Model of Pedestrian Demand (MoPeD 2.0)

Qin Zhang, Kelly Clifton and Rolf Moeckel

Model Overview

The framework of MoPeD is illustrated in Figure 1. The pedestrian prediction tool is integrated with a four-step urban model, in this case, the Portland Metro Regional Model. MoPeD starts with trip generation at the finer spatial resolution. Next, MoPeD adds a step to separate walk trips from other trips. Walk trips are handled by MoPeD, while non-walk trips are processed by the existing urban travel demand model. In the traditional four-step modeling process, mode choice commonly follows the trip distribution.

Nevertheless, this opposite order or model steps was chosen in MoPeD because destination choice works substantially different for walk trips and non-walk trips. By modeling the choice walk/non-walk first, the conditions for the destination choice model are largely improved. Also, since MoPeD implements at a fine spatial resolution, by doing mode choice prior to the destination choice can avoid dealing with massive distance matrices.

MoPeD employs mode-choice models and destination choice models, which were estimated with data from the Oregon Household Activity Survey (OHAS). OHAS is a one-day household travel survey collected for the entire state, including the Portland metropolitan region. The survey was conducted in fall 2011 and included 6,450 households in the Portland Metropolitan area. Person and household characteristics and their travel behavior were collected. The built environment data used in this study for Portland, Oregon was provided by the Portland Metro MPO. The scale of the data is the PAZ level, which consists of a grid of 80 by 80 meters. For each zone, the population and the number of employees by eight job types were provided.

MoPeD has eight trip purposes, including home-based work, home-based school, home-based college, home-based shopping, home-based recreation, home-based other, non-home-based work and non-home-based other. As the purpose of this study is to assess the commuter trips, estimation results for the home-based work trips (HBW) are presented in this paper.

