one or more of the remaining factors influencing the decision. For example, in the context of pedestrian store or shopping-centre choice behaviour, a pedestrian's utility of alternatives is commonly defined as the summation of weighted (evaluation of) physical factors such as store floorspace, store type and variety, distance, traffic condition, and parking facility (see, for example, Borgers and Timmermans, 1986; Oppewal and Timmermans, 1997; Saito and Ishibashi, 1992; Van der Waerden et al, 1998). Sociodemographics can also be incorporated into the utility function to represent their influences (see, for example, Fortheringham and Trew, 1993). Pedestrians are assumed to choose the store bringing them the highest utility. Hoogendoorn and Bovy (2004) developed an extensive framework based on expected utility maximisation to model not only pedestrian route choice but also optimal activity schedule, trajectory, and speed. In terms of microscopic movement, Antonini and Bierlaire (2006) used a multinomial logit model to represent how pedestrians choose local walking directions that are defined as discrete surrounding radial regions. A potential limitation of these models is that they do not explicitly represent decision processes such as information search and representation. Pedestrians are often assumed to have very good knowledge about the environment and the ability to take into account all relevant factors influencing decisions.

In contrast to these models of rational choice behaviour, models of bounded rationality assume that decisions are made on subsets of factors and do not necessarily result in optimal choices. These models often adopt noncompensatory rules and decision heuristics to indicate that decisions may be made on an attribute-by-attribute basis as opposed to a compensatory decision process (see, for example, Gigerenzer et al, 1999; Payne and Bettman, 1988; Tversky, 1972). A few examples of such models have appeared in planning and transportation literature. Often, how well a noncompensatory choice model predicts observed choices has been tested. Foerster (1979) is among the early researchers who attempted to introduce noncompensatory models such as lexicographic, conjunctive, and disjunctive rules into travel-mode decision-process modeling. The potential advantage of noncompensatory models was suggested by positive results. Recker and Golob (1979) proposed an elimination-based model based on sequential consideration of attributes. Applying the model to the problem of vehicle choice and store choice, they estimated critical tolerances for attributes, assuming the sequence of attribute consideration is known. Young (1984) assumed that individuals have a set of minimally acceptable satisfaction levels which are used to judge the satisfactoriness of corresponding attributes. These tolerances were estimated, while attribute importance was provided by respondents, using rating scales. Borgers et al (1986) suggested a hybrid model for residential preferences. They estimated a model which assumes that individuals apply certain thresholds to particular attributes (noncompensatory part), while the remaining attributes are processed in a compensatory fashion. In contrast to the general support for noncompensatory models, Timmermans (1983) found less evidence for the superiority of noncompensatory decision rules. In his study of shopping-centre choice, compensatory and noncompensatory decision rules performed almost equally well.

This brief discussion of the literature suggests that research on models of bounded rationality in planning is relatively scarce. Although existing research has provided some degree of the potential value of exploring further the development of bounded rationality models, the state of the art indicates some shortcomings. First, the models are limited in the sense that they focus on the noncompensatory nature of the decision rule. The process that leads to the selection of factors entering the decision process is not usually modeled. Common practice is that factor selection and search sequence are a priori assumed or obtained from the direct report of respondents. Second, the choice of heuristics to be applied depends largely on the researcher's individual experiences and intuition. Some general approach would be helpful to overcome, to some extent, the arbitrariness involved in this selection process. For example, Swait (2001) shows that incorporating attribute cutoffs and varying utility functions into conventional logit models can approximate disjunctive and conjunctive rule, which means the heuristics can be estimated. However, approximation is not exact; the information search process cannot be identified and cutoffs are self-reported. Third, a particular heuristic model is assumed to apply to all individuals. In reality, however, it is unlikely that different individuals will use the same choice rules. Even for the same individual, decision strategies may be dependent on context. Evidence of such behaviour has accumulated over a long period (see, for example, Beach and Mitchell, 1978; Payne, 1982). What is still missing is a fully operational approach that accounts for heuristic heterogeneity, which would thus enhance models of bounded rationality.

Therefore, we propose a modeling approach that overcomes the limitations discussed above and requires no more information than is required by conventional discrete choice models, while it is based on richer behavioural assumptions, including that: (1) attribute thresholds are incorporated in the utility function as the basic cognitive mechanism; (2) heterogeneous decision heuristics can be exactly identified; (3) mental effort and risk perception of heuristics are defined and their influences on choice of heuristic are estimated. To illustrate this, the approach is applied to the go-home decision of shopping pedestrians, which predicts the duration of pedestrians' shopping trips. It could potentially also be applied to different kinds of decision problems.

The paper is organised as follows. First, we elaborate on the theoretical framework underlying the approach. Next, we discuss the data collection, followed by a discussion of the main findings of the model estimation. The paper is completed with a discussion and conclusions.

2 Conceptual framework

2.1 Preference structure

On the basis of the principle of bounded rationality, any decision or choice process can be understood as a problem-solving process in which an individual processes information to arrive at a decision that achieves a particular goal within some margin of accuracy. We assume that individuals will construct a mental representation of the decision problem. This cognitive process is assumed to consist of at least three sequential processes: filtering of information, factor representation into states, and judging the resulting states, individually and combined. Jointly, these processes lead to preference formation.

