Modeling pedestrians' shopping behavior in downtown areas

Aloys Borgers and Harry Timmermans

Abstract

Modeling how pedestrians move through large downtown shopping areas and which outlets they visit enables planners and designers to assess the likely effects of their proposals to change or upgrade downtown shopping environments and surrounding infrastructure. Such a model was developed and tested for two downtown shopping areas in the Netherlands. The model assumes that pedestrians visiting the downtown area are attracted by different types of retail and services outlets. The attraction of outlets depends on the type and size of the outlets, the distance to the outlets, whether the outlets can be seen, have been passed or visited before, are located in indoor or outdoor areas, and proximity of similar types of outlets. Furthermore, pedestrians' route choice depends on characteristics of the network of shopping streets and the history of selected segments while moving in the shopping area. The model may be estimated from observed or traced trajectories.

A. Borgers (Corresponding author) • H. Timmermans
Urban Planning Group, Eindhoven University of Technology, The Netherlands
Email: a.w.j.borgers@tue.nl

H. Timmermans
Email: h.j.p.timmermans@tue.nl

1 Introduction

The size of pedestrian flows in pedestrianized shopping areas is an important indicator of turnover figures and real estate values in shopping areas. In order to predict the likely effects that policy measures such as changing the urban traffic infrastructure or upgrading the retail environment may have on these pedestrian flows, models of pedestrian behavior in shopping areas have been developed.

One of the first models predicting pedestrians' trajectories of movement through shopping areas was proposed by Crask (1979). He used a gravity type model to predict the successive stores to be visited in a mall. The attraction of each store was assumed dependent on the products needed by the shopper, the straight line distance to the store and a measure of congruity between the image of the typical shopper of the store and the shopper under consideration. The model takes impulse stops into account as well by means of store specific indices. Pedestrians were assumed to move to the next destination along the shortest path. Borgers and Timmermans (1986) used a somewhat similar approach, although they used shopping street segments as destinations, not stores and they did not include a measure of congruity. On the other hand, they predicted route choice by means of a probabilistic model with distance as dependent variable. The model proposed by Helbing (1992) is also based on the assumptions that pedestrians' behavior in a shopping area will be determined mainly by their demand. Although subsequent models proposed by Haklay et al. (2001), Dijkstra et al. (2009), Ali & Moulin (2006), and Zachariadis (2007) are more complex, they all assume some list of (shopping) activities to be performed. However, Ali & Moulin (2006) also allow pedestrians moving around to explore the mall.

Borgers and Timmermans (2005), see also Kemperman et al. (2009), proposed a model to simulate individual pedestrians through shopping areas without assuming a list of activities to be performed. Their motivation was twofold: first, a significant portion of pedestrians shopping in downtown shopping areas may have a hedonic shopping motivation with no or only partially planned activities; second, such a model needs less information and can be estimated from data collected, e.g., by tracing smartphone tracks. Data need not be collected by means of questionnaires or interviews. The model did not include visiting specific outlets, however, the model was extended in order to predict this as well (Borgers and Timmermans, 2012). The latter model was developed using data collected in the downtown shopping area of Eindhoven in 2002. This paper presents a more extended version of this model, estimated and validated on additional data, collected in Maastricht (2003) and Eindhoven (2007).

The remainder of this paper is organized as follows. The model will be specified in the next section. Data collection and estimation results will be discussed in the third section and the simulation results will be presented in section 4. Section 5 concludes the paper.

2 Methodology

In line with the literature on wayfinding (e.g. Gopal et al., 1989), it is assumed that a pedestrian decides about his/her walking direction at intersections of shopping streets. Once a walking direction has been chosen, the pedestrian walks into that direction until he/she reaches the next intersection. There, the pedestrian will choose a walking direction again and so on. When walking, the pedestrian passes outlets and he/she may decide to enter an outlet. If the pedestrian entered an outlet, he/she has to decide about a walking direction when leaving the outlet. If the outlet has multiple exits, the pedestrian has to choose an exit first.

It is also assumed that a pedestrian enters the shopping area at an entry point and that he/she leaves the shopping area at the same point or another exit point nearby the entry point. When entering the shopping area, it is assumed the pedestrian wants to move away from the entry point into the shopping area. This desire diminishes when moving further away from the entry point. At some moment, a desire to move back to the entry point will become apparent. Eventually, the pedestrian will reach an exit point and may leave the shopping area.

The model is based on a network of streets representing the shopping area. In the next subsection, the network will be explained. The model representing pedestrians' behavior will be presented in the second subsection.

2.1 The network

To model pedestrian behavior in a shopping area, the shopping area is represented by a network of shopping streets. Locations of outlets and entry points to the area must be identified and information about the outlets is required. The network consists of nodes and links. A link represents a (semi-)public space where pedestrians can walk, like a part of a street or square, a passage through a mall, or public staircases. A public square is represented by diagonal links and links connecting all outlets along the square. Entrances to outlets are also represented by links.

Primary nodes represent intersections of shopping street segments. Secondary nodes connect the entrances of the outlets to the shopping streets. Outlet nodes represent outlets and contain information about the outlet. Finally, nodes may represent entry points: locations where pedestrians enter or leave the shopping area. A three-dimensional network is used to represent multiple storey shopping environments. However, the network is represented in a 2D plane. To compute distances, only links connecting primary and/or secondary nodes are taken into account, implying that walking into an outlet, moving around in the outlet, and leaving the outlet is not taken into consideration.

2.2 Model Specification

The decisions made at primary and secondary nodes in the network are modeled by means of random utility models. According to these models, the probability a particular alternative will be chosen depends on the utility of each alternative that can be chosen. The utility of alternative i consists of a structural and a random part (U_i=V_i+ε_i). In this study, the multinomial logit model ($p_i = e^{\mu V_i}/\sum_j e^{\mu V_j}$) is used (see e.g. Train, 2003). Successively, the relevant decisions and their corresponding structural utility components will be discussed. As it is assumed that decisions taken previously on the route from entry point e to the current node n may affect decision making at node n, the history of a pedestrian's trajectory is continuously updated.