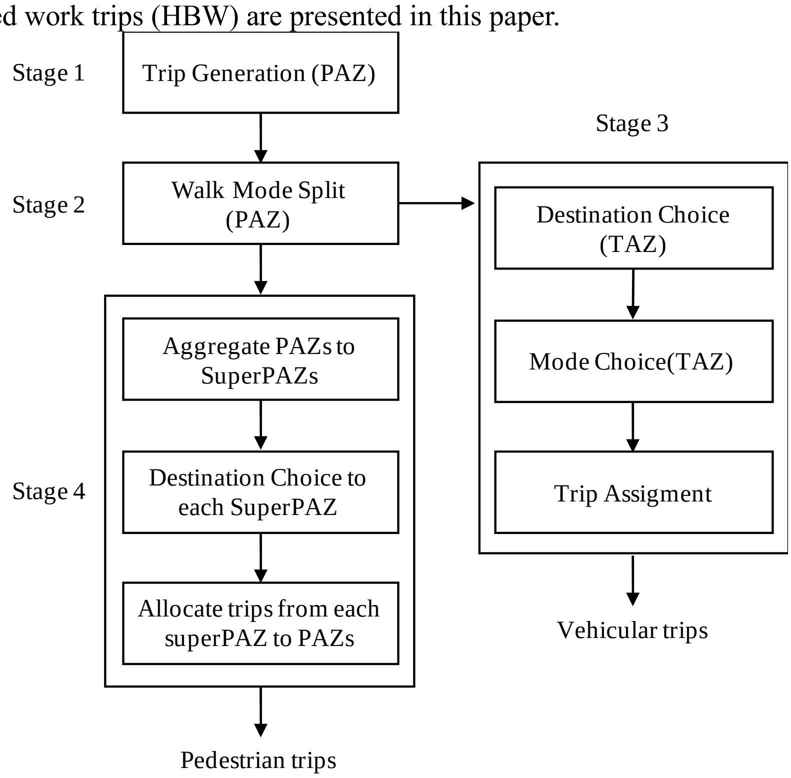

FIGURE 1 Modeling framework of MoPeD (adopted from (Clifton et al., 2016b) **)**

Trip Generation

For trip generation, we adopted Metro's existing trip production models, which use cross-classification to calculate trip production for all home-based purposes (Vogt et al., 2015). While Metro designed model for TAZ-level inputs, we assumed model scalability and applied the trip rates at the PAZ level. Take commuting trips as an example, HBW trips are produced based on the number of workers in a household. For each zone, the number of households by the number of workers (0 workers, 1 worker, 2 workers, and 3+ workers) is multiplied with the corresponding production rate. Afterwards, the number of trips is scaled up to match the regional total number of jobs by applying a calibration factor of 1.36 for HBW trips. For the entire Portland Metro area, the model generates a total of 1,039,872 HBW trips.

Pedestrian Accessibility

Walking behavior is highly correlated to built environment variables. In MoPeD 1.0 (Clifton et al., 2016b), the pedestrian index of the environment (PIE) was used. It consisted of six measures that included activity density, transit access, block size, sidewalk extent (miles of continuous sidewalks within 0.25 mile), bicycle facilities, and urban living infrastructure which means the number of shopping and service destinations in the neighborhood. With a rich spatial database, PIE was able to better represent walking behavior. However, it was less transferable to other applications due to the requirements of detail land use data at a fine spatial resolution. In addition, it was difficult to assess future scenarios and policies, because it was challenging to predict the specific changes to the built environment such as sidewalks and cycle lanes at this fine spatial resolution. In MoPeD 2.0, we contribute a more transferable measurement of pedestrian activity that is more usable for policy analysis.

The new pedestrian accessibility that is used in mode choice is defined as activity density (employment and population) that can be reached within 800-meters network distance. Often, accessibility is calculated based on a circular buffer (Greenwald & Boarnet, 2001; Lee & Moudon, 2006). However, it neglects that people do not make travel decisions in a circular buffer, but rather follow street grids which can be imperfect and lack connectivity. The network distance to reach the outer rim of the buffer area is likely to be longer than the radius of the circle. To overcome the limitation, we use network-distance based isochrones instead of circular buffers. The isochrones are also known as the pedestrian catchment area (PCA). In MoPeD 2.0, we generated the PCA for each PAZ in the sample using a fixed network distance of 800 meters. According to OHAS data, 800-meter distance covers 80% of the walk trip distances and 800-meter distance is equivalent to about 10 minutes' walk which is sufficient to represent the pedestrian accessibility in the neighborhood. The method is based on street segments to generate the buffer polygon. For this task, every block that is enclosed by street segments becomes part of the polygon. If the street segment does not enclose an area, a buffer of 25 meters of the street segment is added to the polygon. We completed this task for the urban growth boundary of Portland that consists of about 160,000 PAZ. Afterwards, we calculated the total number of jobs by type and population that live within each polygon, resulting in the measurement of pedestrian accessibility within 800 meters for the mode choice model. Results of pedestrian accessibility for the Portland metropolitan area are shown in Figure 2.

