Decision Process Models of Pedestrian Walking Direction Choice

Wei Zhu Harry Timmermans

Wei Zhu, corresponding author, postdoctoral research fellow, Center for Adaptive Behavior and Cognition, Max-Planck-Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail: zhu@mpib-berlin.mpg.de
Harry Timmermans, professor, Urban Planning Group, Eindhoven University of Technology, Den Dolech 2, PO Box 513, 5600 MB Eindhoven, The Netherlands. E-mail: h.j.p.timmermans@tue.nl

Introduction

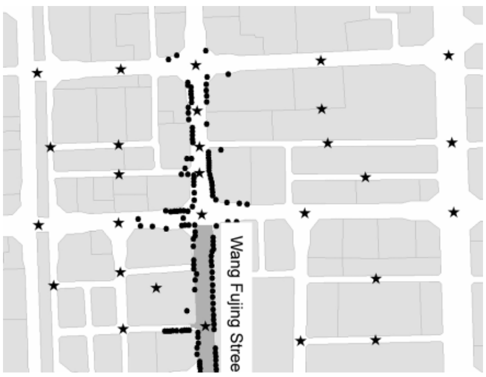
Pedestrian shopping behavior can be modeled as the result of a complex decision system which is composed of many inter-related decisions. Several research frameworks regarding this problem have been proposed. For example, Borgers and Timmermans (1985) decomposed pedestrian shopping trip into three steps: choosing a destination according to a shopping plan, choosing a route given the destination, and choosing whether to make an impulse visit to an unplanned destination. Hoogendoorn and Bovy (2004, 2005) proposed a framework including activity scheduling, activity area choice, and route choice. Dijkstra *et al*. (2005) added a perception module into their multi-agent pedestrian simulation system to model the decision of attending a store.

Although the behavioral aspects modeled are different in these studies, a common methodology used in these studies is the discrete choice models which have become the dominant technique for modeling decisions in transportation research. However, the validity of these models, and rational choice models in general, as process models of human decision making has been questioned (e.g., Hensher *et al*., 2006, 2007) due to that rational choice models assume an unreal decision maker, at least at the cognition level, who uses all relevant information, combines this information according to some weighted-additive utility function, and chooses the alternative with the highest utility. As an alternative theory of decision making, Simon's (1956, 1959) notion of bounded rationality has regained attention. It states that humans use simple decision rules, or heuristics, which aim for satisficing specific goals rather than optimizing results. Especially in psychology, researchers have compared heuristic decision models with rational decision models (e.g., Gigerenzer and Goldstein, 1996; Gigerenzer *et al*., 1999) and tried to identify for what kind of decision problems and in what contexts either type of model is more appropriate as a cognitive process explanation. Evidence showed that people used heuristics more often. In pedestrian behavior research, however, we have not seen any comparative studies on this topic.

This paper aims at initiating such a research line by showing a comparative study between heuristic models and discrete choice models on the problem of pedestrian choosing a walking direction in shopping streets, similar to the route choice problem in previous studies. It is part of a research project focusing on using bounded rationality principles to model pedestrian shopping behavior (Zhu, 2008). We will compare standard multinomial logit models (MNL) and mixed logit models with three heuristic models: conjunctive rule, disjunctive rule (e.g., Dawes, 1964) and lexicographic rule (e.g., Fishburn, 1974). The second section of the paper introduces the data for model estimation. The third section introduces the operationalization of the direction choice problem and gives the specifications of each model. The fourth section compares the models based on model estimation results. The fifth section concludes the paper.

Data

A dataset about pedestrians' shopping diaries was collected in Wang Fujing Street (WFS), the main shopping street in Beijing, China, in 2004. The street is about 1,200 m long within which about 530 m is the pedestrianized section (Figure 1). Twenty undergraduate students from the Department of Regional and Urban Planning, Peking University, administered the survey on May 17 (Monday) and 22 (Saturday). From 11:00 to 20:00, they asked randomly selected pedestrians who indicated that they had completed their shopping trip, to fill out a questionnaire, which recorded each respondent's sequential visits in stores from the beginning of the shopping trip to the end. A total of 760 valid diaries were collected.