2.2.1 Decisions at primary nodes

At a primary node, each adjacent link represents a walking direction and the pedestrian has to choose one of them. The structural utility is defined as:

$$V_{e-n,i}^{Dir} = V_{e-n,i}^{Route} + V_{e-n,i}^{Finish} + V_{e-n,i}^{Attr} + V_{e-n,i}^{Outlets} \tag{1}$$

where $V_{e-n,i}^{Dir}$ is the structural utility of walking direction i at node n, given the pedestrian's entry point is e. Note that indices referring to individual pedestrians have been omitted for reasons of simplicity. The first right hand component of Eq. 1 is defined as:

$$V_{e-n,i}^{Route} = \alpha_1 Dist_{e-n,i} + \alpha_2 LLoS_{e-n,i} + \alpha_3 Stair_i + \alpha_4 Square_i + \alpha_5 Before_{e-n,i} + \alpha_6 Retrace1_{e-n,i} + \alpha_7 Retrace2_{e-n,i} + \alpha_8 Forward_{e-n,i} + \alpha_9 Right_{e-n,i} + \alpha_{10} Left_{e-n,i} \quad (2)$$

$Dist_{e-n,i}$ is related to the assumption that in the beginning the pedestrian wants to move away from the entry point and later, he/she wants to move into the direction of the entry point. It is defined as:

$$Dist_{e-n,i} = \frac{D^{Threshold} - D_{e-n}}{100} \mathcal{D}_{e,n,i} \quad (3)$$

where D_{e-n} is the distance walked so far from entry point *e* to current node *n*; $D^{Threshold}$ is the threshold distance and $\mathcal{D}_{e,n,i}$=1 if choosing walking direction *i* implies moving at least 10m further away from entry node *e*; -1 if direction *i* implies moving at least 10m towards *e* and 0 otherwise. Thus, if the pedestrian has walked less than the threshold distance and moves further away from the entry point, $Dist_{e-n,i}$ will be positive; if he/she has walked more than the threshold distance and moves further away from the entry point, $Dist_{e-n,i}$ will be negative. A positive α_1 implies that pedestrians tend to move away from the entry point in the beginning of their shopping trip and to move back later on.

$LLoS_{e-n,i}$ is the length of the line of sight when choosing direction *i*. $Stair_i$=1 if walking direction *i* represents a staircase, elevator, or escalator and 0 otherwise. $Square_i$=1 if choosing alternative *i* implies crossing a square; otherwise $Square_i$=0. $Before_{e-n,i}$=1 if the pedestrian has chosen this walking direction *i* before at node n and 0 otherwise. It may be expected that choosing the same direction again represents a negative utility. The pedestrian may also have traversed the segment from the opposite site, in that case $Retrace1_{e-n,i}$=1 and 0 otherwise. As pedestrians may retrace their route back to the beginning, the corresponding parameter may be positive. However, retracing a segment more than once ($Retrace2_{e-n,i}$) is expected to generate a negative utility.

The other components in Eq. 2 indicate whether choosing direction *i* implies moving forward, to the right, or to the left. The utility of moving backward is set to 0.0. Choosing this option is expected to be less attractive than moving forward or turning to the left or right.

If walking direction *i* represents a direct link to an entry point of the shopping area, the pedestrian has the opportunity to finish the shopping trip, only if this point was the entry point or if it is located close to that point (within 150m). The utility of finishing the shopping trip is defined as:

$$V_{e-n,i}^{Finish} = \alpha_{11}\, Finish_{n,\mathrm{i}} + \alpha_{12} D_{e-n}^{NotFinish} + \alpha_{13} D_{n,e} \tag{4}$$

$Finish_{n,i}$ is equal to 1, it is a dummy to measure the base utility of leaving the shopping area (α_{11}). However, leaving the shopping area shortly after entering the shopping area may be unlikely. Therefore $D_{e-n}^{NotFinish}$ is equal to 1 if D_{e-n}, the distance walked so far, is not more than $D_{max}^{NotFinish}$m; otherwise $D_{e-n}^{NotFinish} = 0$. In fact, α_{12} adapts the base utility if the pedestrian is still near the entry point of the shopping area. It is expected that the utility of leaving the shopping area decreases with increasing distance from the current node to the entry node e ($D_{n,e}$). If $D_{n,e}$ is equal to 0, the current node is the entry node. If there is no opportunity to leave the shopping area, $V_{e-n,i}^{Finish} = 0$. On the other hand, if choosing direction i implies leaving the shopping area, all other components in Eq. 1 are equal to 0.

The third component of the structural utility of choosing a walking direction at node *n* is defined by attributes representing features of the shopping streets along the line of sight. The first attribute is related to the type of traffic allowed in the streets. Two types are considered: no traffic allowed (pedestrianized area) and traffic allowed. The second attribute is related to the type of shopping street: indoor or outdoor. For each link in the network, the two attributes are represented by dummy variables: traffic allowed and indoor. As links along the line of sight may differ in terms of their characteristics, the characteristics of the links along the line of sight are aggregated as follows:

$$x_{i,k} = \sum_m \frac{X_{i,k,m}}{\sqrt{m}} \Big/ \sum_m \frac{1}{\sqrt{m}} \tag{5}$$

where $x_{i,k}$ is the aggregated score of the k^{th} characteristic of alternative *i*; $X_{i,km}$ is the value of the k^{th} characteristic at the m^{th} meter along the line of sight in walking direction *i*. The value of $x_{i,k}$ represents the weighted proportion of the line of sight with characteristic *k*. The effects are measured as follows:

$$V_{e-n,i}^{Attr} = \alpha_{14} x_{i,traffic} + \alpha_{15} x_{i,indoor} \tag{6}$$

