Aloys Borgers and
Harry Timmermans

A Model of Pedestrian Route Choice and Demand for Retail Facilities within Inner-City Shopping Areas

There are still only a few operational models of pedestrian movement. In particular, the gravity/entropy-maximizing model has received most attention. In this paper a descriptive model of pedestrian movement is presented. It can be considered as an extension of O'Kelly's model of the demand for retail facilities in the presence of multistop, multipurpose trips. The model basically consists of three submodels: one for destination choice, one for route choice, and one for impulse stops. Together, these submodels describe/predict the total demand for retail facilities within inner-city shopping areas. The model is applied to data from the city of Maastricht, The Netherlands.

1. INTRODUCTION

The current redevelopment of many inner-city shopping areas, as exemplified by the construction of new in-town hypermarkets and changes in the transportation network, may change the pattern of inner-city pedestrian movement and hence the viability of particular shopping streets dramatically. It is therefore important to predict the likely effects of such redevelopment schemes on retail turnover in shopping streets when evaluating alternative planning scenarios. Previous research has shown that the commercial viability of inner-city shopping streets is highly influenced by pedestrian movement and that the impact of new retail developments or changes in the transportation network on the turnover in the shopping streets is closely related to the functional linkages between shops/shopping streets, the locational pattern of large magnet stores, and the distribution of transport termini such as bus stops and car parks (Johnston and Kissling 1971; Bennison and Davies 1977a, b; Pacione 1980). This empirical finding implies that models of pedestrian movement should incorporate the functional and spatial linkages between retail sectors and shopping streets if they are to be used successfully for predicting the effects of redevelopment schemes on the commercial viability of shopping streets.

The authors wish to thank two anonymous referees for their useful comments.

Aloys Borgers is a research fellow and Harry Timmermans is lecturer, Department of Architecture, Building, and Planning, Eindhoven University of Technology, The Netherlands.

Geographical Analysis, Vol. 18, No. 2 (April 1986) © 1986 Ohio State University Press
Submitted 6/84. Revised version accepted 9/85.

An examination of the existing literature, however, shows that the few operational models of pedestrian movement do not explicitly incorporate such linkages. Sandahl and Percivall's (1972) model is inherently nonspatial: the number of pedestrians in streets is predicted by such variables as accumulated demand and land use using regression analysis. Factors such as pedestrian flows, route choice, sequence of stops, and the strength of the functional relationships are not considered by their model. Likewise, Butler's (1978) entropy-maximizing model does not incorporate the strength of the functional linkages between retail sectors and the interdependencies between shops.

Models of trip chaining behavior, however, which have been developed in a different context, are explicitly concerned with interrelated choices underlying multipurpose trips; as such they might be a viable starting point for developing a model of pedestrian movement. Two approaches should be further explored in this respect. First, the utility-maximizing models of trip chaining behavior (Lerman 1979; Horowitz 1979, 1980; Kitamura 1984) might be refined by explicitly incorporating the functional relationships between shopping streets and retail sectors. Second, the descriptive models of multipurpose trip behavior might be extended by including explanatory variables to account for the observed functional and spatial relationships in multipurpose trips underlying consumer choice processes.

In this study the latter approach is adopted. The aim of the present paper, therefore, is to develop and test a model of multipurpose trips with particular reference to route choice and destination shopping behavior within city centers. In particular, the model may be considered as an extension of O'Kelly's (1981) model of the demand for retail facilities in the presence of multistop, multipurpose trips, which is based on Howard's (1971) ideas on time-varying Markov processes. The basic strength of his model is its ability to provide detailed insight into the relative interdependence between locations and retail sectors in shopping behavior while avoiding some of the criticisms that have been raised against the use of simple Markov models (Horton and Schuldiner 1967; Horton and Wagner 1968; Sasaki 1971; Wheeler 1972; Kondo 1974) in modeling multipurpose trips (Cullen et al. 1972; Jones 1976; Hanson 1979).

The paper is divided into four sections. The next section provides an outline of the theoretical considerations underlying the model. This is followed in section 3 by a description of the model itself. Section 4 then presents the results of an empirical test of the model in the city center of Maastricht, the Netherlands. The paper is concluded by evaluating this empirical test and suggesting some lines for further research.

2. THEORETICAL CONSIDERATIONS

Pedestrian destination and route choice within city centers is conceptualized as follows. Assume that pedestrians have selected some entry point of a particular city center. They are then faced with the problem of choosing in some sequence a number of destinations, given their decision to purchase a set of goods. Hence, the destination choice process is conceptualized here as a multipurpose trip. We hypothesize that the choice itself is the result of a decision-making process in which pedestrians evaluate each of a potential large number of choice alternatives on the basis of a set of attributes relevant to their decision-making task and select the alternative that receives the highest evaluation. The set of attributes is assumed to consist of both locational and nonlocational attributes.

Suppose now that the first destination has been selected. It represents the first stop in the shopping trip. We hypothesize that the choice of the next destination is also the result of an evaluation (utility)-maximizing process in which pedestrians trade off the locational and nonlocational attributes of potential choice alternatives and arrive at a choice. We assume that this process continues until all goods are purchased. Hence we conceptualize the destination choice process as a multistep decision-making process in which utilities are *sequentially* maximized.