The non-pedestrianized section

The pedestrianized section

Figure 1. Survey area of WFS

Models

We assumed that a pedestrian made direction choice decisions each time after visiting a store or resting at some place. Although respondents did not explicitly report their chosen directions, it is not difficult to infer the choice outcome from the destination of their next movement relative to their current location. The situation in WFS is relatively simple because the street is almost linear, and only two directions, north and south, relative to the current location of the pedestrian need to be identified as choice alternatives. The current location is the store or place where the pedestrian just conducted some activities.

For each direction, three factors were considered relevant for the pedestrians' decision. The first factor is whether the direction is the same as the one that the pedestrian just came from, represented by a dummy variable d^Y (*Y* = *N* (North), *S* (South)). Because there is a natural tendency that pedestrians follow the previous direction and try to minimize the number of back-turns, a positive influence is expected from this factor. The second factor is the total retail floorspace in the direction, q^Y. Although a pedestrian does not actually know the total amount of floorspace, the variable substitutes the pedestrian's estimate about the attractiveness of retail activities based on his/her perception of the environment. The third factor, l^Y, is the length of the pedestrianized street in the direction, representing the amenity of walking. Because there is a considerable number of observations in the data showing that pedestrians turned back at the end of the pedestrianized section and did not go further into the non-pedestrianized sections, it is reasonable to hypothesize that the longer the pedestrianized part the more attractive the waling direction.

Heuristic models

Conjunctive rule

The conjunctive rule is a typical heuristic decision rule which states that all criteria (or thresholds) of related factors have to be met in order to arrive at a positive overall judgment. Applying this rule, the decision process of direction choice is assumed to consist of two stages. The first stage is the screening stage in which the pedestrian judges whether a direction is satisfactory. If the rule for judging satisfaction is conjunctive, this process can be represented as,

$$\begin{aligned} p_1^Y &= \prod_x p_1^{Yx} \quad x = d, q, l \\ p_1^{Yd} &= \alpha^d (1 - d^Y) + \beta^d d^Y \\ p_1^{Yq} &= G^q (q^Y - \alpha^q, \beta^q, \theta^q) \\ p_1^{Yl} &= G^l (l^Y - \alpha^l, \beta^l, \theta^l) \end{aligned} \tag{1}$$

where p_1^Y is the probability that the direction is satisfactory, p_1^{Yx} is the probability that a factor is considered satisfactory. For the factor of previous direction, *d*, there are only two possible values. Being 1 means the direction is the same as the previous direction, and being 0 means it is not the same. Because the pedestrian may be satisfied with either situation, β^d is the parameter representing the probability of being satisfied when the direction is the same as the previous direction, and α^d represents the probability of being satisfied if the direction is opposite to the previous direction. For the other two factors, they are considered satisfactory when their value x^Y exceed the threshold δ^x. We assumed that the thresholds of the factors follow the distributions, $\delta^x \sim \alpha^x + \Gamma(\beta^x, \theta^x)$. Here Γ represents the standard gamma distribution, α is a constant, β is the shape parameter and θ is the scale parameter. The cumulative function is G^x.

If only one direction survives the screening, then the decision process stops with this direction chosen. If both directions are satisfactory or unsatisfactory, then the decision enters the second stage, which we simply assumed as a random choice with 50% probability of choosing either direction. Aggregating the two stages, the overall probability that a direction is chosen, p^Y, is

$$
\begin{aligned}
p^{Y} &= p_1^{Y} p_0^{\bar{Y}} + (p_1^{Y} p_1^{\bar{Y}} + p_0^{Y} p_0^{\bar{Y}})0.5 \\
p_0^{Y} &= 1 - p_1^{Y}
\end{aligned}
\tag{2}
$$

Disjunctive rule

The disjunctive rule says that an alternative is satisfactory as long as at least a factor meets the threshold value. Under this rule, the first expression in Equation 1 needs to be replaced by

$$
p_1^{Y} = \sum_x p_1^{Yx} - p_1^{Yd} p_1^{Yq} - p_1^{Yd} p_1^{Yl} - p_1^{Yq} p_1^{Yl} + \prod_x p_1^{Yx} \tag{3}
$$

which is a function of “or” relationship between the three factors.