The fourth component of Eq. 1 is related to the supply of outlets. Overall, it is assumed that the contribution of a specific outlet to the utility of the walking direction increases with increasing floorspace and decreasing distance to the outlet. An outlet may be located along the line of sight or not. If an outlet is located along the line of sight, the distance to the outlet is equal to the shortest distance between the current node *n* and the nearest node the outlet *o* is connected to (D_{n,n_o}). If the outlet is not located along the line of sight, the distance to the outlet is equal to the shortest distance ($D_{n'|i,n_o}$) between the next primary node in walking direction i ($n'|i$) and the nearest node outlet *o* is connected to plus the shortest distance between nodes *n* and $n'|i$. The rationale behind this is that the first option to change direction when walking into direction *i* is at node $n'|i$. So, the distance to an outlet *o* is defined as:

$$d_{n,i,n_o} = \begin{cases} D_{n,n_o} & if\ n_o \in \boldsymbol{LOS}_{n,i} \\ D_{n,n'|i} + D_{n'|i,n_o} & if\ n_o \notin \boldsymbol{LOS}_{n,i} \end{cases} \tag{7}$$

where d_{n,i,n_o} is the distance from node *n* to the nearest node (n_o) outlet *o* is connected to, if walking direction *i* is chosen and $\boldsymbol{LOS}_{n,i}$ is the collection of nodes connected to the line of sight in walking direction *i* at node *n*.

It is assumed that the contribution of outlets in the utility of walking direction *i* can be defined as:

$$V_{e-n,i}^{Outlets} = \sum_{o} (\beta_{T_o} + \beta_{T_o}^{nlos} + \beta_{T_o}^{p} + \beta_{T_o}^{nlos,p} + \beta_{T_o}^{v} + \beta_{T_o}^{indr}) f(F_o)/g(d_{n,\mathrm{i},n_o}) + \beta_{T_o}^{aggl} A_o \tag{8}$$

Where $f(F_o)$ is a function of the floorspace (in m^2) of outlet *o*, $g(d_{n,i,n_o})$ is a function of the distance (in m) to the node outlet *o* is connected to, and A_o is a measure of the accumulation of other outlets of type T_o around outlet *o*, defined as:

$$A_o = \sum_{o' \in \boldsymbol{O}} f(F_{o'}) \frac{g(D_{max}^{Agglom}) - g\left(D_{n_o,n_{o'}}\right)}{g\left(D_{max}^{Agglom}\right)}, \quad \mathrm{o}' \neq \mathrm{o} \tag{9}$$

Where $D_{n_o,n_{o'}}$ is the shortest distance between the nodes outlets o and o' are connected to, D_{max}^{Agglom} is a maximum distance, and the collection $\boldsymbol{O}$ contains all outlets of type T_o within a range of D_{max}^{Agglom} meters from *o*.

The parameters are specific to the type of outlet (β_{T_o}), whether the outlet is not connected to the line of sight ($\beta_{T_o}^{nlos}$), has been passed before ($\beta_{T_o}^{p}$), is not connected to the line of sight and has been passed before ($\beta_{T_o}^{nlos,p}$), has been visited before ($\beta_{T_o}^{v}$), or is located in an indoor shopping environment ($\beta_{T_o}^{indr}$). Furthermore, the parameter $\beta_{T_o}^{aggl}$ measures the agglomeration effect of similar outlets around outlet *o*.

2.2.2 Decisions at secondary nodes

Secondary nodes provide access to outlets. It is assumed that if a pedestrian has reached a secondary node, he/she will consider visiting each outlet (randomly ordered) connected to that node. Dijkstra *et al.* (2009) developed a model to predict whether a pedestrian will be activated or not to visit a store. However, in this study we specified a binary choice model. The utility of visiting outlet *o* ($V_{e-n,o}^{Visit}$) is defined as:

$$V_{e-n,o}^{Visit} = (\beta_{T_o} + \beta_{T_o}^{p} + \beta_{T_o}^{v})\, f(F_o) + \beta_{T_o}^{Ơp} \sqrt{Ơ_{T_o}^{p}} + \beta_{T_o}^{Ơv}\, Ơ_{T_o}^{v} \qquad (10)$$

The parameters β_{T_o}, $\beta_{T_o}^{p}$, and $\beta_{T_o}^{v}$ measure the effect of the floorspace, just as in Eq. 8. However, note that distance now can be ignored as the pedestrian is in front of the outlet. $Ơ_{T_o}^{p}$ is the number of times other outlets of the same type (T_o) have been passed so far and $Ơ_{T_o}^{v}$ is equal to 1 if other outlets of the same type have been visited before, otherwise $Ơ_{T_o}^{v}$ is 0. The utility of not visiting outlet o ($V_{e-n,o}^{NotVisit}$) is:

$$V_{e-n,o}^{NotVisit} = \gamma_1 + \gamma_2 \sqrt{N^v} \qquad (11)$$

Not visiting an outlet is assumed to depend on a general propensity of not visiting outlets (γ_1) which is expected to be positive, and the total number of outlets already visited (N^v). If none of the outlets connected to node *n* will be visited, the pedestrian moves on in the walking direction previously chosen. However, if a pedestrian decided to enter an outlet, he/she will leave the outlet subsequently. In the case of a single entrance/exit outlet, the pedestrian will return to the previous secondary node and will consider visiting other outlets connected to that node, if any. If no other outlets are visited, the pedestrian has to decide in which direction to continue the shopping trip. In general, there will be two options: continue walking in the direction previously chosen, or return into the direction of the previous primary node. In fact, this choice problem is similar to choosing a walking direction at a primary node. Thus, to choose a walking direction, the mechanisms represented by Eq. 1 will be applied (although $V_{e-n,i}^{Finish}$ will be equal to 0 by definition). However, Eq. 2 has to be adapted as follows:

$$V_{e-n,i}^{Route} = \alpha_1 Dist_{e-n,i} + \alpha_2 LLoS_{e-n,i} + \alpha_3 Stair_i + \alpha_4 Square_i + \alpha_{16} Return_{e-n,i} \quad (12)$$

Variable $Return_{e-n,i}$ is equal to 1 if the walking direction is back to where the pedestrian came from and 0 if the pedestrians continues in the other direction.