The result of the destination choice process is a set of destinations where the goods are purchased. Next, we conceptualize the route choice process as one in which pedestrians choose a route that connects two successive destinations in a shopping trip from among all possible routes. Again, we assume that route choice is the result of a decision-making process in which pedestrians integrate their utilities, defined on the attribute levels of the choice alternatives, into some overall utility measure and then choose the alternative that receives the highest utility.

Finally, we assume that the shopping trip itself may give rise to impulse buying. That is, additional goods may be purchased by pedestrians along their routes. Hence, we assume that impulse stops are conditional upon pedestrian destination and route choice behavior. Together, the regular stops and the impulse stops determine the commercial viability of the shopping streets within the city center.

In sum, we assume that consumer choice behavior in inner-city shopping areas may be represented as a multipurpose trip that consists of both intended and impulse stops. The intended stops are assumed to be the result of a multistep decision-making process of destination choice in which utilities are sequentially maximized, whereas the impulse stops are a function of route choice behavior, which, in turn, is conditional upon pedestrian destination choice. It follows that the modeling approach should include separate submodels of destination choice, route choice, and impulse stops.

3. THE MODEL

Suppose the streets of a city center are defined in terms of N links of a network. The nodes of this network consist of the intersections of the shopping streets. The set of links consists of two subsets. Let $i, j = 1, 2, \ldots, N'$ index the links that are associated with the city center entry and departure points. The indices $i, j = N' + 1, N' + 2, \ldots, N$ denote the links or shopping streets that are visited for shopping at successive stops during a trip. Each of these links has a number of associated shops. The purpose associated with each stop is represented by $g, h = 1, 2, \ldots, Z$. The type of shop that is visited at successive stops is indexed by $g, h = 2, 3, \ldots, Z$, while $g, h = 1$ indicates entry or departure. The stops are indicated by $m = 1, 2, \ldots, M$ and the functions $s(m)$ and $d(m)$ respectively denote the purpose and link associated with the mth stop. Hence, by definition it follows that

$$s(m) = 1 \text{ implies } d(m) \in \{1, 2, \ldots, N'\} \tag{1}$$

$$s(m) \neq 1 \text{ implies } d(m) \in \{N' + 1, N' + 2, \ldots, N\} \tag{2}$$

$$s(1) = 1 \tag{3}$$

$$s(m) \neq 1 \text{ if } 1 < m < M \tag{4}$$

$$s(M) = 1. \tag{5}$$

Assume now that the initial distribution of pedestrians over the city center entry points is given by

$$\mathbf{B}(1) = \{b_i^{(g)}(1)\}, i = 1,2,\ldots,N; g = 1,2,\ldots,Z, \tag{6}$$

where $\mathbf{B}(1)$ is a $(1 \times NZ)$ vector of the first stops at link i for purpose g. Due to the definitions it follows that

$$b_i^{(g)}(1) = Pr\{d(1) = i, s(1) = g\}, g = 1; i = 1,2,\ldots,N' \tag{7}$$

$$b_i^{(g)}(1) = 0, g = 1; i = N' + 1, N' + 2,\ldots,N \tag{8}$$

$$b_i^{(g)}(1) = 0, g = 2,3,\ldots,Z; i = 1,2,\ldots,N. \tag{9}$$

The problem may then be stated as follows: Given (1) the state space $\mathbf{D} = \{D_{ig}\}$, $i = 1,2,\ldots,N$; $g = 1,2,\ldots,Z$ of mutual exclusive links associated with purpose g, and (2) the initial distribution $\mathbf{B}(1)$ of pedestrians, what is the probability that each of the N shopping streets (links) will be chosen for each of the Z purposes?

Consider first the general case. The conditional probability that a trip that stopped for purpose g at stop $(m-1)$ stops for purpose h at the mth stop is given as

$$a^{gh}(m|m-1) = Pr\{s(m) = h|s(m-1) = g\}. \tag{10}$$

These conditional probabilities can be included in a $(Z \times Z)$ matrix $\mathbf{A}$ that has the following structure

$$\mathbf{A}(m|m-1) = \left[\begin{array}{c|c} 0 & \mathbf{A}^{**}(m|m-1) \\ \hline \mathbf{Y}(m|m-1) & \mathbf{A}^{*}(m|m-1) \end{array}\right], \tag{11}$$

where

$$\mathbf{A}^{**}(m|m-1) = \begin{cases} a^{gh}(m|m-1), g = 1; h = 2,3,\ldots,Z; m = 2 \\ 0 \text{ otherwise} \end{cases} \tag{12}$$

$$\mathbf{A}^{*}(m|m-1) = \begin{cases} a^{gh}(m|m-1), g, h = 2,3,\ldots,Z; m = 3,4,\ldots,M-1 \\ 0 \text{ otherwise} \end{cases} \tag{13}$$

$$\mathbf{Y}(m|m-1) = \begin{cases} a^{gh}(m|m-1), g = 2,3,\ldots,Z; h = 1; m = M \\ 0 \text{ otherwise} \end{cases} \tag{14}$$

$\mathbf{A}^*$ describes the probability that a trip that stopped at the $(m-1)$th stop for purpose g stops at the mth stop for purpose h. Similarly, $\mathbf{A}^{**}$ describes the probability that the various types of shops are visited, given that the pedestrian just entered the city center. $\mathbf{Y}$ describes the probability that the trip is terminated at the mth stop, given that shop type g was visited at the $(m-1)$th stop. Notice that these formulations are analogous to Howard (1971) and O'Kelly (1981): the transition probabilities are not necessarily constant over a trip, but may vary as a function of the stops.