Let $X = \{x_j \mid j = 1, 2, ..., J\}$ represent the set of factors (or attributes) influencing the decision of interest. We assume that individuals do not necessarily take all these factors into account, but mentally construct or reconstruct the problem and select a subset of the 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), will all factors be considered. Mathematically, this can be expressed as:

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

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

Once the relevant factors have been filtered, bounded rationality suggests that individuals tend not to discriminate between all possible values of factors. Rather, they will categorise the continuous factors into discrete classes or states, or recategorise discrete factors. We assume that in the case of continuous factors, this process of factor representation involves the application of a monotonically increasing set of threshold values that discretise the continuous factors into an ordered set of discrete classes. Let $\Delta_j = \{\delta_{j1} < \delta_{jn} < \delta_{jN} \,|\, n = 1, 2, ..., N\}$ 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 representational simplicity, the subscript j is ignored.) A factor may then meet one or more of these increasingly stricter activation thresholds and hence the stimulus becomes stronger. The relevant equations then become

$$s_{jn} = \begin{cases} 0, & \text{if } x_j < \delta_{jn} , \\ 1, & \text{if } x_j \geqslant \delta_{jn} , \end{cases} \tag{2}$$

$$X' = \{x_j \,|\, s_{jn} = 1\} .$$

Thus, filtering and factor representation transforms categorical and continuous external factors into a set of activated and nonactivated 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 and making the decision. Attaching judgment values to states implies that the state is evaluated. Weights indicate the relative importance of states. In our approach, these values and weights are combined into a single value, w_{jn}, which can be interpreted as a part-worth utility, because they are both unknown parameters. 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. To facilitate operation, 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, we assume that the overall values are also categorised by checking them against a set of $\Lambda = \{\lambda_1 < \lambda_m < \lambda_M \,|\, m = 1, 2, ..., M\}$ of successively increasing overall thresholds, resulting in the overall states, p_{im}. This can be expressed as:

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

In case this representation involves only two preference orders (for example, reject or accept), only one λ is needed and $p_i = 0$ defines rejecting the alternative, whereas $p_i = 1$ implies accepting it. For representational simplicity, in the remainder of this paper, we will assume that only two orders 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 $\bar{v}_k$ represent any factorial combination from value judgments in the sets, that is,

$$\bar{v}_k = \sum_j v_{j1}, \qquad l \in [1, ..., N_j + 1]. \tag{6}$$

Ordering all the $\bar{v}_k$ in ascending order forms an overall value set,

$$\bar{V} = \left\{\bar{v}_1 < \bar{v}_k < \bar{v}_K | k = 1, 2, ..., K; \quad K = \prod_j (N_j + 1)\right\}. \tag{7}$$

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 $\bar{V}$ can be divided into a subset $\bar{V}_0$ of rejected overall value judgments and a set $\bar{V}_1$ 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,

$$\Phi = \left\{\begin{matrix} \bar{v}_k \in \bar{V}_0 | \bar{v}_k < \lambda \\ \bar{v}_k \in \bar{V}_1 | \bar{v}_k \geqslant \lambda \end{matrix}\right\}. \tag{8}$$

2.2 Decision heuristics

We assume that in every choice context, individuals will consciously or unconsciously define a set of threshold values and apply decision 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 decision heuristics. One extreme is the strictest preference structure in the sense that no single value (judgment) combination survives the overall threshold,

$$\Phi = \{\bar{v}_k \in \bar{V}_0 | \bar{v}_k < \lambda\}. \tag{9}$$

That means that, regardless of the states of the factors, the choice alternative under evaluation will be rejected. In this case, no information search is implied (or the heuristic of 'no action', since the individual does not need to consider any information). Relaxing λ a little leads to a preference structure where only the value combination of factor states with the highest threshold values is accepted,

$$\Phi = \left\{\begin{matrix} \bar{v}_k \in \bar{V}_0 | \bar{v}_k < \lambda \\ \bar{v}_k \in \bar{V}_1 | \bar{v}_k \geqslant \lambda, \bar{v}_k = \sum_j v_{jN+1} \end{matrix}\right\}. \tag{10}$$

This preference structure implies a conjunctive heuristic 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, representing the case that all factor combinations will be accepted,

$$\Phi = \{\bar{v}_k \in \overline{V}_1 | \bar{v}_k \geqslant \lambda\}. \tag{11}$$

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 nonactivated factor states are accepted,

$$\Phi = \left\{ \begin{array}{l} \bar{v}_k \in \overline{V}_0 | \bar{v}_k < \lambda, \bar{v}_k = \sum_j v_{j1} \\ \bar{v}_k \in \overline{V}_1 | \bar{v}_k \geqslant \lambda \end{array} \right\}. \tag{12}$$

Disjunctive heuristics can be inferred from this preference structure since any factor state (except the nonactivated 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 such a preference structure,

$$\Phi = \left\{ \begin{array}{l} \bar{v}_k \in \overline{V}_0 | \bar{v}_k < \lambda, \sum_k \sum_{t=1}^{n} s_{jt|k} = 0 \\ \bar{v}_k \in \overline{V}_1 | \bar{v}_k \geqslant \lambda, \prod_k \prod_{t=n'}^{N} s_{jt|k} = 1 \end{array} \right\} \text{for } n < n'. \tag{13}$$

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 of other factors is needed.

2.3 Choice of heuristics

We assume that individuals in different contexts may apply different preference structures and corresponding decision heuristics to solve problems. That is, individuals will 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 included mathematically in the overall threshold, so that we get $\lambda \sim f$, where f is a probability density function. Because $\overline{V}$ is a discrete set, between consecutive pairs of $\bar{v}_k$, there is a range of λ, satisfying $\bar{v}_{k-1} < \lambda \leqslant \bar{v}_k$. 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:

$$p_k = \int_{\bar{v}_{k-1}}^{\bar{v}_k} f \, \mathrm{d}t. \tag{14}$$

Equivalently, we may view this as the probability of applying decision heuristics implied by the preference structure. To elaborate, assume that the notion of bounded rationality implies that individuals cannot, or may not, feel the need to discriminate to the fullest extent between the identified preference structures. Preference structures with similar factor relaxation can be treated as similarly effective in satisfying a particular need. We called this the preference tolerance, represented by a set of values $\Gamma = \{\gamma_g | \, g = 1, 2, ..., G; \, G \leqslant K\}$, grouping preference structures into $G + 1$ subsets according to $\overline{V}$,

$$Z = \begin{Bmatrix} z_1 \,|\, \bar{v}_k < \gamma_1 \\ z_2 \,|\, \gamma_1 \leqslant \bar{v}_k < \gamma_2 \\ \ldots \\ z_{g'} \,|\, \gamma_{g'-1} \leqslant \bar{v}_k < \gamma_{g'} \\ \ldots \\ z_{G+1} \,|\, \gamma_G \leqslant \bar{v}_k \end{Bmatrix}. \tag{15}$$