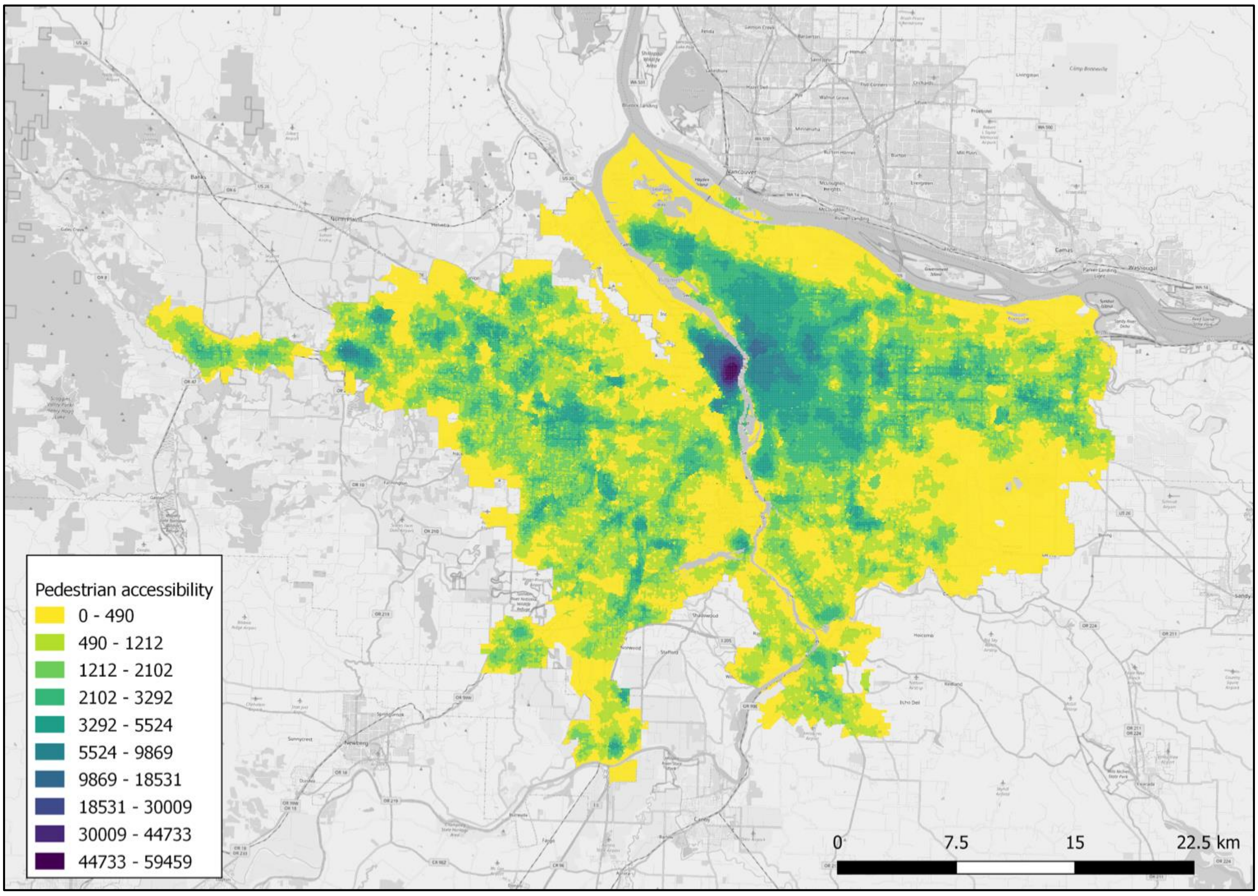

FIGURE 2 Map of the pedestrian accessibility

Walk Mode Choice Model

MoPeD employs a binary logit model to estimate the probability of choosing to walk. Note that in contrast to traditional four-step models, the MoPeD mode choice model does not know the destination yet. After removing the non-significant variables, the final model estimation results are shown below.

The models include three household attributes (income category, number of vehicles, and children) and pedestrian accessibility as independent variables. Pedestrian accessibility was transformed to log-form to improve the model fit. It shows a positive impact in the model, which indicates that households living in denser neighborhoods with better street networks tend to be more likely to walk. The log-transformation suggests that differences in pedestrian accessibility matter a lot at the lower end of accessibilities. Once a certain level of pedestrian accessibility has been reached, additional growth in accessibility has less impact on the likelihood to walk. This suggests that pedestrians need a certain level of accessibility. Once this level is satisfied, commuters are much more likely to walk to work. The statistical significance also indicates that the new measurement of built environment is a good indicator of walking activity while controlling for all other variables. The R^2 appears low in comparison to traditional mode choice models. However, traditional mode choice models appear after destination choice, and thereby, can model the walk share based on travel distance. In this case, the walk share is modeled first to skim off walk trips from the total trip production. As the travel distance is not known yet, the challenge for MoPeD's mode choice model is substantially higher. Figure 3 presents how the walk share spatially distributed in the Portland metropolitan area.