2.2.3 Decisions in outlets

If a pedestrian entered a multiple entrance/exit outlet, he/she has to choose one of the exits to leave the outlet. By considering each exit as a walking direction, Eq. 1 can be used again, however, both $V_{e-n,i}^{Finish}$ and $V_{e-n,i}^{Attr}$ will be 0. Regarding the utility of outlets ($V_{e-n,i}^{Outlets}$), there is no line of sight in an outlet, thus none of the outlets is located along the line of sight. The route component is now defined as:

$$V_{e-n,i}^{Route} = \alpha_1 Dist_{e-n,i} + \alpha_{17} Exit_i + \alpha_{18} LevelExit_i \quad (13)$$

$Exit_i$ and $LevelExit_i$ are dummy variables indicating whether the exit was also used to enter the outlet and whether the exit is on the same level (storey) as the entrance. If the pedestrian has chosen one of the exits, he/she will arrive at a secondary node and will have to choose a walking direction again.

3 Data collection and model estimation

The data used for this study were collected in the downtown shopping areas of Eindhoven and Maastricht, the Netherlands. In Eindhoven, data were collected in March 2002 during a Friday (including late night shopping) and a Saturday. In Maastricht, late night shopping is on Thursdays. Therefore, data were collected during a Thursday night, a Friday, and a Saturday. Data collection took place in November 2003. A third set of data was collected in Eindhoven in March 2007. In 2005, a large multi-level mall in the northern part of the Eindhoven downtown shopping area was opened.

Although the data might have been collected just by means of observation, the data for this study was collected by interviewing pedestrians when leaving the shopping area at the main entry points. Respondents were asked where they had entered the shopping area, which outlets they had visited and which route they had walked. Over 1700 complete and circular trajectories were collected (694 in Eindhoven, 2002; 452 in Maastricht, 2003; 587 in Eindhoven, 2007). Two thirds of the trajectories were used for estimating the model. Choice sets were created by decomposing the trajectories in choice sets representing choices regarding walking directions, entering outlets (yes or no), choosing an outlet exit, or leaving the shopping area. In total, approximately 215.000 choice sets were generated to estimate the parameters of the model.

The distance threshold ($D^{Threshold}$) in Eq. 3 was set to 200m for pedestrians in Eindhoven and to 350m and 500m for pedestrians entering the western respectively the eastern part of the Maastricht downtown area. Note that the Maastricht downtown area is separated by a river. The maximum value for $D_{max}^{NotFinish}$ (Eq. 4) is set to 100m and the maximum distance D_{max}^{Agglom} in Eq. 9 is set to 50m. The functions *f* and *g* (Eqs. 8 and 9) are defined as respectively the cube root of the floorspace and square root of the distance. These values and functions were obtained by estimating the model for different settings. The parameters $\beta_{T_o}^{indr}$ could not be estimated for each type of outlet separately, therefore one common parameter β^{indr} was estimated. Nlogit (Econometric Software Inc., 2012) was used to estimate the model. Non-significant parameters and wrong sign parameters were removed from the model step by step. A scale factor was included to measure scale differences between Eindhoven and Maastricht. The results are listed in Table 1.

The distance parameter (α_1) is positive indicating that pedestrians first move away from the entry point and then want to move back to the entry point. Regarding the length of the line of sight, the quadratic function $\left(LLoS_{e-n,i} - 400\right)^2$ with $LLoS_{e-n,i}$ being the length of the line of sight (in m) performed better than a linear function, although this effect is significant for Maastricht only. Pedestrians tend to move forward at intersections and the utility of turning right is somewhat higher than that of turning left. Parameter α_{16} is positive, indicating that pedestrians tend to move back into the direction they came from when leaving an outlet (the value of continuing was set to zero). Pedestrians in Maastricht tend to cross a square if possible, in Eindhoven this tendency is weaker. The parameter for choosing a staircase, escalator or elevator (α_3) is positive. This may be counter-intuitive, however, the positive effect may represent the additional utility

Table 1. Parameter estimates

Variable	Param	Value	Contrast[c]	Variable	Param	Value	Contrast[c]
Dist	α_1	.0763*		Finish	α_{11}	9.62*	-.999*
LLOS[a]	α_2	-.31E-5*		$D^{NotFinish}$	α_{12}	-.792*	
Stair	α_3	1.37*		$D_{n,e}$	α_{13}	-.117*	
Square	α_4	.0797*	-.0602*	$x_{traffic}$[b]	α_{14}	-.608*	
Before	α_5	-.571*	-.312*	x_{indoor}	α_{15}	.299*	.216*
Retrace1	α_6	1.19*		Return	α_{16}	.691*	
Retrace2	α_7	-.269*		Exit	α_{17}	1.35*	
Forward	α_8	3.55*		LevelExit	α_{18}	.447*	
Right	α_9	2.83*		Not Visit	γ_1	3.64*	
Left	α_{10}	2.80*		$\sqrt{N^v}$	γ_2	.149*	.205*

*significant at α=0.05.
[a]Maastricht only.
[b]Eindhoven only.
[c]Eindhoven: add to Value, Maastricht: subtract from Value.

of outlets which can be seen at the other storey. Pedestrians do not prefer to choose the same walking direction for a second time at a particular node (see also Zacharias, 2006). However, retracing a previously chosen street segment (α_6) is quite popular, in contrast to retracing the segment a second time (α_7).