In the case of high-involvement decisions, one would expect strict tolerances and potentially many subsets. In contrast, in the case of low involvement, preference structures may be grouped into a small number of broad classes. The process of selecting a decision heuristic is a two-stage process, starting with an individual applying context-specific preference tolerances, followed by selecting a decision heuristic within the relevant preference-structure subset. Let $h = 1, 2, ..., H$ be the symbol for heuristics. We assume that within each subset, the choice of a heuristic can be modeled probabilistically based on the expected value of applying each heuristic,

$$p_{h \in z_{g'}} = p_{z_{g'}} \frac{\exp u_h}{\sum_{h' \in z_{g'}} \exp u_{h'}}. \tag{16}$$

The probability of every subset being selected, $p_{z_{g'}}$, can be derived empirically from equation (14) and is equal to the sum of the probabilities of preference structures within that set, given the estimated preference tolerances. The expected value of a heuristic, u_h, may be composed of various factors. Here we consider two of those factors. The first is the expected amount of mental effort, e_h, which is defined to be inflicted during searching a factor, considering the factors into the decision, representing their states, and attaching values. The second factor is risk perception, r_h, which represents subjective outcome diversity.

A complicating aspect is that individuals cannot be sure about the amount of mental effort that could be involved before making the decision. They can only estimate it subjectively based on their beliefs p_{jn} that the factors occupy states that make any further searching of subsequent factors useless. To illustrate, let three factors x_1, x_2, and x_3 have, A, B, and C states, respectively, ($a = 1, 2, ..., A$; $b = 1, 2, ..., B$; $c = 1, 2, ..., C$). Assume that the heuristic under consideration implies the search sequence $x_1 \rightarrow x_2 \rightarrow x_3$. Let e_1, e_2, and e_3 denote the amount of mental effort inflicted when considering factor 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 overall 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,

$$e_h = e_1 + \sum_a \Big(p_a e_2 I_{ab} + \sum_b p_a p_b e_3 I_{abc} \Big), \tag{17}$$

$$I_{ab} = \begin{cases} 0, & \text{if } v_{abc} < \lambda \vee v_{abc} \geqslant \lambda, \qquad \forall b, \forall c, \\ 1, & \text{otherwise}, \end{cases} \tag{18}$$

$$I_{abc} = \begin{cases} 0, & \text{if } v_{abc} < \lambda \vee v_{abc} \geqslant \lambda, \qquad \forall c, \\ 1, & \text{otherwise}. \end{cases} \tag{19}$$

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, the expected effort is 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 I_{ab}, an identity function defined by equation (18). I_{ab} represents a judgment process, in which an individual checks whether all possible value combinations given the current factor state, $v_{abc} | a$, are inactivated against λ, or are all activated. If all value combinations are inactivated, the unsearched factor is not considered and no additional mental effort is involved. If all factors are activated, it means 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 case, $I_{ab} = 0$; in contrast, $I_{ab} = 1$, and x_2 needs to be searched. According to the same logic, the second term relates to searching x_3 when at a state of x_2. Effort e_3 is weighted by $p_a p_b$, the joint probability of being in the previous two factor states, and I_{abc} is another identity function judging whether the simultaneous conditions $v_{abc} | a, b$ against λ are satisfied or not. By this definition, due to the fact that the effort for searching factors may differ and different factor values may cause earlier or later termination of the decision process when the expected overall values are homogeneous against the overall threshold, the expected effort of consideration sequences may differ also.

Preference structures with either very strict thresholds or very relaxed thresholds are defined to be risky because most information about factors falls into the nonactivated states or the activated states, respectively. This makes the decision of rejecting or accepting alternatives very certain, but increases the potential risk of false rejection or false acceptation, respectively. Thus, outcome diversity and risk perception are inversely related. Risk perception is defined in terms of Shannon's information entropy measure. Let $\overline{p}_k$, corresponding to $\overline{v}_k$, be the factorial joint product of the probabilities of factor states. The probability of getting a positive, r_h^+ or a negative, r_h^-, outcome is, respectively,

$$r_h^+ = \sum_k \overline{p}_k , \quad \forall \overline{v}_k \geqslant \lambda , \tag{20}$$

$$r_h^- = 1 - r_h^+ .$$

Then, the risk perception of a heuristic is equal to

$$r_h = -r_h^+ \ln r_h^+ - r_h^- \ln r_h^- . \tag{21}$$

Because the sequence of factor search does not influence r_h, the risk perceptions are the same for heuristics of the same preference structure. Thus, risk perception only differentiates the values of heuristics of different preference structures.

In total, we assume that the value of a decision heuristic equals some weighted linear trade-off between mental effort and risk perception,

$$u_h = \beta_e e_h + \beta_r r_h . \tag{22}$$

3 Illustration

3.1 Data

We will illustrate the suggested modeling approach in the context of pedestrian go-home decisions, which determine the total duration of pedestrian's shopping trips. The data used were collected in May 2004 as part of a pedestrian survey in Wang Fujing Street, in the city centre of Beijing, China. Twenty students from the Department of Urban and Regional Planning, Peking University administered the survey by randomly asking pedestrians, who were near the end of their shopping trips, two categories of questions: (1) respondents' sociodemographics; (2) their sequential set of stops and activities in the street. The total sample consists of 694 respondents. Except for the final stop, the data do not provide any explicit information about the go-home decision. Because we could not observe when go-home decisions were made, we assumed that this decision is made implicitly or explicitly every time a store has been visited and an additional approach was applied to estimate the time of decisions (Zhu et al, 2006). This resulted in 2741 decision cases for model estimation.