This formulation may be extended to the special case of a spatial model (O'Kelly 1981). First, define

$$p_{ij}^{gh}(m|m-1) = Pr\{d(m) = j,\, s(m) = h | d(m-1) = i,\, s(m-1) = g\} \tag{15}$$

as the conditional probability that the jth link will be chosen on the mth stop for purpose h, given that the ith link was chosen for purpose g on stop $(m-1)$. These conditional probabilities may be collected in matrix $\overline{\mathrm{IP}}(m|m-1)$:

$$\overline{\mathrm{IP}}(m|m-1) = \mathbf{P}^{gh}(m|m-1);\, g, h = 1,2,\ldots, Z;\, m = 2,3,\ldots, M, \tag{16}$$

where

$$\mathbf{P}^{gh}(m|m-1) = p_{ij}^{gh}(m|m-1), i, j = 1,2,\ldots, N \tag{17}$$

$$p_{ij}^{gh}(m|m-1) = 0 \text{ if } \begin{cases} m = 2 \text{ and } g = 2,3,\ldots, Z \\ m = M \text{ and } h = 2,3,\ldots, Z \\ g = 1 \text{ and } h = 1 \\ g = 1 \text{ and } i = N' + 1, N' + 2,\ldots, N \\ h = 1 \text{ and } j = N' + 1, N' + 2,\ldots, N \end{cases} \tag{18}$$

The probability that a pedestrian will choose link-purpose combination (j, h), given that he has chosen combination (i, g) on a previous stop l, then equals:

$$q_{ij}^{gh}(m, l) = \sum_{a=1}^{Z} \sum_{v=1}^{N} \sum_{b=1}^{Z} \sum_{w=1}^{N} \cdots \sum_{d=1}^{Z} \sum_{y=1}^{N} p_{iv}^{ga}(l+1|l)\, p_{vw}^{ab}(l+2|l+1) \cdots p_{yj}^{dh}(m|m-1). \tag{19}$$

Or, in matrix notation

$$\mathbf{Q}(m, l) = \prod_{k=l+1}^{m} \overline{\mathrm{IP}}(k|k-1), 1 \leqslant l < m;\, m = 2,3,\ldots, M. \tag{20}$$

The total number of interactions between link-purpose combination (i, g), which was visited on the lth stop, and link-purpose combination (j, h), visited on or before the Mth stop, then equals

$$n_{ij}^{gh}(M, l) = \sum_{k=l+1}^{M} q_{ij}^{gh}(k, l). \tag{21}$$

The demand for retail facilities at link j in retail sector h is then easily computed as

$$b_j^{(h)}(M) = \sum_{g=1}^{Z} \sum_{i=1}^{N} b_i^{(g)}(1) \sum_{k=2}^{M} q_{ij}^{gh}(k, 1). \tag{22}$$

The above equation is one of the basic results of the model and is similar to O'Kelly's equation for modeling multipurpose trips involving two or more shopping centers. If the transition probabilities are estimated by maximum likelihood procedures on the basis of observed pedestrian flows, the model will perfectly describe observed pedestrian destination choice behavior, implying that we are dealing with a fundamentally descriptive model.

The development of a predictive model of destination choice would require a submodel that predicts the observed transition probabilities. Several such models may be formulated, but in the present study a simple model was assumed.

Destination Choice

In particular, it is assumed that the transition probabilities are a multiplicative nonlinear function of the total amount of floor space in retail sector g at link j and the distance separation between the links. A suitable formulation would then be

$$p_{ij}^{(g)}(m|m-1) = \frac{F_j^{(g)\alpha_{gm}} \exp\left(c_{ij}^{-\beta_{gm}}\right)}{\sum_{j'=N'+1}^{N} F_{j'}^{(g)\alpha_{gm}} \exp\left(c_{ij'}^{-\beta_{gm}}\right)}, \tag{23}$$

$$g = 2,3,\ldots,Z;\ m = 2,3,\ldots,M-1;$$

$$j = N'+1, N+2,\ldots,N;$$

$$i = 1,2,\ldots,N' \text{ if } m = 2;$$

$$i = N'+1, N'+2,\ldots,N \text{ if } m = 3,4,\ldots,M-1,$$

where

$p_{ij}^{(g)}(m|m-1)$ is the probability that link j will be chosen for purpose g at stop m, given that the previous stop was at link i
$F_j^{(g)}$ is the total amount of floor space for shop type g at link j
c_{ij} is the distance separation between link i and link j
α_{gm} and β_{gm} are the purpose and stop specific parameters to be estimated.

Route Choice

The model described so far predicts pedestrian destination choice behavior. It results in the probability that each of the destinations will be selected at a particular stop. The next step involves predicting pedestrian route choice behavior, given the choice of the destinations. Assume that pedestrians attach a utility value to each alternative route r. Then

$$u_r = f(\mathbf{x}_r), \tag{24}$$

where u_r is the utility of route r, and $f(\mathbf{x}_r)$ is a function of a vector of attributes of alternative r.