Lexicographic rule

The lexicographic rule assumes that alternatives are compared on an attribute-by-attribute basis following some information search sequence which is organized according to descending factor importance. Thus, alternatives are first compared in terms of the most important attribute; if they tie the factor next in importance is evaluated and so on until a choice can be made or all factors have been evaluated. In the latter case, the two alternatives are indifferent. The comparisons between factors depend on the levels of each factor. There must be at least two levels to differentiate the alternatives. In this simplest situation, let δ_j be the threshold which divides the factor into a higher level when $x_j \geq \delta_j$ and into a lower level when $x_j < \delta_j$. It is also assumed to be a distribution like those above.

According to the lexicographic rule, once the pedestrian finds one direction better than the other direction, this direction will be chosen. The probabilities of factor comparison are equal to,

$$
\begin{aligned}
p_B^{Yx} &= p_1^{Yx} p_0^{\bar{Y}x} \\
p_W^{Yx} &= p_0^{Yx} p_1^{\bar{Y}x} \\
p_T^{Yx} &= 1 - p_B^{Yx} - p_W^{Yx}
\end{aligned}
\tag{4}
$$

Here, p_1^{Yx} represents the probability that factor *x* of direction *Y* is at the higher level, and p_0^{Yx} refers to the probability of being at the lower level; p_B^{Yx} is the probability that the factor of this direction is better than that of the other direction, while p_W^{Yx}

means it is worse. Parameter p_T^{Yx} represents the probability of a tie. In the last situation, another factor is used to compare the two directions. If they still cannot be discriminated after the last factor is compared, a random choice is assumed. The probability of a direction being chosen given this assumed decision process, assuming the sequence of factor comparison is $d \to q \to l$, can thus be expressed as,

$$
\begin{aligned}
p^Y &= p_B^{Yd} + p_T^{Yd} p' \\
p' &= p_B^{Yq} + p_T^{Yq} p'' \\
p'' &= p_B^{Yl} + p_T^{Yl} 0.5
\end{aligned}
\tag{5}
$$

Discrete choice models

Assume that the pedestrian chooses the direction that has the highest utility. Under the MNL framework, the probability of choosing a particular direction then equals,

$$
\begin{aligned}
p^Y &= \frac{\exp(v^Y)}{\sum_{Y'} \exp(v^{Y'})} \quad Y, Y' = N, S \\
v^Y &= \beta^d d^Y + \beta^q q^Y + \beta^l l^Y
\end{aligned}
\tag{6}
$$

where p^Y is the choice probability, v^Y is the observable utility and the βs are parameters for the respective variables. If taste variation is considered, we assume the parameters are normal distributions, which is a standard mixed logit model specification.

Results

The comparison includes 11 models, conjunctive and disjunctive model, 6 lexicographic models under the full permutation of factor search sequences, the standard MNL model, MNL model with the variables taken natural log transformation, and mixed logit model. Due to page limit, we cannot show the full model estimation results. Only the parameter estimation of the optimal model is illustrated (Table 1). The selection of the optimal model is based on Consistent Akaike Information Criterion. The overall fitting statistics of the rest models are show in Table 2.

Table 1. Estimation results of the optimal model

Lexicographic $q \rightarrow d \rightarrow l$	
Parameter	Estimate
α^d	0.381 *
β^d	0.767 *
α^q	17,999.620 *
β^q	-
θ^q	-
α^l	348.636 *
β^l	-
θ^l	-
N^C	2,098
N^P	4
LL	-946
CAIC	1,926

N^C : number of cases

N^P : number of free parameters

LL: log-likelihood

CAIC: consistent Akaike Information Criterion

*: parameters counted as free parameter

Table 2. Comparison of model fit

Model	N^P	LL	CAIC
Conjunctive	5	-966	1,975
Disjunctive	3	-987	2,000
Lexicographic $d \rightarrow q \rightarrow l$	5	-963	1,968
Lexicographic $d \rightarrow l \rightarrow q$	5	-962	1,968
Lexicographic $q \rightarrow d \rightarrow l$	4	-946	1,926
Lexicographic $q \rightarrow l \rightarrow d$	4	-953	1,941
Lexicographic $l \rightarrow d \rightarrow q$	4	-970	1,974
Lexicographic $l \rightarrow q \rightarrow d$	4	-953	1,941
MNL standard	2	-989	1,996
MNL with logged variables	3	-991	2,008
Mixed logit	6	-988	2,029

All the heuristic models are statistically better than the MNL models. In general, the lexicographic models perform better than the other models, suggesting that attribute-based comparison of choice alternatives is more appropriate for the direction choice decision. The influence of factor search sequence is notable. The sequences starting from *d* have an inferior CAIC compared to other sequences.