If the pedestrian can leave the shopping area, he or she is likely to do so given the high value for α_{11}, especially when the pedestrian entered the shopping area at the same point: then $D_{n,e}$ is equal to zero and α_{13} has no effect. However, if the pedestrian just entered the shopping area (walked less than 100m), the intention to leave the shopping area is weaker (α_{12}).

In Eindhoven, pedestrians prefer pedestrianized street segments over segments allowing other traffic as well (α_{14}). This does not hold for Maastricht, probably because a number of segments giving access to the Maastricht shopping area are mixed traffic streets. On the other hand, the effect of indoor shopping streets (α_{15}) is stronger in Eindhoven than in Maastricht. This is as expected because at the time of data collection, the indoor shopping arcades in Eindhoven were much more attractive than the one in Maastricht.

For outlet type 'Other', the general contribution in the utility of a walking direction and the utility of entering an outlet (β_{T_o}) is set to zero. Except for the 'Services', the other types of outlets have a positive sign. For most types of outlets, the attraction decreases if the outlet is not connected

Table 1. (cont.)

Type of outlets	Overall	Outlet not connected to L-O-S	Outlet passed	Outlet not connected to L-O-S & passed	Outlet visited	Agglomeration	Other outlets passed	Other outlets visited
T_o	β_{T_o}	$\beta_{T_o}^{nlos}$	$\beta_{T_o}^{p}$	$\beta_{T_o}^{nlos,p}$	$\beta_{T_o}^{v}$	$\beta_{T_o}^{aggl}$	$\beta_{T_o}^{Op}$	$\beta_{T_o}^{Ov}$
Groceries	.262[a]	-.055[a]	-.273[a]		-.776[a]			.658[a]
Personal Care	.277[a]		-.169[a]	.045[a]	-.522[a]		-.405[a]	.670[a]
Dept. Stores	.205[a]	-.137[a]	-.073[a]	.150[a]	-.196[a]	.033[a]	-.169[a]	
Clothing	.008[a]	-.012[a]	-.039[a]		-.264[a]		-.085[a]	.946[a]
Clothing Large	.183[a]	-.183[a]	-.263[a]	.176[a]	-.190[a]	.058[a]		1.40[a]
Other Fashion	.021[c]	-.093[a]	-.139[a]	.068[a]		.039[a]	-.191[a]	1.24[a]
Household	.107[a]			.074[a]	-.153[b]		-.166[a]	
Sports	.195[a]	-.164[a]		.187[a]	-1.01[a]		-.722[a]	2.14[a]
Reading	.250[a]	-.209[a]	-.168[a]	.251[a]	-.822[a]		-.268[a]	1.18[a]
Electronics	.285[a]		-.141[a]		-.232[a]		-.554[a]	1.43[a]
Other Retail	0.0	-.059[a]	-.041		-.267[a]			1.99[a]
Market	.518[a]	-.144[a]	-.428[a]	1.44[a]	-1.96[a]			
Bars/Restaurants	.075[a]	-.092[a]	-.086[a]		-1.24[a]		-.121[a]	
Services	-.059[a]			-.157[a]	-.805[a]	.033[a]		
Indoor (β^{indr})	.010[a]							

[a]significant at α=0.05.
[b]significant at α=0.10.
[c]significant at α=0.15.

to the line of sight ($\beta_{T_o}^{nlos}$). Also, the attraction decreases for most types if the outlet has been passed in the prior part of the trajectory ($\beta_{T_o}^{p}$). If both effects occur for a particular outlet, the negative effects may be compen sated for partially ($\beta_{T_o}^{nlos,p}$). For some types (e.g. 'Market'), the sum of these three effects is positive. Apparently, the concerning outlets increased in attractiveness and may be visited later. If an outlet already has been visited ($\beta_{T_o}^{v}$), its attraction decreases for almost all types. In general, outlets in indoor shopping areas are more attractive ($\beta_{T_o}^{indr}$). The types 'Department stores', 'Large clothing stores', 'Other fashion', and 'Services' are sensitive to agglomeration effects ($\beta_{T_o}^{aggl}$). The utility of entering an outlet decreases if more other outlets of the same type have been passed ($\beta_{T_o}^{Np}$), but increases if at least one other outlet of the same type has been visited ($\beta_{T_o}^{Nv}$).

The base utility of not entering an outlet has a relatively high positive value (γ_1). This implies that the probability of visiting an outlet when passing it is rather small. In Eindhoven, the utility of not entering an outlet increases with increasing number of outlets visited (γ_2). In Maastricht, the tendency is that pedestrians enter more outlets with increasing number of outlets visited. Overall, the model performs well: parameters have expected signs and the rho^2 is equal to 0.88. According to the scale-parameter μ, all parameters have to be multiplied by 1.234 in the case of Maastricht.

4 Simulation

The estimated model was used to simulate the routes of the respondents in the estimation set and the holdout set. For each respondent, the entry node was taken as the starting position. Then, at each decision point, the model was used to compute the probabilities for each alternative. By means of Monte Carlo Simulation, one of the alternatives was selected. This process was repeated until a complete route was generated. For each respondent, 50 routes were simulated and each route was weighted by 1/50. The simulation results can be compared with the observed routes. Figs. 1-3 display the results; estimation and holdout sets have been summed.

Table 2 summarizes some statistics regarding the observed and simulated routes. Mean route length and average number of outlets visited per route are well reproduced. Also aggregated link and outlet loadings match well according to the correlation coefficients. The indices have been scaled down to 100 pedestrians for the purpose of comparison. Over all, the results are satisfying, although there is a structural overprediction of the route length. Another structural trend is that the maximum number of pedestrians across the links in the network is underpredicted. This is also visible in the Figs. 1-3. Apparently, the probability of choosing less frequented street segments is too high.