3.2 Operationalisation

The go-home decision can be influenced by many factors such as the fulfillment of the shopping list, feeling tired, or turning to other scheduled activities. Not all of them are easy to measure in surveys. Here, we use time as the major factor influencing this decision. Two types of time are distinguished: relative time, t^{R}, and absolute time, t^{A}. Time t^{R} refers to the time elapsed, in minutes, since the pedestrian started the shopping trip. It correlates with the progress of purchasing the planned items during the shopping trip, visiting schedules, and how tired the pedestrian has become. Time t^{A} refers to the time difference between the current time and the base time 0:00. It correlates with available time budgets, reflecting when pedestrians must turn to other business. It should be noted that other factors may need to be included in future models. In this study, however, our primary focus was on developing the modeling principles and thus we kept the number of explanations limited.

Let the threshold values for t^{R} be $\boldsymbol{\Delta}^{\mathrm{R}} = [\delta_1^{\mathrm{R}}, ..., \delta_m^{\mathrm{R}}, ..., \delta_M^{\mathrm{R}}]$, a row vector with M elements, and the threshold values for t^{A} be $\boldsymbol{\Delta}^{\mathrm{A}} = [\delta_1^{\mathrm{A}}, ..., \delta_n^{\mathrm{A}}, ..., \delta_N^{\mathrm{A}}]$, a row vector with N elements. $\boldsymbol{W}^{\mathrm{R}} = [w_1^{\mathrm{R}}, ..., w_m^{\mathrm{R}}, ..., w_M^{\mathrm{R}}]^{\mathrm{T}}$ and $\boldsymbol{W}^{\mathrm{A}} = [w_1^{\mathrm{A}}, ..., w_n^{\mathrm{A}}, ..., w_N^{\mathrm{A}}]^{\mathrm{T}}$ are column vectors for corresponding state values. An unobserved factor ε with a standard normal distribution is also assumed [since all terms in equation (23) have free parameters nonunit standard deviations can be reduced to unit standard deviation], while λ represents the overall threshold. The go-home decision model therefore is,

$$I(t^{\mathrm{R}} \geqslant \boldsymbol{\Delta}^{\mathrm{R}})\boldsymbol{W}^{\mathrm{R}} + I(t^{\mathrm{A}} \geqslant \boldsymbol{\Delta}^{\mathrm{A}})\boldsymbol{W}^{\mathrm{A}} + \varepsilon \geqslant \lambda\,, \tag{23}$$

where $I(\zeta)$ is an element-wise identity function assigning 1 when the relation ζ is true, and assigning 0 when ζ is false. Literally, if the sum of the observed and unobserved values exceeds the overall threshold, the pedestrian will decide to go home. The probability of the observed utility under the normal distribution of mean λ and unit standard deviation, equivalently the probability of going home p^{H}, can be expressed as,

$$p^{\mathrm{H}} = F[I(t^{\mathrm{R}} \geqslant \boldsymbol{\Delta}^{\mathrm{R}})\boldsymbol{W}^{\mathrm{R}} + I(t^{\mathrm{A}} \geqslant \boldsymbol{\Delta}^{\mathrm{A}})\boldsymbol{W}^{\mathrm{A}}, \lambda, 1]\,, \tag{24}$$

where F represents the cumulative density function of the normal distribution. Parameters can be estimated in terms of a maximum likelihood estimator. The numbers of threshold values M and N were not set a priori but estimated as part of the estimation process. Because the model involves a discontinuous multidimensional step function, the likelihood function is also discontinuous, multipeaked, and difficult to solve analytically. Therefore, a hybrid estimation algorithm composed of a genetic algorithm for global search and a quasi-Newtonian algorithm for local search provided by MATLAB was used. To prevent overestimation, the consistent Akaike information criteria (CAIC) (see, for example, Dayton and Lin, 1997) was applied to select a well-fitted but relatively parsimonious model. Several models with different combinations of threshold numbers were tried. The model with the lowest CAIC was finally selected.

3.3 Preference structures

Table 1 shows the estimation results for the preference structures. The estimated number for $\boldsymbol{\Delta}^{\mathrm{R}}$ is 2, implying that pedestrians categorise relative time (in minutes)

Table 1. Model estimation results. (See text for definitions of symbols.)

Proposed model		Multinomial logit model		Mixed logit model	
parameter	estimate	parameter	estimate	parameter	estimate
δ_1^{R}: threshold 1 for t^{R}	90	β^{R}: for $\ln t^{\mathrm{R}}$	-1.4705	$\beta_{\mathrm{m}}^{\mathrm{R}}$: mean of β^{R}	-14.3453
δ_2^{R}: threshold 2 for t^{R}	180	β^{A}: for $\ln t^{\mathrm{A}}$	-8.5103	$\beta_{\mathrm{sd}}^{\mathrm{R}}$: standard deviation of β^{R}	0.9074
w_1^{R}: value for state 1 of t^{R}	0.8957	β^{H}: for dummy variable	-67.1172	$\beta_{\mathrm{m}}^{\mathrm{A}}$: mean of β^{A}	-83.9423
w_2^{R}: value for state 2 of t^{R}	0.6764			$\beta_{\mathrm{sd}}^{\mathrm{A}}$: standard deviation of β^{A}	1.8076
δ_1^{A}: threshold 1 for t^{A}	840			$\beta_{\mathrm{m}}^{\mathrm{H}}$: mean of β^{H}	-661.3274
δ_2^{A}: threshold 2 for t^{A}	960			$\beta_{\mathrm{sd}}^{\mathrm{H}}$: standard deviation of β^{H}	11.2442
δ_3^{A}: threshold 3 for t^{A}	1140				
w_1^{A}: value for state 1 for t^{A}	1.1826				
w_2^{A}: value for state 2 for t^{A}	0.8374				
w_3^{A}: value for state 3 for t^{A}	0.7065				
λ: overall threshold	3.3883				
q^C: number of observations	2741		2741		2741
q^P: number of parameters [a]	6		3		6
LL0: log-likelihood of null model	-1900		-1900		-1900
LL: log-likelihood	-1036		-1085		-1079
CAIC	2113		2197		2211

