MODELLING SHORTEST PATH DECISIONS USING AN ACTIVITY-BASED SEGMENTATION

Katrien Ramaekers, Mario Cools, Sofie Reumers and Geert Wets
Transportation Research Institute
Hasselt University
Wetenschapspark 5 bus 6
B-3590 Diepenbeek, Belgium
E-mail: katrien.ramaekers@uhasselt.be; mario.cools@uhasselt.be; sofie.reumers@student.uhasselt.be; geert.wets@uhasselt.be

KEYWORDS: route choice modelling, shortest path, activity-based approach, trip purpose

ABSTRACT

The aim of this research is to identify the relationship between activity patterns and route choice decisions. The focus is turned to the relationship between the purpose of a trip and whether or not the shortest path is chosen for the relocation. The data for this study were collected in 2006 and 2007 in Flanders, the Dutch speaking and northern part of Belgium. To estimate the relationship between the choice for the shortest path or not and the corresponding activity-travel behaviour a logistic regression model is developed. The results point out that, when analyzing the relationship between the activities of the people and whether or not the shortest path is chosen, there is no significant influence by the activity-based segmentation. However, when the deviation from the shortest path is related to the activities people perform, a significant relationship is found. Furthermore, next to trip-related attributes (trip distance), also socio-demographic variables and geographical differences play an important role.

INTRODUCTION

To support policy makers, traffic and transportation models can be used to make better long-term decisions. On an international level, activity-based models have become the norm to model travel behaviour (Davidson et al., 2007). The most important characteristic of these models is that the travel behaviour of persons or families is a product of the activities that they wish or have to perform, procuring a more realistic description and a better understanding of people's travel behaviour. Because of these advantages, researchers and policy makers in the United States have switched from conventional models to activity-based models. Although this trend is most visible in the United States, a similar evolution can be noticed in Europe.

Governments require reliable predictions of travel behaviour, traffic performance, and traffic safety to support long-term decisions. A better understanding of the events that influence travel behaviour and traffic performance will lead to better forecasts and consequently policy measures that are based on more accurate data.

An important issue that the models should account for is the decision process that one undergoes when conducting a trip. One of many considerations is the route choice. Therefore, it is important to examine to what extent route choice is related to the type of trip. The term ‘type of trip’ indicates the purpose of the trip and submits a link to the pattern of human activities. The term ‘route choice’ includes the many attributes of the path chosen to conduct the trip and establishes the link with the behaviour pattern of an individual.

The aim of this research is to identify statistically significant relationships between activity patterns and the behaviour regarding route choice. The focus is on the relationship between the purpose of the trip and whether or not the shortest route is chosen.

Studies regarding the relationship between the purpose of the trip and the travel distance are frequently available. This indicates that travel distance is considered an important attribute of route choice. Ramming (2002) states that car travellers want to minimize their travel time, regardless of the purpose of the trip. Zhang and Levinson (2008), however, state that the shortest path is chosen less when the purpose of the trip is shopping or paying a visit rather than travelling to work or to a leisure activity. Goldenbeld et al. (2007) study the route choice of car travellers in the Netherlands. Almost half of the respondents indicates ‘shortest path’ as one of the main reasons for choosing a route. When the purpose of the trip is work, ‘shortest path’ is considered more important in route choice than for other trip purposes. In a study of Papinski et al. (2009), route directness is indicated as the second most important factor when choosing a route.

In addition, literature shows that other factors besides route attributes (personal, household and situational characteristics) play a role in the route selection, e.g. Bayarma et al. (2007), Zhang and Levinson (2008) and Scheiner (2009). Therefore, personal characteristics as age, gender, income, profession and province and the situational characteristic time of day are considered in the analyses.

In the next section, the data is described. In the third section, the adopted methodology approach is explained in further detail and the results are discussed. Finally, in the last section, the most important findings are summarised and directions for further research are highlighted.

DATA

The data for this study were collected in 2006 and 2007 in Flanders, the Dutch speaking and northern part of Belgium, in the context of a large scale survey, conducted on 2500 households in the study area. In this section, first, more information concerning the large scale survey is provided. Next, some of the important data processing steps are highlighted and the variables that are considered for the analysis are described.