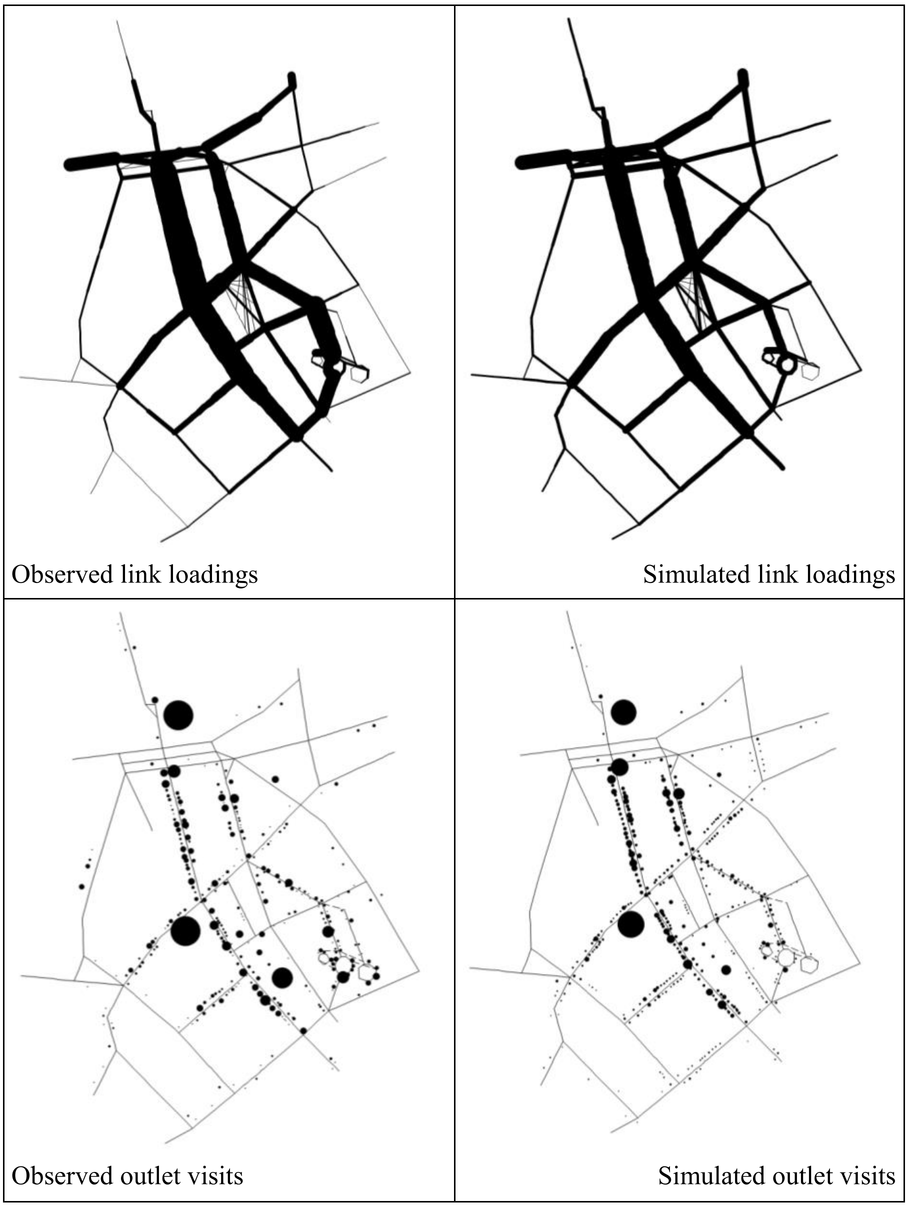

Fig. 1. Link loadings and outlet visits for Eindhoven 2002

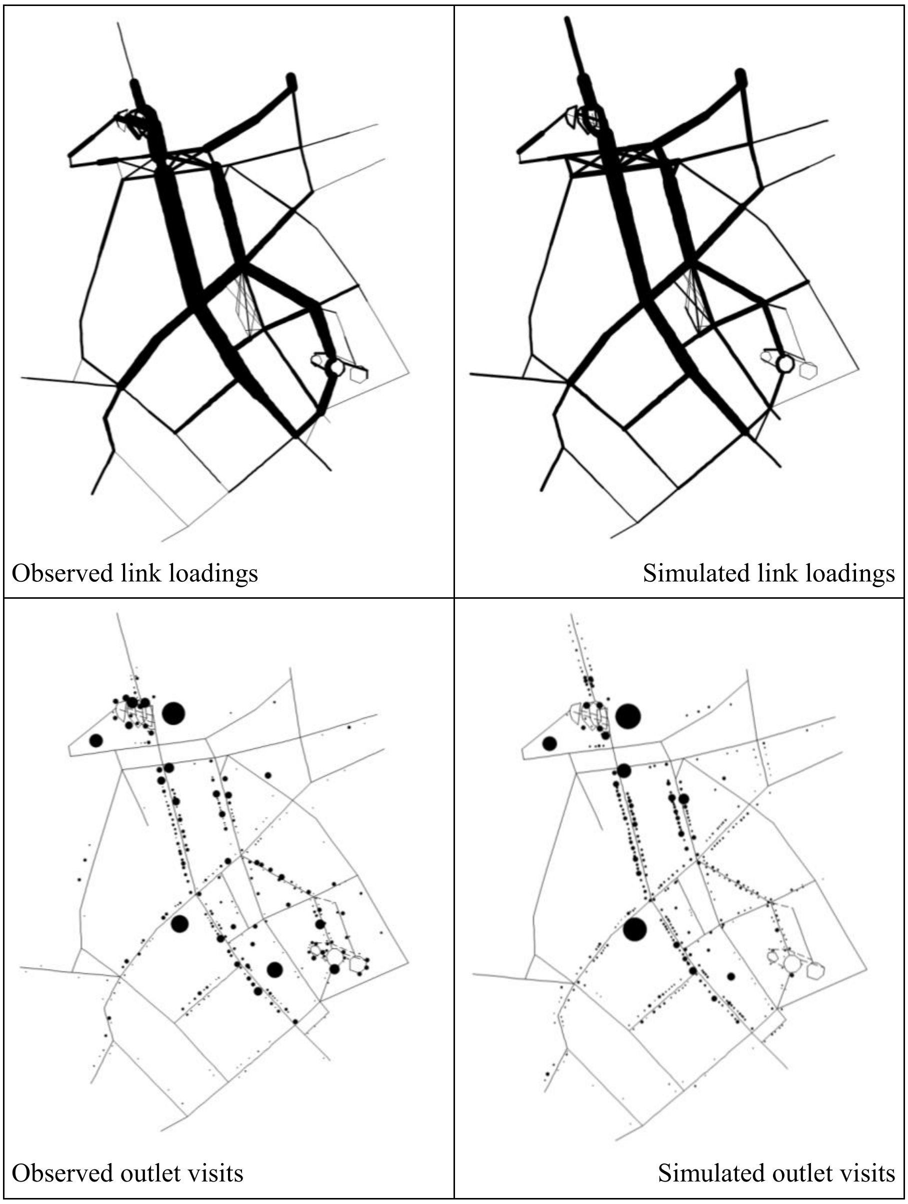

Fig. 2. Link loadings and outlet visits for Eindhoven 2007