[a] Thresholds are not counted as free parameters because their values control the effectiveness of the states.

into three discrete states: [<90, 90–180, >180). The number for Δ^{A} is 3, implying four discrete states for absolute time (in minutes) [<840, 840–960, 960–1140, >1140). The threshold values show surprising regularity. The thresholds for relative time are about 1.5 hours and 3 hours. The thresholds for absolute time are 14:00, 16:00, and 19:00. They conform to people's habit of using critical time spots as decision references. Value judgments were bounded to positive values in the estimation procedure under the assumption that pedestrians' inclination to go home should increase with time. Value estimates do not show any conflicts with this assumption because conflicts should have generated some values close to 0. A nonlinear relationship between time and the impetus to go home is suggested as the increase in the values of each factor slows down as time passes.

Additionally, two discrete choice models were estimated for comparison. The first is a multinomial logit model with a linear utility function, which is,

$$\begin{aligned} p^{\mathrm{H}} &= \frac{\exp u^{\mathrm{H}}}{\exp u^{\mathrm{H}} + \exp u^{\mathrm{S}}}, \\ u^{\mathrm{S}} &= \beta^{\mathrm{R}} \ln t^{\mathrm{R}} + \beta^{\mathrm{A}} \ln t^{\mathrm{A}}, \\ u^{\mathrm{H}} &= \beta^{\mathrm{H}}, \end{aligned} \tag{25}$$

where the observable utility of shopping, u^{S}, is a function of natural-logged time variables, representing the marginally decreasing utility with the increase in time. The utility of going home, u^{H}, is represented by an alternative-specific constant, β^{H}. The second is a mixed logit model, which is gaining increasing popularity. The model assumes that parameters are distributions, which may capture more heterogeneity in people's utility functions. However, the proposed model outperforms both discrete choice models in terms of log-likelihood and CAIC.

3.4 Decision heuristics

By identifying the combinations of activated and nonactivated states of t^{R} and t^{A} and summing their corresponding value judgments, we can derive the overall judgment value for each of the twelve factor–state combinations.

Figure 1 portrays the combinations as a tree structure, and the value combinations, $\bar{v}_k$, are ranked in ascending order according to their values. For example, $\bar{v}_5$ consists of state 1 for t^{R} and state 3 for t^{A}, keeping in mind the principle of accumulating thresholds, and thus equals $0+0+1.18+0.84 = 2.02$. By aggregating the random factor and the overall threshold, we have $\lambda \sim \mathrm{N}(3.39, 1)$, from which different decisions heuristics can be derived. For example, if $\bar{v}_{11} < \lambda \leqslant \bar{v}_{12}$, only $\bar{v}_{12}$ can be accepted. Pedestrians will then go home only if t^{R} is in state 3 and t^{A} is in state 4. The preference structure is,

$$\phi_{12} = \left\{ \begin{array}{l} \bar{v}_k \in \bar{V}_0 | \bar{v}_k < 4.29 \\ \bar{v}_k \in \bar{V}_1 | \bar{v}_k \geqslant 4.29 \end{array} \right\}. \tag{26}$$

This represents two conjunctive heuristics, one starts from searching t^{R} and the other from t^{A}. Pedestrians will continue shopping if either t^{R} is not in state 3 or t^{A} is not in state 4. Similarly, disjunctive heuristics are observed when $\bar{v}_1 < \lambda \leqslant \bar{v}_2$, with the preference structure,

$$\phi_2 = \left\{ \begin{array}{l} \bar{v}_k \in \bar{V}_0 | \bar{v}_k < 0.89 \\ \bar{v}_k \in \bar{V}_1 | \bar{v}_k \geqslant 0.89 \end{array} \right\}. \tag{27}$$

Pedestrians will decide to go home if t^{R} is in state 2 or 3 when searching t^{R} first or t^{A} is in state 2, 3 or 4 when searching t^{A} first. Figure 2 shows the decision tree of these two heuristics.

If we just make the overall threshold a little bit stricter so that $\bar{v}_4 < \lambda \leqslant \bar{v}_5$, a lexicographic heuristic appears. It is implied by the preference structure,

$$\phi_5 = \left\{ \begin{array}{l} \bar{v}_k \in \bar{V}_0 | \bar{v}_k < 2.02 \\ \bar{v}_k \in \bar{V}_1 | \bar{v}_k \geqslant 2.02 \end{array} \right\}. \tag{28}$$

Figure 3 shows the two heuristics that are consistent with this preference structure. Different from the above preference structures, however, the lexicographic heuristic is valid only when t^{A} is searched first. That is, pedestrians can decide to go home if t^{A} is in state 3 or 4, and decide to continue shopping if t^{A} is in state 1. They need to

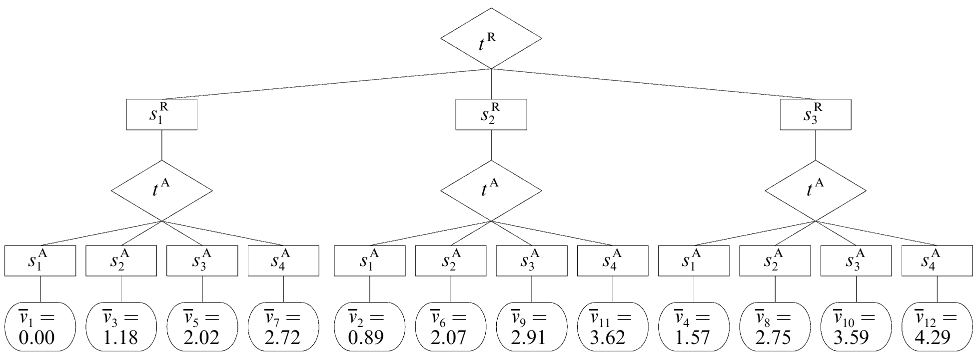

Figure 1. Factor – state combinations underlying the go-home decision. See text for definitions of symbols.