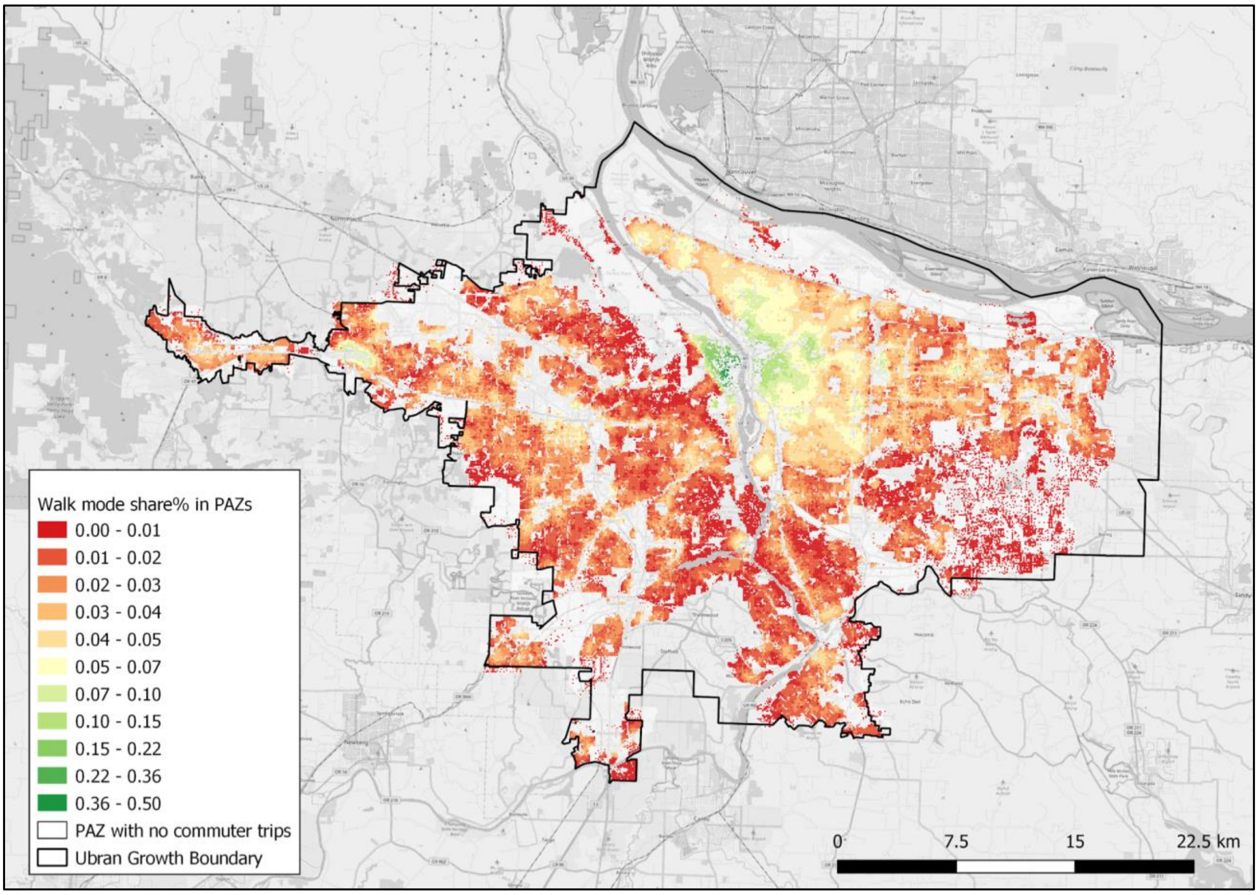

FIGURE 3 Estimated walk mode share of commute trips in the Portland metropolitan area

Walk Destination Choice Model

For the estimation of the choice model, the definition of the choice set is a major challenge. There is little literature regarding this topic for pedestrian models, and there is no conclusive evidence of the preferable approach. The debate continues in the literature between different sampling methods or to include the whole universe of choices. Clifton et al. (K. J. Clifton et al., 2016a) used a random sample of 10 trips that were shorter than the distance of the 99% percentile in the observed data, Berjisian and Habibian (Berjisian & Habibian, 2019) used a 90% percentile threshold with a complete sample at a different spatial scale. We took the percentile 99% of the network distance that was 4.8 kilometers, and we used the whole choice set to estimate the destination choice model.