To account for measurement error and heterogeneity, it is assumed that the utility function consists of two parts, viz. a deterministic part v_r and an error term. Assume that these two parts are independent and additive. It follows that

$$u_r = v_r + \epsilon_r. \tag{25}$$

If it is assumed that pedestrian route choice behavior is the result of a utility-maximizing process, and the error terms are identically and independently Weibull distributed, the choice probabilities are obtained by the well-known multinomial logit model. It can be expressed as

$$p(r|R_{ij}) = \frac{\exp(\gamma \mathbf{x}_r)}{\sum\limits_{r' \in R_{ij}} \exp(\gamma \mathbf{x}_{r'})}, \tag{26}$$

where $\mathbf{x}_r$ is a vector of attributes of the rth alternative, γ is a vector of parameters to be estimated, and R_{ij} is the set of possible routes between link i and link j. The probability that a pedestrian then passes through the kth link given that he has chosen route r equals

$$p(k|r) = \begin{cases} 1, \text{if } k \in r \\ 0 \text{ otherwise}. \end{cases} \tag{27}$$

Thus, the total number of pedestrians in link k, C_k, equals

$$C_k = \sum_i \sum_j \sum_{r \in R_{ij}} p(k|r) p(r|R_{ij}) E_{ij}, \tag{28}$$

where E_{ij} is the total number of individuals going from link i to link j.

Impulse Stops

To extend the model further, assume that pedestrian shopping behavior consists of two components: a fixed component that represents the intended stops at specific links, and a component that relates the impulse buying. Assume that the demand derived from impulse stops is a function of link attributes and the number of pedestrians passing through the link. Then

$$b_i^{*(g)} = f'(\mathbf{x}_i^{(g)}, C_i), g = 2,3,\ldots, Z; i = N' + 1, N' + 2,\ldots, N, \tag{29}$$

where $b_i^{*(g)}$ is the demand related to impulse stops at link i for shop type g, and $\mathbf{x}_i^{(g)}$ is a vector of attribute values of link i for shop type g. Thus, the total demand for retail facilities within the system can be computed by

$$\tilde{b}_i^{(g)} = b_i^{(g)}(M) + b_i^{*(g)}. \tag{30}$$

For planning purposes, the following indices may be calculated:

$$T_i^{(g)} = \tilde{b}_i^{(g)} e_i^{(g)}, g = 2,3,\ldots, Z; i = N' + 1, N' + 2,\ldots, N \tag{31}$$

$$V_i^{(g)} = T_i^{(g)}/F_i^{(g)}, g = 2,3,\ldots, Z; i = N' + 1, N' + 2,\ldots, N, \tag{32}$$

where $T_i^{(g)}$ is the retail turnover in the ith link for shop type g, $e_i^{(g)}$ is the average per capita expenditure in the ith link for shop type g, and $V_i^{(g)}$ is the turnover to floorspace ratio in the ith link for shop type g.

4. TEST OF THE MODEL

The model outlined in the previous section was tested using data pertaining to the city center of Maastricht. The constructed network consisted of 88 links and 6 entry/departure points. Most of the data used to calibrate the submodels were collected by on-street interviews. These interviews were in the form of a structured questionnaire. Only pedestrians who were leaving the city center were asked to complete the questionnaire. Each respondent was asked to mark on a map the route taken within the city center, the links of the network where they stopped, and the type of shop associated with each stop. The shops were classified into five categories: groceries, clothing, department stores, markets, and other. Hence, for each respondent the successive link-purpose combinations can be derived.

The sample consisted of 426 respondents. They were also asked to indicate their impulse stops. The submodels for the destination choice and route choice were calibrated on the basis of non-impulse stops only. Consequently, the data of 345 respondents were used to calibrate the destination choice submodel and route choice submodel. The respondents were equally distributed over the entry points. The process of calibrating and testing the model itself consisted of three phases: (1) the submodel related to pedestrian destination choice was calibrated and tested for each purpose and number of stop separately as well as for the total pedestrian flows; (2) the submodel of pedestrian route choice behavior was calibrated and tested, given the predicted destination choices; and (3) the submodel of impulse stops was calibrated and empirically tested. The results of these three phases are presented in the following subsections.

Destination Choice

The first step in the process involved the calibration of the submodel that predicts the probability that a shopping street will be chosen to buy some goods (eq. 23). This submodel assumes that the probability of choosing a particular shopping street is proportional to a function of the floor space in that street and inversely proportional to a function of the distance separation between stops and the attractiveness of competing streets. The submodel was disaggregated by the five identified categories of shops and number of stops. Due to the small number of observations, the number of stops was limited to three. The data pertaining to the three stops and more were combined. Hence, the submodel of destination choice was calibrated separately for 15 cases.