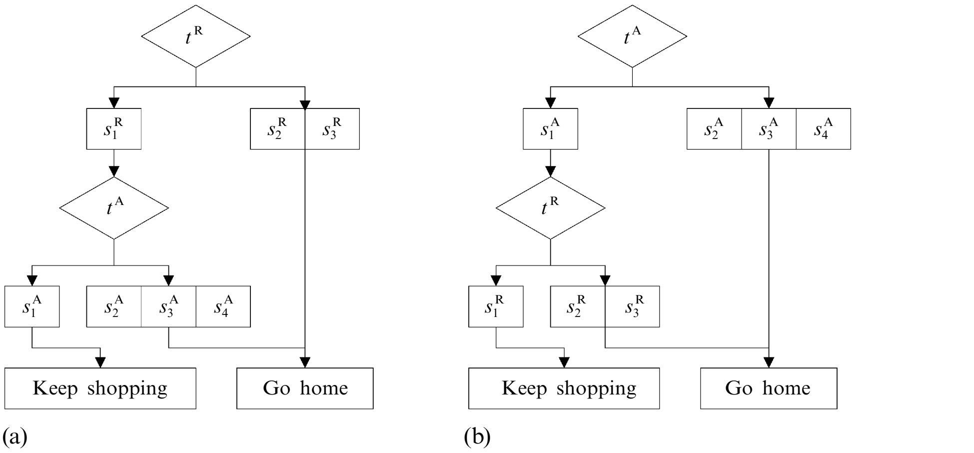

Figure 2. Two disjunctive heuristics: (a) searching for t^R first; (b) searching for t^A first. (See text for definitions of symbols.)

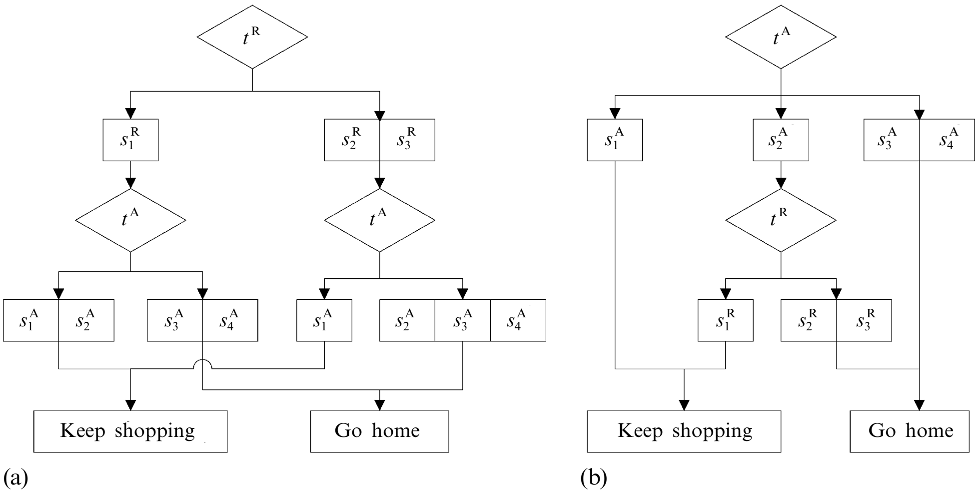

Figure 3. Lexicographic heuristic when considering t^A first: (a) nonlexicographic when searching t^R first; (b) lexicographic when searching t^A first. (See text for definitions of symbols.)

consider t^R if t^A is in state 2. The lexicographic interpretation does not hold when t^R is searched first because t^R, regardless of its state, will not generate a definite outcome.

3.5 Choice of heuristics

To illustrate calculating the probabilities of preference structures and implied heuristics, consider the above conjunctive heuristics. These heuristics are invariant within an overall threshold range $3.62 < \lambda \leqslant 4.92$, which corresponds to a probability 0.226 under the normal distribution $\lambda \sim \mathrm{N}(3.39, 1)$. This suggests that there is 22.6% probability that a pedestrian applies these specific conjunctive heuristics. The probabilities of the thirteen preference structures are shown as white bars in figure 4 (including the preference structure that no factor states are activated, ϕ_{13}, and the preference structure that all factor states are activated, ϕ_1). As can be seen from the figure, the probabilistic distribution zigzags as the strictness of preference structures increases.

The heuristic choice model implies estimating the following parameters: (1) the preference tolerance $\Gamma = \{\gamma_g \,|\, g = 1, 2, ..., G;\ G \leqslant 12\}$, which is used to derive $p_{z_{g'}}$ of

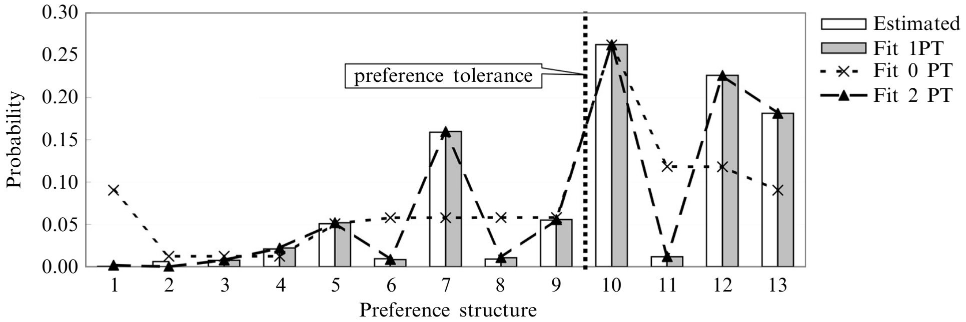

Figure 4. Probabilities of preference structures. PT = preference tolerance.

each preference structure group; (2) e^{R} and e^{A}, the mental effort of searching e^{R} and e^{A}; (3) $P^{\mathrm{R}} = [p_1^{\mathrm{R}}, ..., p_m^{\mathrm{R}}, ..., p_M^{\mathrm{R}}]$ and $P^{\mathrm{A}} = [p_1^{\mathrm{A}}, ..., p_n^{\mathrm{A}}, ..., p_N^{\mathrm{A}}]$, the probability beliefs of the pedestrians regarding the factors states; and (4) β_h, the parameter for r_h. The estimation procedure uses the sum of squared error as the goodness-of-fit criterion and the hybrid algorithm, described earlier.