Using PAZ, the choice set would include an average of 10,000 possible destinations. Such a large choice set violates the assumption of discrete choice models that the number of choices should be small enough to be comparable against each other. Therefore, we aggregated the PAZ structure to a new scale that we called superPAZ. A superPAZ consists of 5 by 5 PAZ. For the purpose of estimation, we calculated the network distance shortest path between all superPAZ. Next, we assigned origin and destination superPAZ to each observation in the OHAS survey data for model estimation. In terms of superPAZ, the possible choice set for each trip is about 576 (all superPAZ within a 4.8-kilometer buffer). By aggregating PAZ to superPAZ, the issue of very large samples was compensated to some degree.

The specification (see equation below) was inspired by the previous MoPeD model yet simplified slightly to improve model sensitivities. Newly introduced were a log transformation of total employment, the proportion of industrial jobs, and the distance to be controlled by car ownership. We also introduced a β_0 as a constant of intrazonal trips. The estimation results are presented in Table 2 and Table 3.

$$U_{ij} = \beta_0 + \beta_{dist} DIST_{ij} Auto + \beta_{size} \ln(Total\ jobs_j) + \beta_{industrial} \frac{Industrial\ jobs_j}{Total\ jobs_j}$$

Distance was a significant and sensitive factor in the model. If the destination is one kilometer further, then its probability of being chosen is reduced by 75% compared to a destination at the origin of the trip. Supporting the previous destination choice model, we found a significant interaction between distance and auto ownership. Households with no cars tend to walk slightly longer than those who own cars. The size variable shows a significant and positive impact, while the share of industrial jobs has a barrier impact on choosing a destination. It suggests that destinations with more retail and service jobs and fewer industrial jobs are more attractive to choose. In contrast to the previous destination choice model, we added a dummy variable for checking if the destination zone is equal to the origin zone. In other words, it is a constant for intrazonal trips. The result represents that the origin zone has a higher probability of being chosen. This confirms expectations as we are using a superPAZ of 400 by 400 meters in the destination choice model, which leads to a significant number of intrazonal trips.

Table 4 and Table 5 present the estimation results of the PAZ-level destination choice models. In general, they had low goodness of fit. This could be due to fewer variations across small-scale destination zones or the lack of important factors in the model. Future studies need to investigate more factors, such as micro-level or street-level built environment variables (e.g., pavement condition and the number of trees).

The parameters are calibrated to match the trip length frequency distribution and the average trip length reported for HBW trips in the OHAS data. After calibration, the estimated average trip length is 1.14 km (observed: 1.11km) and the cumulative trip length distribution of both estimated and observed results are similar (Figure 4).