The calibrated parameters of the submodel are provided in Table 1, which also includes some goodness-of-fit measures. Table 1 clearly demonstrates that the model fits the observed data well. The Pearson product-moment correlation coefficient between observed and predicted pedestrian interaction flows for the categories groceries, clothing, department stores, and markets ranges from 0.518 to 0.999. The model describes the observed pedestrian flows less well for the category other. This is likely due to the fact that a wide range of different types of shops with different locational patterns is included in this category. For planning purposes, however, it is more important to assess how well the model predicts observed destination totals. Table 1 shows that the destination submodel predicts the data well as indicated by the product moment correlation coefficient of > 0.82 for groceries, > 0.66 for clothing, and 0.99 for department stores and markets. The latter results should, to a large degree, be explained by the highly concentrated locational pattern of department stores—and especially markets—within the city center of Maastricht. The results for the category other compare less favorably with the results for the other categories: the correlation between observed and predicted destination totals ranges from 0.558 to 0.753. Again, this result might be explained by the heterogeneity of this category.

TABLE 1
Parameters of the Destination Choice Submodel

	Parameters		Correlation Coefficient	
Purpose	α	β	Interactions	Destination Totals
Stop 1				
Groceries	1.122	0.021	0.630	0.852
Clothing	1.030	0.036	0.737	0.912
Department stores	1.759	0.005	0.999	0.999
Markets	1.648	0.162	0.998	0.999
Other	0.604	0.013	0.341	0.558
Stop 2				
Groceries	1.201	0.038	0.737	0.891
Clothing	1.203	0.026	0.721	0.865
Department stores	0.847	0.069	0.995	0.999
Markets	1.670	0.002	0.942	0.996
Other	1.152	0.119	0.511	0.753
Stop 3, 4, . . .				
Groceries	0.724	0.157	0.736	0.827
Clothing	0.465	0.134	0.518	0.662
Department stores	1.923	0.079	0.999	0.999
Markets	1.256	1.770	0.999	0.999
Other	1.145	0.213	0.532	0.614

The estimated parameter values can be used to predict the transition probability matrices of the Markov chain and then equations (19)–(22) can be used to predict the demand for retail facilities within the network of the city center. Since the destination submodel involves a series of stops, errors may accumulate and hence it is important to assess how well the destination submodel predicts observed pedestrian behavior, not disaggregated by number of stops. The predictive validity of the submodel was assessed on the basis of the following quantities: the goodness-of-fit of the predicted pedestrian flows between the links of the network for each type of shop (purpose), respectively for the total pedestrian flows; the goodness-of-fit of the predicted arrivals and departures from the links of the network for each type of shop and the total pedestrian flows; the goodness-of-fit of the predicted demand in each link of the network for each type of shop and total demand respectively, and the goodness-of-fit for each type of shop. To avoid dependence on a particular goodness-of-fit measure, several of these measures were calculated.

The results of this testing procedure are given in Tables 2 and 3. Table 2 clearly illustrates that the predictive validity of the destination choice model is quite good. The pedestrian flows between the links of the network, as well as the arrivals and departures for the various types of shops, are well predicted. The Pearson product-moment correlation coefficients vary between 0.478 and 0.992 for the pedestrian flows, and between 0.781 and 0.999 for the arrivals; while the minimum value for this goodness-of-fit measure is 0.986 for the departures. In addition, the values of Robinson's agreement measure suggest that the relationship between the observed and the predicted flows, arrivals, and departures is fairly linear.

The results of the testing procedure for the total pedestrian flows are given in Table 3. It clearly shows that the goodness-of-fit of the model is good. Hence, the relatively unsatisfactory results for the category other are now corrected for. Table 3, however, also shows that the mean percentage error in the predicted arrivals is still 35.9 percent. Since the other goodness-of-fit measures are good, this finding is likely due to the small sample size as compared to the total number of links in the network. The average number of stops/number of links ratio is approximately 1.5 if one leaves the main shopping streets out. Hence, even a small absolute error in prediction may, under these circumstances, result in high percentage errors.

TABLE 2
Goodness-of-Fit of Destination Choice Submodel

Measure	Flows	Arrivals	Departures
Groceries			
Correlation coefficient	0.727	0.909	0.986
Robinson's agreement measure	0.836	0.947	0.992
Root mean square	0.138	1.904	0.623
Mean absolute error	0.017	0.810	0.297
Mean percentage error	0.593	20.898	7.795
Clothing			
Correlation coefficient	0.752	0.939	0.992
Robinson's agreement measure	0.854	0.967	0.996
Root mean square	0.123	1.380	0.467
Mean absolute error	0.016	0.610	0.232
Mean percentage error	0.636	13.034	7.452
Department Stores			
Correlation coefficient	0.991	0.999	0.992
Robinson's agreement measure	0.996	0.999	0.996
Root mean square	0.067	0.114	0.604
Mean absolute error	0.003	0.017	0.262
Mean percentage error	0.086	0.273	5.365
Markets			
Correlation coefficient	0.992	0.999	0.995
Robinson's agreement measure	0.995	0.999	0.997
Root mean square	0.050	0.403	0.369
Mean absolute error	0.002	0.063	0.159
Mean percentage error	0.081	1.172	5.525
Other			
Correlation coefficient	0.478	0.781	0.986
Robinson's agreement measure	0.688	0.845	0.993
Root mean square	0.115	1.467	0.455
Mean absolute error	0.017	0.809	0.246
Mean percentage error	0.657	15.035	7.987

TABLE 3
Goodness-of-Fit of the Destination Choice Submodel on the Basis of Total Pedestrian Flows