Although judging and following the previous direction is much easier, the condition of shopping environments seems to be more important. This is understandable since most pedestrians visit the street for fulfilling their needs. In particular, floorspace appears to be more important than pedestrianized street, which is reflected in that the optimal search sequence is $q \rightarrow d \rightarrow l$ and its counter sequence $l \rightarrow d \rightarrow q$ is much worse. It looks like a wise decision strategy, well balanced between accuracy and speed. Comparing floorspace first, although it could involve more mental effort, has the highest probability to guarantee that pedestrian needs will be better realized in the chosen direction. Moreover, when the directions are not differentiated under floorspace, comparing with the previous direction is quick and easy. In the optimal model, both threshold values are scalar. Directions with a total floorspace larger than 18,000 m^2 and a length of pedestrianized street longer than 350 m will be considered satisfactory for these factors respectively.

Conclusion

We proposed using heuristic models as alternatives to rational choice models for modeling decision processes of pedestrians, because heuristic models are theoretically more appropriate as cognition level explanations of decision making. This study empirically tested this conjecture by comparing the two types of models on the decision problem of choosing walking direction in shopping streets. It involved three typical heuristic models (conjunctive, disjunctive and lexicographic rule) and discrete choice models (MNL and mixed logit model). The results showed that heuristic models are better than discrete choice models in general and a lexicographic model is the best in particular, which shows the empirical validity of using heuristic models. The most probable information search sequence is from shopping attraction, to walking history, then to walking condition. This can be sensibly interpreted.

References

Borgers, A. and H. Timmermans, 1995, "A Model of Pedestrian Route Choice and Demand for Retail Facilities within Inner-City Shopping Areas", *Geographical Analysis*, 18 (2), 115-128.

Dawes, R. M., 1964, "Social Selection based on Multidimensional Criteria", *Journal of Abnormal and Social Psychology*, 68 (1), 104-109.

Dijkstra, J., H. Timmermans and Bauke de Vries, 2005, "Modeling Behavioral Aspects of Agents in Simulating Pedestrian Movement", in *Proceedings of the 9th International Conference on Computer in Urban Planning and Urban*

Management, London, paper No. 63, 11 pp.
Fishburn, P. C., 1974, "Lexicographic Orders, Utilities and Decision Rules: A Survey", *Management Science*, 20 (11), 1442-1471.
Gigerenzer, G. and G. Goldstein, 1996, "Reasoning the Fast and Frugal Way: Models of Bounded Rationality", *Psychological Review*, 103, 650-669.
Gigerenzer, G., P. M. Todd and ABC Research Group, 1999, *Simple Heuristics that Make us Smart*, Oxford University Press, Oxford University Press, New York.
Hensher, A. H., J. Rose and S. Puckett, 2006, "Selective Developments in Choice Analysis and a Reminder about the Dimensionality of Behavioural Analysis", resource paper for the *2006 International Association of Traveler Behavior Conference*, Kyoto, Japan, 44 pp.
Hensher, A. H., J. Rose and T. Bertoia, 2007, "The Implications on Willingness to Pay of a Stochastic Treatment of Attribute Processing in Stated Choice Studies", *Transportation Research E*, 43, 73-89.
Hoogendoorn, S. P. and Bovy, P. H. L., 2004, "Pedestrian Route-Choice and Activity Scheduling Theory and Models", *Transportation Research B*, 38, 169-190.
Hoogendoorn, S. P., and Bovy, P. H. L., 2005, "Pedestrian Travel Behavior Modeling", *Networks and Spatial Economics*, 5 (2), 193-216.
Simon, H. A., 1956, "Rational Choice and the Structure of the Environment", *Psychological Review*, 63 (2), 129-138.
Simon, H. A., 1959, "Theories of Decision-Making in Economics and Behavioral Science", *The American Economic Review*, 49 (3), 253-283.
Zhu, W., 2008, *Bounded Rationality and Spatio-Temporal Pedestrian Shopping Behavior*, doctoral thesis, Eindhoven University of Technology, The Netherlands.