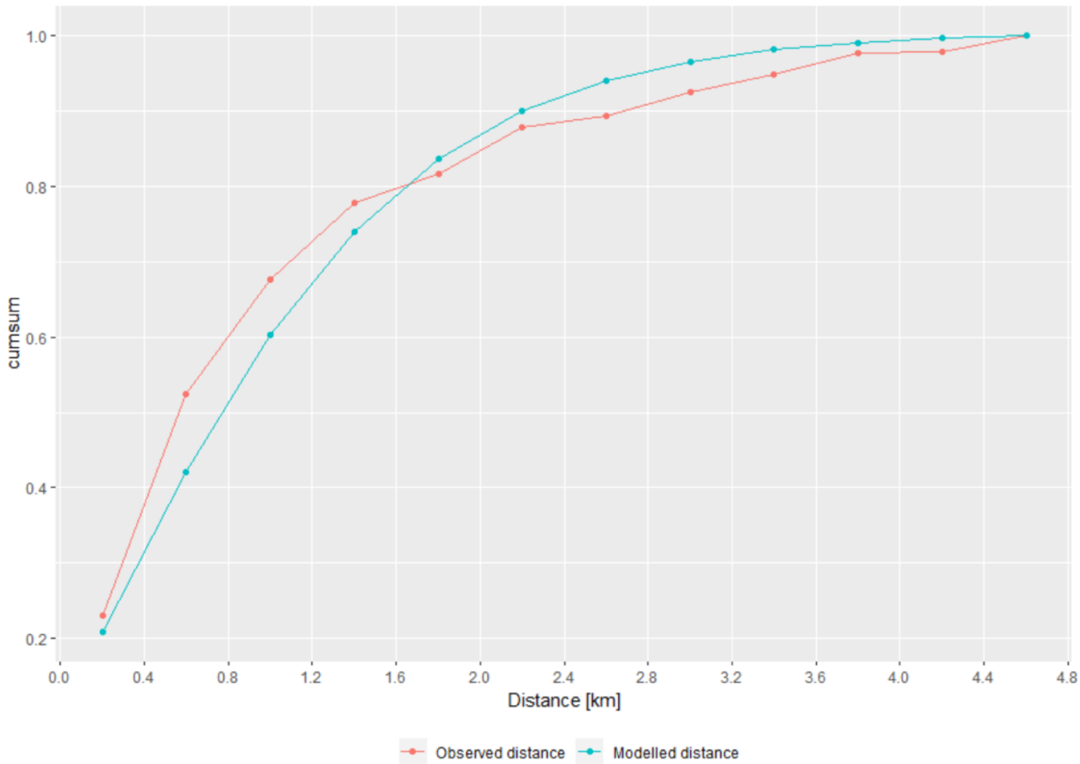

Figure 4 Cumulative trip length frequency distribution of modeled and observed trips

TABLE 1 Binary logit mode choice model estimation in MoPeD

	Home-based purposes			Non-home-based purposes		
	Estimate	$\Pr(>\lvert z\rvert)$		Estimate	$\Pr(>\lvert z\rvert)$	
(intercept)	-8.392	0.000	***	-7.411	0.000	***
Income category 2				-0.205	0.261	
Income category 3				0.222	0.046	*
Income category 4			.	0.448	0.000	***
Number of vehicles (0)	1.001	0.000	***	1.375	0.000	***
Number of vehicles (2)	-0.226	0.002	**	-0.898	0.000	***
Number of vehicles (2+)	-0.394	0.000	***	-0.963	0.000	***
Number of children (0)	-0.554	0.000	***			
Number of children (2)	-0.574	0.000	***			
Number of children (2+)	-0.718	0.000	***			
Child (Yes)			***	-0.162	0.039	*
log(pedestrian accessibility)	0.754	0.000	***	0.686	0.000	***
HBS	1.029	0.000	***			
HBO	1.046	0.000	***			
HBR	1.566	0.000	***			
NHBW				-0.362	0.000	***
Log-Likelihood:	-4189			-2624		
McFadden R^2:	0.135			0.228		
Accuracy	10.14%			11.96%		

Table 2 Results of the SuperPAZ destination choice model (HBW, HBS, HBR)