Measure	Flows	Arrivals =	Departures
Correlation coefficient	0.942	0.993	0.993
Robinson's agreement measure	0.971	0.997	0.997
Root mean square	0.309	2.620	2.620
Mean absolute error	0.068	1.469	1.469
Mean percentage error	2.111	35.873	35.873

TABLE 4
Goodness-of-Fit of the Destination Choice Submodel on the Basis of the Demand Matrix

Measure	Demand per Link and Type of Shop	Total Demand per Type of Shop	Total Demand per Link
Correlation coefficient	0.987	1.000	0.991
Robinson's agreement measure	0.993	1.000	0.995
Root mean square	1.146	0.002*	1.852
Mean absolute error	0.385	0.006*	0.734
Mean percentage error	8.402	0.006*	17.936

*Rounding errors

The final step in testing the destination submodel involved calculating the degree of correspondence between the observed and predicted (5 × 88) type of shop × link demand matrix. The results of the analysis are given in Table 4, which again illustrates that the destination submodel fits the observed data well. The correlation measures fluctuate around 0.99, the absolute and percentage errors are small, and the minimum discrimination information statistic is not statistically significant beyond the 5 percent level. The predictive validity on the basis of the total demand per link is somewhat less than the predictive validity of the destination choice submodel on the basis of the demand per link and type of shops, which might suggest that the errors accumulate in certain links of the network. Evidently, one would expect that the destination choice submodel predicts the total demand for the types of goods perfectly. Table 4 shows that this is true within rounding errors.

Route Choice

The second step in the process of assessing model performance involved the specification and calibration of the submodel that predicts route choice probabilities, given the predicted destinations associated with the different stops. It is assumed that route choice behavior is the result of a utility-maximizing process whereby individuals integrate their utilities defined on the attribute levels of the shopping streets into an overall utility. Although the specification of the utility function could involve several variables, in the present study a simple specification was first tested by assuming that pedestrians attempt to minimize distance. Hence, the route choice submodel used in the present study is a multinomial logit model with only a single independent variable, namely distance.

Maximum likelihood procedures were used to estimate the submodel's parameter. The data for this analysis consisted of the observed routes taken by the sample respondents. Five hundred two observations were used to estimate the parameter. The best fitting equation can be expressed as

$$p(r|R_{ij}) = \frac{\exp(-0.04X_r)}{\sum\limits_{r' \in R_{ij}} \exp(-0.04X_{r'})}, \tag{33}$$

where X_r represents the length of route r. The percentage of observations predicted correctly by the model was 52.4. The route choice submodel was also assessed by calculating some additional aggregate goodness-of-fit measures. The results, given in Table 5, clearly demonstrate that, in aggregate, the route choice submodel gives a good description of observed pedestrian route choice behavior. The Pearson product-moment correlation coefficient between observed and predicted route probabilities is 0.839. The value of Robinson's agreement measure suggests that the predictions are reasonable linearly related to the observations.

Once the route choice submodel is estimated, it can be linked to the destination choice submodel to see whether these two models, when combined, still provide a good description of pedestrians' actual route choice behavior. This step involves identifying the pedestrians' choice sets, estimating the utility associated with each route, and then predicting the probability that each route will be chosen, given the successive destinations associated with the stops. The identification of the routes within the choice sets requires special attention, since the enumeration of all possible routes would require too much computing time. Hence, it was decided to limit the choice sets by implementing the following rules:

- The length of a route will not be more than 2.5 the length of the shortest route (91.2 percent of the observed routes in the survey satisfy this condition).
- A route has a maximum of 13 links (this rule was satisfied by 94.7 percent of the observed routes). If a destination cannot be reached within 13 links, this rule was relaxed by calculating the minimum number of required links.
- If a pedestrian has two successive stops in the same street, he is supposed not to have left the street between these successive stops (this rule applied to 94.0 percent of the cases in the survey).

- A route consists of mutually exclusive nodes (this rule was satisfied by 86.8 percent of the observations).
- If the total number of routes included in a choice set is greater than 50, the 50 shortest routes were identified, and these 50 routes are assumed to constitute the choice set.

Again, the performance of these two combined submodels was assessed by calculating goodness-of-fit measures that quantify the degree of correspondence between the observed and predicted number of pedestrians in all links of the network.

The results are given in Table 6, which shows that the route choice submodel, when combined with the destination choice submodel, provides a good fit of the data.

Impulse Stops

As noted previously, the present model is based on a distinction between impulse and non-impulse stops. In turn, these impulse stops are assumed to be some function of pedestrian route choice and destination choice. Given the predicted destinations and routes, the final step in the approach, therefore, involves specifying and testing a submodel that predicts the distribution of the impulse stops within the city center of Maastricht. The following submodel was tested

$$p_i^{\star(g)} = \frac{F_i^{(g)\delta_g} C_i^{\theta_g}}{\sum_j F_j^{(g)\delta_g} C_j^{\theta_g}}, \tag{34}$$

where $p_i^{\star(g)}$ is the probability that an impulse stop for purpose g will be made at link i, and δ_g, θ_g are purpose-specific parameters to be estimated. Hence, it is assumed that the probability of an impulse stop is some function of the opportunities and the number of pedestrians in the various shopping streets. The demand related to impulse stops at link i for shop type g can be calculated as

$$b_i^{\star g} = p_i^{\star(g)} I^{(g)}, \tag{35}$$

where $I^{(g)}$ is the total number of impulse stops made for shop type g.