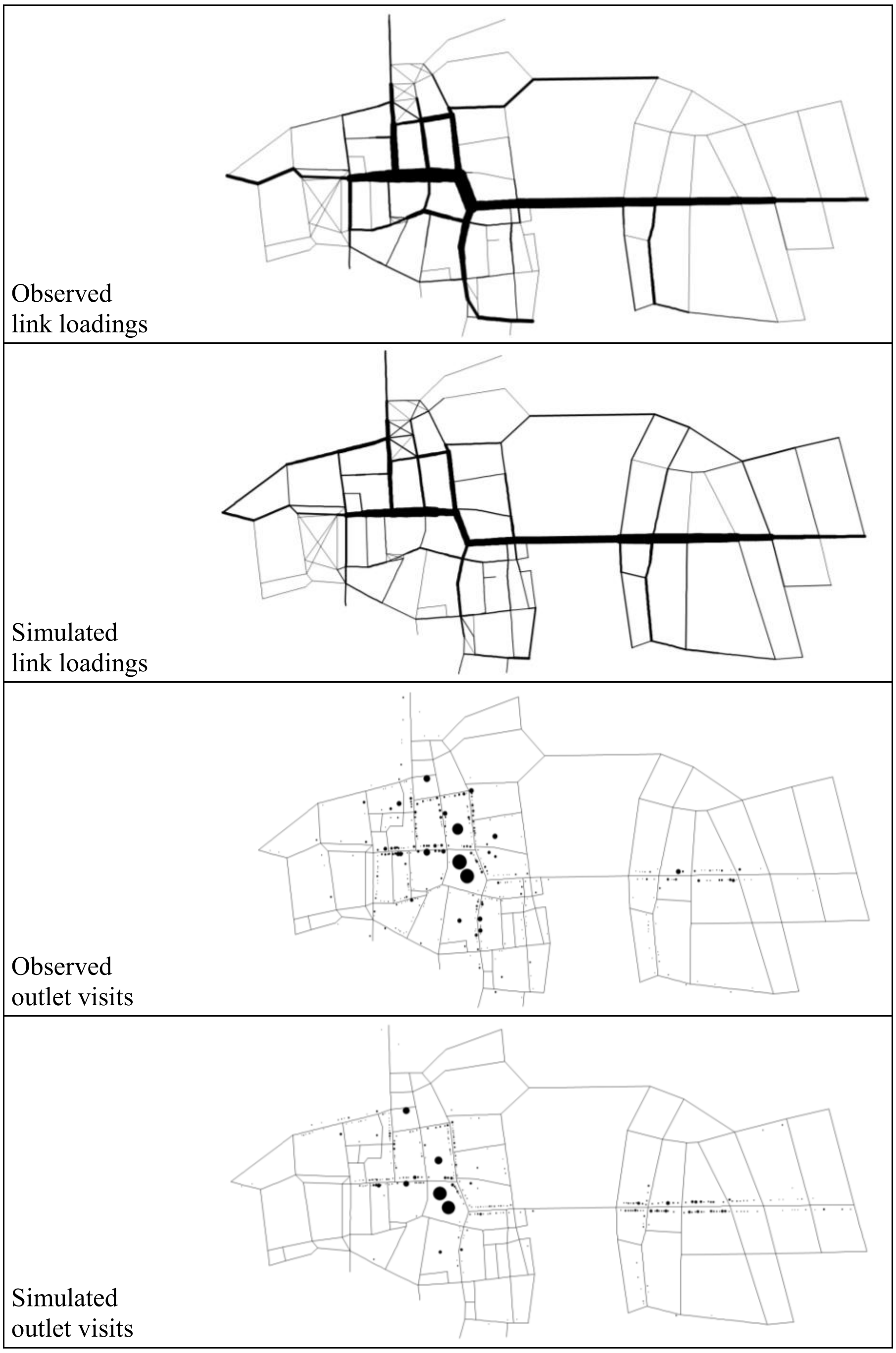

Fig. 3. Link loadings and outlet visits for Maastricht 2003

Table 2. Simulation results (per 100 simulated routes)