<table>
<tr><th></th><th colspan="3">HBW</th><th colspan="3">HBS</th><th colspan="3">HBR</th></tr>
<tr><th></th><th>Estimate</th><th>Pr(>|z|)</th><th></th><th>Estimate</th><th>Pr(>|z|)</th><th></th><th>Estimate</th><th>Pr(>|z|)</th><th></th></tr>
<tr><td>Distance (Km)</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>x Auto (Yes)</td><td>-1.536</td><td>0.000</td><td>***</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>x Auto (No)</td><td>-1.372</td><td>0.000</td><td>***</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>x Child (Yes)</td><td></td><td></td><td></td><td>-2.182</td><td>0.000</td><td>***</td><td>-2.321</td><td>0.000</td><td>***</td></tr>
<tr><td>x Child (No)</td><td></td><td></td><td></td><td>-1.776</td><td>0.000</td><td>***</td><td>-1.955</td><td>0.000</td><td>***</td></tr>
<tr><td>Network density (Km)</td><td>0.141</td><td>0.008</td><td>**</td><td>0.049</td><td>0.209</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Size term (ln)</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Service</td><td rowspan="2">0.445</td><td rowspan="2">0.000</td><td rowspan="2">***</td><td></td><td></td><td></td><td rowspan="2">0.133</td><td rowspan="2">0.000</td><td rowspan="2">***</td></tr>
<tr><td>Retail</td><td>0.977</td><td>0.000</td><td>***</td></tr>
<tr><td>Finance</td><td rowspan="3">0.352</td><td rowspan="3">0.000</td><td rowspan="3">***</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Government</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>All other non-industrial</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Household</td><td></td><td></td><td></td><td></td><td></td><td></td><td>0.054</td><td>0.126</td><td></td></tr>
<tr><td>Industrial prop.</td><td>-1.249</td><td>0.021</td><td>*</td><td>-1.306</td><td>0.003</td><td>**</td><td></td><td></td><td></td></tr>
<tr><td>Slope (mean)</td><td>-0.167</td><td>0.024</td><td>*</td><td>-0.386</td><td>0.000</td><td>***</td><td>-0.139</td><td>0.001</td><td>***</td></tr>
<tr><td>Crossing Motorway</td><td>-0.321</td><td>0.18</td><td></td><td>-0.279</td><td>0.149</td><td></td><td>-0.568</td><td>0.032</td><td>*</td></tr>
<tr><td>Park (Yes)</td><td></td><td></td><td></td><td></td><td></td><td></td><td>0.662</td><td>0.000</td><td>***</td></tr>
<tr><td>Sample size</td><td colspan="3">289</td><td colspan="3">646</td><td colspan="3">626</td></tr>
<tr><td>Null model Log-Likelihood:</td><td colspan="3">-1516</td><td colspan="3">-3402</td><td colspan="3">-3261</td></tr>
<tr><td>Final model Log-Likelihood:</td><td colspan="3">-952</td><td colspan="3">-1598</td><td colspan="3">-1910</td></tr>
<tr><td>Pseudo R^2:</td><td colspan="3">0.37</td><td colspan="3">0.53</td><td colspan="3">0.41</td></tr>
</table>

Table 3 Results of the SuperPAZ destination choice model (HBO, NHBW, NHBO)

	HBO			NHBW			NHBO		
	Estimate	Pr(>|z|)		Estimate	Pr(>|z|)		Estimate	Pr(>|z|)	
Distance (Km)	-2.217	0.000	***	-1.883	0.000	***	-2.141	0.000	***
x Auto (Yes)									
x Auto (No)									
x Child (Yes)									
x Child (No)									
Network density (Km)	0.214	0.000	***	0.185	0.000	***	0.184	0.000	***
Size term (ln)									
Service Retail Finance Government All other non-industrial	0.389	0.000	***	0.667	0.000	***	0.516	0.000	***
Household									
Industrial prop.				-0.749	0.117				
Slope (mean)	-0.381	0.000	***	-0.157	0.006	**	-0.060	0.220	
Crossing Motorway	-0.828	0.000	***	-0.718	0.000	***	-1.361	0.000	***
Park (Yes)	0.510	0.000	***						
Sample size	1042			723			697		
Null model Log-Likelihood:	-5438			-3728			-3621		
Final model Log-Likelihood:	-2939			-1774			-1762		
Pseudo R^2:	0.46			0.52			0.51		