Three models with preference tolerances 0, 1, and 2, respectively, were fitted to the thirteen preference structure probabilities. Figure 4 shows that the 0-tolerance model performs poorly. In contrast, the 1-tolerance and 2-tolerance threshold models fit the data almost perfectly. Trading-off between goodness of fit and model parsimony, we selected the 1-tolerance threshold model and show the results in table 2. The model implies that the preference structures are differentiated into two groups, split by preference tolerance at 10. Group 1 includes ϕ_1 to ϕ_9, with a total choice probability 31.8%, which can be interpreted as relatively relaxed preference structures, and group 2 includes ϕ_{10} to ϕ_{13} with a total choice probablity 68.2%, which can be interpreted as relatively strict preference structures. The signs of the amount of mental effort for both factors are negative, as hypothesized. Searching t^{A} seems to cost almost five times as much effort as searching t^{R}, probably because searching absolute time requires more frequent, external references such as checking one's watch or personal schedule. In contrast, searching relative time may depend more on internal references, such as effort consumption and need fulfillment, which can be instantly felt subconsciously. As for the probability belief of factor states, it seems that pedestrians' beliefs that the decision will be made within the first half hour of the shoppig trip is very weak,

Table 2. Results of the heuristic choice model. (See text for definitions of symbols.)

Parameter	Estimate
e^{R}: effort of t^{R}	−11.7149
p_0^{R}: probability for state 1 of t^{R}	0.0635
p_1^{R}: probability for state 2 of t^{R}	0.5481
p_2^{R}: probability for state 3 of t^{R}	0.3884
e^{A}: effort of t^{A}	−53.1271
p_0^{A}: probability for state 1 of t^{A}	0.3350
p_1^{A}: probability for state 2 of t^{A}	0.1377
p_2^{A}: probability for state 3 of t^{A}	0.2215
p_3^{A}: probability for state 4 of t^{A}	0.3058
β_h: parameter for r_h	63.2634
γ: preference tolerance	10
R^2	0.9996

as p_1^{R} is only 6.4%. It reflects a tendency that pedestrians are not quite willing to make very relaxed decisions which probably lead to early going-home decisions and deter them from enjoying more shopping. Parameter β_h turns out to be positive, suggesting that pedestrians are risk averse and decision heuritics leading to diversified outcomes are preferred, given the same effort. Overall, preference structures with large overall thresholds are probably chosen more than those with small overall thresholds, suggesting that saving effort and making quick decisions is more important than reducing decision risk in the go-home decision.

Table 3 shows the estimated amount of values, and the corresponding probabilities of heuristics. In group 1, heuristics searching t^{A} first have a six times higher total probability than heuristics searching t^{R} first, even though e^{A} is much larger. Excluding the influence of risk perception, such a relationship generally still holds. This means that pedestrians' beliefs about a factor state play a more important role. In less strict decisions, a heuristic starting with more mental effort but reaching a decision quicker is still preferable to a heuristic starting with less mental effort but taking more steps to reach an outcome. When the thresholds become stricter, as in group 2, t^{R} receives more attention and the total probability of considering t^{R} first becomes equal to the probability of considering t^{A} first.

Table 3. Estimated heuristic utilities and probabilities. See text for definitions of symbols.

Group	Preference structure	Heuristic utility		Heuristic probability	
		$t^{\mathrm{R}} \rightarrow t^{\mathrm{A}}$	$t^{\mathrm{A}} \rightarrow t^{\mathrm{R}}$	$t^{\mathrm{R}} \rightarrow t^{\mathrm{A}}$	$t^{\mathrm{A}} \rightarrow t^{\mathrm{R}}$
1	1	0	0	0.0009	0.0010
	2	−5.6933	−47.6631	<0.0001	<0.0001
	3	2.0786	−10.7615	0.0077	<0.0001
	4	3.1386	−11.3146	0.0221	<0.0001
	5	−6.1074	3.9926	<0.0001	0.0520
	6	−5.3176	2.1876	<0.0001	0.0085
	7	−2.3892	5.1159	0.0001	0.1598
	8	1.3868	1.9402	0.0038	0.0067
	9	1.7888	3.9552	0.0057	0.0501
Group sum				0.0404	0.2781
2	10	−1.2116	0.9548	0.0270	0.2355
	11	−6.8118	−2.0506	0.0001	0.0117
	12	0.9144	−23.4541	0.2262	<0.0001
	13	0	0	0.0906	0.0906
Group sum				0.3439	0.3378

4 Discussion and conclusions

In this paper, we have suggested a modeling approach based on principles of bounded rationality. In particular, the model is based on the assumption that (1) individuals may not necessarily take into consideration all factors potentially influencing their choice behaviour, (2) that individuals do not necessarily discriminate between all values of factors but rather group these into discrete factor states, (3) that individuals judge the combination of activated and nonactivated factor states against some overall acceptance threshold. We have shown that by adapting this framework and estimating the various thresholds and number of factor states, a flexible model is obtained with some interesting properties.