This submodel was calibrated by a gradient search technique that minimizes the sum of squared deviations. Table 7 gives the parameter estimates and an indication of the model's goodness-of-fit. Table 7 suggests that the clothing and markets retail sectors are relatively sensitive to the number of pedestrians and the opportunities in the various shopping streets. It also demonstrates that the predictive validity of this

TABLE 5
Some Goodness-of-Fit Measures for the Route Choice Submodel

Measure	Value
Correlation coefficient	0.839
Robinson's agreement measure	0.888
Root mean square	1.614

TABLE 6
Predictive Validity of the Route Choice Submodel

Measure	Value
Correlation coefficient	0.904
Robinson's agreement measure	0.947
Root mean square	28.710

beyond the 5 percent level. The predictive validity on the basis of the total demand per link is somewhat less than the predictive validity of the destination choice submodel on the basis of the demand per link and type of shops, which might suggest that the errors accumulate in certain links of the network. Evidently, one would expect that the destination choice submodel predicts the total demand for the types of goods perfectly. Table 4 shows that this is true within rounding errors.

Route Choice

The second step in the process of assessing model performance involved the specification and calibration of the submodel that predicts route choice probabilities, given the predicted destinations associated with the different stops. It is assumed that route choice behavior is the result of a utility-maximizing process whereby individuals integrate their utilities defined on the attribute levels of the shopping streets into an overall utility. Although the specification of the utility function could involve several variables, in the present study a simple specification was first tested by assuming that pedestrians attempt to minimize distance. Hence, the route choice submodel used in the present study is a multinomial logit model with only a single independent variable, namely distance.

Maximum likelihood procedures were used to estimate the submodel's parameter. The data for this analysis consisted of the observed routes taken by the sample respondents. Five hundred two observations were used to estimate the parameter. The best fitting equation can be expressed as

$$p(r|R_{ij}) = \frac{\exp(-0.04X_r)}{\sum_{r' \in R_{ij}} \exp(-0.04X_{r'})}, \tag{33}$$

where X_r represents the length of route r. The percentage of observations predicted correctly by the model was 52.4. The route choice submodel was also assessed by calculating some additional aggregate goodness-of-fit measures. The results, given in Table 5, clearly demonstrate that, in aggregate, the route choice submodel gives a good description of observed pedestrian route choice behavior. The Pearson product-moment correlation coefficient between observed and predicted route probabilities is 0.839. The value of Robinson's agreement measure suggests that the predictions are reasonable linearly related to the observations.

Once the route choice submodel is estimated, it can be linked to the destination choice submodel to see whether these two models, when combined, still provide a good description of pedestrians' actual route choice behavior. This step involves identifying the pedestrians' choice sets, estimating the utility associated with each route, and then predicting the probability that each route will be chosen, given the successive destinations associated with the stops. The identification of the routes within the choice sets requires special attention, since the enumeration of all possible routes would require too much computing time. Hence, it was decided to limit the choice sets by implementing the following rules:

- The length of a route will not be more than 2.5 the length of the shortest route (91.2 percent of the observed routes in the survey satisfy this condition).
- A route has a maximum of 13 links (this rule was satisfied by 94.7 percent of the observed routes). If a destination cannot be reached within 13 links, this rule was relaxed by calculating the minimum number of required links.
- If a pedestrian has two successive stops in the same street, he is supposed not to have left the street between these successive stops (this rule applied to 94.0 percent of the cases in the survey).

TABLE 7
Parameters and Goodness-of-Fit of the Submodel for Impulse Stops

Type	α_g	β_g	Correlation
Groceries	0.552	0.524	0.837
Clothing	0.679	0.858	0.975
Department stores	0.481	0.646	1.000
Markets	0.921	0.984	0.995
Other	1.204	0.997	0.797

TABLE 8
Predictive Validity of the Submodel for Impulse Stops

Measure	Impulse Stops per Type and Link	Total Number of Impulse Stops per Link
Correlation coefficient	0.972	0.968
Robinson's agreement measure	0.986	0.984
Root mean square	0.576	1.701
Mean absolute error	0.226	0.924
Mean percentage error	5.032	18.501

submodel is good. Pearson's product-moment correlation coefficient exceeds 0.97 for three retail sectors and still is equal to 0.84 and 0.80 for, respectively, groceries and other.

Given these estimation results, this submodel was linked to the previous two submodels and the distribution of the impulse stops was predicted, given the predicted destination and route choice behavior of the pedestrians. Again, several goodness-of-fit measures were calculated to assess the validity of the submodel when it is combined with the other two submodels. The results are given in Table 8, which clearly indicates that the combined submodels are capable of reproducing the observed impulse stops per type and link as well as the total number of impulse stops per link in a satisfactory and convincing way. The correlation coefficients are all greater than 0.96; the percentage error is only 5 and 18 percent for impulse stops per type and link and total number of impulse stops per link, respectively. The minimum discrimination information statistic is not statistically significant beyond the 5 percent probability level.