Table 4 Results of the PAZ destination choice model (HBW, HBS, HBR)

	HBW			HBS			HBR		
	Estimate	Pr(>|z|)		Estimate	Pr(>|z|)		Estimate	Pr(>|z|)	
OriginPAZ	2.068	0.000	***	0.623	0.132		2.704	0	***
Distance (Km)	-1.335	0.000	***	-2.120	0.000	***	-1.974	0.000	***
Size term (ln)									
Retail	0.541	0.000	***	0.820	0.000	***	0.123	0.028	*
Service				0.188	0.000	***			
Finance									
Government									
Household	-0.433	0.000	***	-0.169	0.000	***	-0.560	0.000	***
Industrial prop.	1.629	0.000	***				-1.602	0.003	**
Park (acre)				-0.651	0.285		1.473	0.013	*
Sample size	289			646			626		
Null model Log-Likelihood:	-939			-2070			-1428		
Final model Log-Likelihood:	-813			-1617			-813		
Pseudo R^2:	0.13			0.22			0.14		

Table 5 Results of the PAZ destination mode choice model (HBO, NHBW, NHBO)

	HBO			NHBW			NHBO		
	Estimate	Pr(>|z|)		Estimate	Pr(>|z|)		Estimate	Pr(>|z|)	
OriginPAZ	3.162	0.000	***	0.654	0.000	***	1.628	0.000	***
Distance (Km)	-2.348	0.000	***	-2.894	0.000	***	-2.163	0.000	***
Size term (ln)									
Retail	0.145	0.000	***	0.316	0.000	***	0.355	0.000	***
Service									
Finance	0.559	0.000	***	0.062	0.016	***	0.137	0.001	***
Government									
Household	-0.507	0.000	***	-0.051	0.097	.	-0.082	0.022	*
Industrial prop.	-0.477	0.086	.				-0.694	0.058	.
Park (Yes)									
Sample size	1042			723			697		
Null model Log-Likelihood:	-3362			-2332			-2255		
Final model Log-Likelihood:	-2617			-2116			-1982		
Pseudo R^2:	0.22			0.09			0.12		

Reference

Berjisian, E., & Habibian, M. (2019). Developing a pedestrian destination choice model using the stratified importance sampling method. *Journal of Transport Geography*, *77*, 39–47. https://doi.org/10.1016/j.jtrangeo.2019.04.009

Bierlaire, M. (2018). PandasBiogeme: a short introduction. In *Technical report TRANSP-OR 181219. Transport and Mobility Laboratory, ENAC, EPFL.* Retrieved from https://wilpf.org/wilpf/who-we-are/

Clifton, K. J., Singleton, P. A., Muhs, C. D., & Schneider, R. J. (2016a). Development of destination choice models for pedestrian travel. *Transportation Research Part A: Policy and Practice*, *94*, 255–265. https://doi.org/10.1016/j.tra.2016.09.017

Clifton, K. J., Singleton, P. A., Muhs, C. D., & Schneider, R. J. (2016b). Representing pedestrian activity in travel demand models: Framework and application. *Journal of Transport Geography*, *52*, 111–122. https://doi.org/10.1016/j.jtrangeo.2016.03.009

Greenwald, M. J., & Boarnet, M. G. (2001). Built Environment as Determinant of Walking Behavior: Analyzing Nonwork Pedestrian Travel in Portland, Oregon. *Transportation Research Record:*

Journal of the Transportation Research Board, *1780*(1), 33–41. https://doi.org/10.3141/1780-05

Lee, C., & Moudon, A. V. (2006). Correlates of Walking for Transportation or Recreation Purposes. *Journal of Physical Activity and Health*, *3*(s1), S77–S98. https://doi.org/10.1123/jpah.3.s1.s77

Vogt, M., Researcher, S., Researcher, S., Bosa, P., Patton, T., Pederson, C., … Stein, B. (2015). *2015 Trip-Based Travel Demand Model Methodology Report Modeling Services Manager*.