	Eindhoven 2002		Eindhoven 2007		Maastricht 2003	
ESTIMATION SET	Ob-served	Simu-lated	Ob-served	Simu-lated	Ob-served	Simu-lated
mean route length (m)	1101	1160	1051	1126	1545	1556
mean # pedestrians/link	18.9	17.7	15.7	15.3	16.6	14.3
max. # pedestrians links	79.0	61.9	68.2	59.3	87.4	62.3
number of pedestrians per link:						
- correlation	0.951		0.945		0.895	
- abs. difference	11.2		9.1		12.6	
average # outlets/route	3.01	3.12	3.35	3.28	3.67	3.47
max. # visitors outlets	29.56	28.00	29.77	31.43	34.22	32.08
number of visits per outlet:						
- correlation	0.906		0.876		0.949	
- abs. difference	0.53		0.47		0.33	
number of visits/type of outlets:						
- correlation	0.965		0.967		0.945	
- abs. difference	5.32		3.68		5.26	
HOLDOUT SET	Ob-served	Simu-lated	Ob-served	Simu-lated	Ob-served	Simu-lated
mean route length (m)	1088	1184	1127	1155	1524	1560
mean # pedestrians/link	18.2	18.1	16.9	15.6	16.4	14.3
max. # pedestrians links	82.0	62.2	66.0	55.8	95.4	63.6
number of pedestrians per link:						
- correlation	0.943		0.936		0.876	
- abs. difference	11.1		10.1		14.4	
average # outlets/route	3.07	3.22	3.28	3.26	4.05	3.49
max. # visitors outlets	34.56	28.68	25.77	30.62	38.4	33.1
number of visits per outlet:						
- correlation	0.906		0.823		0.902	
- abs. difference	0.53		0.54		0.41	
number of visits/type of outlets:						
- correlation	0.933		0.957		0.931	
- abs. difference	6.55		4.24		8.66	

5 Conclusions and recommendations

The model presented in this paper is able to predict route choice behavior and shopping behavior in downtown shopping areas. At each intersection of shopping streets, the pedestrian chooses one of the adjacent street segments as a walking direction. While moving through a street segment, the pedestrian may choose to enter each of the outlets located along the segment. Data to estimate the model can be collected in different ways, for example, by means of tracing pedestrians' trajectories. The model performs relatively well and is sensitive to the location, size, and type of stores. Therefore, the model can be used to assess the likely effects of store replacements, extensions of the shopping area, and changes in the street network.

Most other models that have been developed to predict pedestrian behavior in shopping areas assume a list of products to be bought or activities to be performed. This model, similar to the model proposed by Zhu and Timmermans (2009), does not require this kind of information. On the other hand, this model differs from the model by Zhu and Timmermans (2009) in the sense that all decisions are integrated in one model such that all parameters are simultaneously estimated. Furthermore, unlike most other models, the model has been estimated using data from two cities. For one of these cities, the data was collected prior to and after a large multi-level mall was added to the downtown shopping area.

The most important recommendation for future research is to investigate why to model tends to underpredict the choice of the most frequented shopping streets. Are important variables missing? Is a more advanced type of model required? Should data from other downtown shopping areas with different spatial structures be included in the estimation of the model?

References

Ali, W., & Moulin, B. (2006). How artificial intelligent agents do shopping in a virtual mall: a 'believable' and 'usable' multiagent-based simulation of customers' shopping behavior in a mall. In L. Lamontagne & M. Marchand (Eds.), Canadian AI 2006 (pp. 73-85). Berlin, Heidelberg: Springer-Verlag.

Borgers, A., & Timmermans, H.J.P. (1986). City centre entry points, store location patterns and pedestrian route choice behaviour: a microlevel simulation model. Socio-Economic Planning Sciences, 20(1), 25-31.

Borgers, A., & Timmermans, H. (2005). Modelling pedestrian behaviour in downtown shopping areas. Paper presented at the 9th International Conference on Computers in Urban Planning and Urban Management, London.

Borgers, A., & Timmermans, H. (2012). Shopping behaviour in downtown shopping areas: a detailed pedestrian model. Paper presented at the 17th International Conference of Hong Kong Society for Transportation Studies, Hong Kong.

Crask, M.R. (1979). A simulation model of patronage behaviour within shopping centers. Decision Science, 10(1), 1-15.

Dijkstra, J., Timmermans, H., De Vries, B. (2009). Modeling impulse and non-impulse store choice processes in a multi-agent simulation of pedestrian activity in shopping environments. In H. Timmermans (Ed.), Pedestrian behaviour: Models, data collection and applications (pp. 63-85). Bingley: Emerald.

Econometric Software Inc. (2012). NLOGIT 5 [Computer software]. Plainview, New York.

Gopal, S., Klatzky, R.L., Smith, T.R. (1989). Navigator: a psychologically based model of environmental learning through navigation. Journal of Environmental Psychology, 9, 309-331.

Haklay, M., O'Sullivan, D., Thurstain-Goodwin, M., Schelhorn, T. (2001). "So go downtown": simulating pedestrian movement in town centers. Environment and Planning B, 28, 343-359. doi:10.1068/b2758t.

Helbing, D. (1992). Models for pedestrian behavior. In Natural Structures Part II, Sonderforschungsbereich 230 (pp. 93-98). Stuttgart.

Kemperman, A.D.A.M., Borgers, A.W.J., Timmermans, H.J.P. (2009). Tourist shopping behavior in a historic downtown area. Tourism Management, 30, 208-218. doi:10.1016/j.tourman.2008.06.002.

Train, K.E. (2003). Discrete choice methods with simulation. Cambridge: Cambridge University Press.

Zachariadis, V., (2007). Modelling pedestrian movement and choices from micro to urban scale: Issues, patterns and emergence. Paper presented at the 10th International Conference on Computers in Urban Planning and Urban Management, Iguassu Falls, Brasil.

Zacharias, J. (2006). Exploratory spatial behaviour in real and virtual environments. Landscape and Urban Planning. 78, 1-13. Doi:10.1016/j.landurbplan.2005.05.002.

Zhu, W., & Timmermans, H. (2009). Modeling and simulating pedestrian shopping behaviour based on principles of bounded rationality. In H. Timmermans (Ed.), Pedestrian behaviour: Models, data collection and applications (pp. 63-85). Bingley: Emerald.