First, the results of the model can indicate the extent to which observed behavioural patterns can be explained in terms of rational choice behaviour. Evidence of rational choice behaviour would be obtained if estimated thresholds are such that all factors are taken into account, the number of states for each factor is high as this would indicate detailed discrimination, and the overall threshold is high, implying that only the choice rule with the highest value judgement is accepted. Any deviations from this outcome, in contrast, would support aspects of bounded rationality. Second, heterogeneous decision heuristics are automatically depicted. To further qualify this conclusion, discrete choice models typically (if at all) depict consumer heterogeneity in terms of estimate distributions around the parameters of one and the same utility function and hence do not allow for different utility specifications as part of the same model. Third, it acknowledges the fact that different decision heuristics may be equally satisfying, and that the ultimate choice among these equivalent heuristics depends on mental and risk perception. It reflects the notion of selective information processing and context dependency. Fourth, the illustration of the approach on pedestrians' go-home decision showed evidence of face validity. The overall model fit outperformed that of two typical discrete choice models and parameter estimates can be explained reasonably.

As a general decision model, this approach can be useful for understanding the process of decision making in other contexts, while it requires no more data than are needed to estimate conventional discrete choice models. With the knowledge of people's decision strategies and their focus and representation of information, better predictions can be made; information can be provided in more efficient forms. Undoubtedly, to assure the potential advantages of the approach, more concrete tests are required. It will be interesting to compare the estimated heuristics with data collected from self-reported decision protocols or controlled experiments where information is presented differently to respondents.

Acknowledgements. The authors thank De Wang of the Department of Urban Planning, Tongji University and Yanwei Chai of the Department of Urban and Regional Planning, Peking University, for their cooperation during the data collection.

References

Antonini G, Bierlaire M, Weber M, 2006, "Discrete choice models of pedestrian walking behaviour" *Transportation Research B* **40** 667 – 687

Beach L R, Mitchell T R, 1978, "A contingency model for the selection of decision strategies" *Academy of Management Review* **3** 439 – 449

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** 115 – 128

Borgers A, Timmermans H J P, Veldhuisen J, 1986, "A hybrid compensatory – noncompensatory model of residential preference structures" *Journal of Housing and Environmental Research* **1** 227 – 234

Dayton C M, Lin T H, 1997, "Model selection information criteria for non-nested latent class models" *Journal of Educational and Behavioral Statistics* **22** 249 – 264

Dijkstra J, Timmermans H, 2002, "Towards a multi-agent model for visualizing simulated user behaviour to support the assessment of design performance' *Automation in Construction* **11** 135 – 145

Foerster J F, 1979, "Mode choice decision process models: A comparison of compensatory and non-compensatory structures" *Transportation Research A* **13** 17 – 28

Fortheringham A S, Trew R, 1993, "Chain image and store-choice modeling: the effects of income and race" *Environment and Planning A* **25** 179 – 196

Gigerenzer G, Todd P M, ABC Research Group, 1999 *Simple Heuristics that Make Us Smart* (Oxford University Press, Oxford)

Haklay M, O'Sullivan D, Thurstain-Goodwin M, Schelhorn T, 2001, " 'So go downtown': simulating pedestrian movement in town centres" *Environment and Planning B: Planning and Design* **28** 343 – 359

Helbing D, Molnár P, 1995, "Social force model for pedestrian dynamics" *Physical Review* E51, 4282 – 4286

Helbing D, Molnár P, Farkas I J, Bolay K, 2001, "Self-organizing pedestrian movement" *Environment and Planning B: Planning and Design* **28** 361 – 383

Hoogendoorn S P, Bovy P H L, 2004, "Pedestrian route-choice and activity scheduling theory and models" *Transportation Research B* **38** 169 – 190

Hoogendoorn S P, Daamen W, Landman R, 2007, "Microscopic calibration and validation of pedestrian models: cross-comparison of models using experimental data", in *Pedestrian and Evacuation Dynamics 2005* Eds N Waldau, P Gattermann, H Knoflacher, M Schreckenberg (Springer, Berlin) pp 254 – 265

Kerridge J, Hine J, Wigan M, 2001, "Agent-based modeling of pedestrian movements: the questions that need to be asked and answered" *Environment and Planning B: Planning and Design* **28** 327 – 341

Oppewal H, Timmermans H J P, 1997, "Modelling the effects of shopping centre size and store variety on consumer choice behaviour" *Environment and Planning A* **29** 1073 – 1090

Payne J W, 1982, "Contingent decision behaviour" *Psychological Bulletin* **92** 382 – 401

Payne J W, Bettman J R, 1988, "Adaptive strategy selection in decision making" *Journal of Experimental Psychology: Learning, Memory and Cognition* **14** 534 – 552

Recker W W, Golob T F, 1979, "A non-compensatory model of transportation behaviour based on sequential consideration of attributes" *Transportation Research B* **13** 269 – 280

Saito S, Ishibashi K, 1992, "A Markov chain model with covariates to forecast consumer's shopping trip chain within a central commercial district", presented at the *Fourth Congress of Regional Science Association International, Marloca* http://resweb2.jhk.adm.Fukuoka-u.ac.jp/FukuokaUnivHtml/info/3369/R110J.html

Swait J, 2001, "A non-compensatory choice model incorporating attribute cutoffs" *Transportation Research B* **35** 903 – 928

Timmermans H J P, 1983, "Non-compensatory decision rules and consumer spatial choice behaviour: a test of predictive ability' *The Professional Geographer* **35** 449 – 455

Tversky A, 1972, "Elimination by aspects" *Psychological Review* **79** 281 – 299

Van der Waerden P, Borgers A, Timmermans H, 1998, "The impact of the parking situation in shopping centres on store choice behaviour" *GeoJournal* **45** 309 – 315

Willis A, Gjersoe N, Harvard C, Kerridge J, Kukla R, 2004, "Human movement behaviour in urban spaces: implications for the design and modeling of effective pedestrian environments" *Environment and Planning B: Planning and Design* **31** 805 – 828

Young W, 1984, "A non-tradeoff decision making model of residential location choice" *Transportation Research A* **18** 1 – 11

Zhu W, Timmermans H J P, Wang D, 2006, "Temporal variation in consumer spatial behavior in shopping streets" *Journal of Urban Planning and Development* **132** 166 – 171