5. CONCLUSION

The main thrust of the present paper has been to outline and test a model that predicts pedestrian route choice and demand for retail facilities within inner-city shopping areas. The model is based on time-varying Markov chains, an approach that has already been used elsewhere by geographers, although in a different context. The present study, however, has extended this approach by formulating a separate submodel that predicts the transition probability matrices, thereby putting the approach in a forecasting format. In addition, a route choice submodel and a submodel for the prediction of impulse stops has been developed. Together, these submodels constitute, in theory at least, a flexible approach to the modeling of pedestrian route choice and destination choice behavior and could be used to assess the likely effects of changes in the locational patterns of retail facilities and the initial distribution of pedestrians in a city center on the profitability of different retail sectors in the various shopping streets.

The empirical analysis, conducted in the city center of Maastricht, generally tends to support the model. Notwithstanding these results, it should be noted that the present model is but one example of an approach towards modeling pedestrian movement which incorporates the functional and spatial linkages between shops/shopping streets. In addition, it is only a simple operationalization of the more general theoretical considerations underlying this study. The emphasis of the study has been primarily to improve existing models of pedestrian choice behavior by developing a model that incorporates the functional and spatial linkages within inner-city shopping centers influencing such behavior. With such a simple model providing a useful description of observed pedestrian flows, further theoretical and empirical work could be conducted. In work underway, several of such possible improvements are being tested. First, effort is being directed towards refining the destination choice submodel by introducing more and subjectively measured independent variables, and testing alternative specifications, partly derived from recent developments in discrete choice theory. Second, the effect of introducing heterogeneity in the Markov chain will be tested. Third, the route choice submodel is based on the rigorous assumption that the routes are independent. This assumption may be relaxed by estimating a full variance-covariance matrix for the routes, using a probit model. Finally, the operational decision to assume distance-minimizing behavior may be substituted by the more general assumption of utility-maximizing involving more independent variables.

LITERATURE CITED

Bennison, D. J., and R. L. Davies (1977a). "The Movement of Shoppers within the Central Area of Newcastle upon Tyne." Seminar Paper no. 34, Department of Geography, University of Newcastle upon Tyne, Newcastle.

______(1977b). "The Local Effects of City Centre Shopping Schemes: A Case Study." PTRC Summer Annual Meeting, Proceedings of Seminar A, P148 Retailing, University of Warwick, 8-19.

Butler, S. (1978). "Modelling Pedestrian Movements in Central Liverpool." Working Paper 98, Institute for Transport Studies, University of Leeds.

Cullen, I., V. Godson, and S. Major (1972). "The Structure of Activity Patterns." In *Patterns and Processes in Urban and Regional Systems*, edited by A. G. Wilson, pp. 281–96 London: Pion.

Hanson, S. (1979). "Urban-Travel Linkages: A Review." In *Behavioral Travel Modelling*, edited by D. A. Hensher and P. Stopher, pp. 81–100. London: Croom Helm.

Horowitz, J. (1979). "Disaggregate Demand Model for Nonwork Travel." *Transportation Research Record*, 673, 65–71.

______(1980). "A Utility-Maximizing Model of the Demand for Multi-Destination Non-Work Travel." *Transportation Research B*, 14, 369–86.

Horton, F. E., and P. W. Schuldiner (1967). "The Analysis of Land Use Linkages." *Highway Research Record*, 165, 96–107.

Horton, F. E., and W. E. Wagner (1968). "A Markovian Analysis of Urban Travel Behavior: Pattern Response by Socio-Economic-Occupational Groups." *Highway Research Record*, 283, 19–29.

Howard, R. A. (1971). *Dynamic Probabilistic Systems Vol. I.* New York: John Wiley.

Johnston, R. J., and C. C. Kissling (1971). "Establishment Use Patterns within Central Places." *Australian Geographical Studies*, 9, 116–32.

Jones, P. M. (1976). "The Analysis and Modelling of Multi-trip and Multi-Purpose Journeys." Working Paper no. 6, Transport Studies Unit, University of Oxford.

Kitamura, R. (1984). "Incorporating Trip Chaining into Analysis of Destination Choice." *Transportation Research B*, 18, 67–81.

Kondo, K. (1974). "Estimation of Person Trip Patterns and Model Split." In *Transportation and Traffic Theory*, edited by D. J. Buckley, pp. 715–42. Wales: N. S. Wales University.

Lerman, S. R. (1979). "The Use of Disaggregate Choice Models in Semi-Markov Process Models of Trip Chaining Behavior." *Transportation Science*, 13, 273–91.

O'Kelly, M. E. (1981). "A Model of the Demand for Retail Facilities, Incorporating Multistop, Multipurpose Trips." *Geographical Analysis*, 13, 134–48.

Pacione, M. (1980). "Redevelopment of a Median-Sized Central Shopping Area: A Case Study of Clydeband," *Tijdschrift voor Economische en Sociale Geografie*, 71, 159–68.

Sandahl, J., and H. Percivall (1972). "A Pedestrian Traffic Model for Town Centers." *Traffic Quaterly*, 26, 359–72.

Sasaki, T. (1971). "Estimation of Person Trip Patterns through Markov Chains." In *Proceedings of the Fifth International Symposium on the Theory of Traffic Flow and Transportation*, edited by G. Newell, pp. 119–30. New York: Elsevier.

Wheeler, J. O. (1972). "Trip Purpose and Urban Activity Linkages." *Annals of the Association of American Geographers*, 62, 641–54.