Data Collection

Traditionally, travel surveys have been collected by paper and pencil or over the phone. The introduction of activity-based analysis, which prompts the need for considerably more detailed data on travel behaviour, identified the advantages of collecting activity or time use diary data (see Ettema et al. (1997) for an overview). At the same time, however, the use of diary data virtually precluded the use of telephone interviews and in addition substantially increased respondent burden and error proneness (see e.g. Dowling and Colman (1995) and Sun et al. (1995)). To avoid or at least reduce such error, computer-assisted diary instruments were developed.

The data for this study stem from a large scale activity-based data collection effort conducted on households since the household context, in which individuals operate, has a very strong influence on individuals' decisions, particularly when household resources are shared, there are shared household responsibilities and there are decisions that are made jointly by multiple household members. The survey used a mixed-mode survey design, using a PDA application on the one hand, and using traditional paper and pencil diaries on the other hand. Cools et al. (2009) demonstrated that the use of this mixed survey design turns out to be a suitable way of collecting detailed information about planned and executed activity-travel behaviour of households as the survey mode has no direct impact on the quantities investigated.

The PDA application, called PARROTS (PDA (Personal Digital Assistant) system for Activity Registration and Recording of Travel Scheduling) has been developed in such that respondents could easily provide information about their activity-travel behaviour (Bellemans et al., 2008). Whenever an activity or trip is registered in PARROTS, a number of attributes for this activity or trip were collected using a customized GUI. The most important activity and trip attributes PARROTS collected are: activity type, date, start and end time, location, mode of transportation, travel time and travel party. Besides PARROTS uses the integrated Global Positioning System (GPS) to automatically record location data. This combination of GPS and diary responses provides great insight into the route choice decision-making process (Papinski et al., 2009). Jan et al. (2000) showed that GPS is a viable tool to study travellers' route choice decisions as GPS can reveal important travel behavioural information that is impossible to discern with earlier conventional survey methods such as interviews, respondent-administered questionnaires, or driver simulators. Moreover, conventional methods have proved burdensome, time consuming, and error prone (Wolf et al., 1999).

Data Processing

In order to analyze the reported and recorded travel data, advanced post-processing is necessary to make the information usable for route choice modelling (Schuessler et al., 2010). In this research only displacements made by car are taken into account. Displacements made with any other mode of transport are filtered out of the database. Next, the GPS-data are compared to the data reported by the respondents in the diaries. If there is a mismatch between both data sources, the displacements are not used in the analyses since it is possible that the reported displacements are incorrect. Furthermore, only respondents that filled in all personal characteristics are considered because these characteristics are used in the analyses. Given the network that is used to analyze the trips is a national network, cross-border displacements are removed from the database.

The data processing step leads to a dataset containing car displacements on the Belgian road network for respondents of whom the personal characteristics are known and for whom the GPS-data is consistent with the data reported in the diaries. The dataset contains 1423 car displacements, made by 299 different respondents.

Data description

This study focuses on the relationship between the purpose of a trip and whether or not the shortest path is chosen. In this paragraph, the variables that are used in the analyses are described.

Roads are divided in three categories, following the functional road classification of Weijermars et al (2008), namely through-roads (primary roads), distributor roads (secondary roads) and access roads (local roads).

Furthermore, trip-related attributes are considered. In literature these attributes are often pinpointed as predominant variables including: trip purpose (de Palma and Picard, 2005), trip distances (Scheiner, 2010) and congestion (Jan et al., 2000). Five types of trip purposes are distinguished: work, leisure, shopping, home and other. Congestion is coded as a dummy taking value one for trips made during congested periods (6:00-9:00 and 16:00-19:00) and taking value zero during other periods of the day.

Besides trip-related attributes, other factors such as socio-demographic and geographical characteristics play an important role in the route selection, as discussed by de Palma and Picard (2005), Bayarma et al. (2007) and Li et al. (2005). Therefore, the personal characteristics age, gender, net personal income, profession and the geographical characteristic province are considered in the analyses.

METHODOLOGY

Recall that the focus of this study is to assess the relationship between the choice for the shortest path or not and the corresponding activity-travel behaviour. To estimate this relationship a logistic regression is developed (Agresti, 2002). In the previous section, an elaborate description of the considered variables is provided. To assess the significance of the various trip-related and non-trip related predictors, a type III analysis of the effects is made, displayed in Table 1.

There is no significant influence of the activity-based segmentation: there is no significant relationship between whether or not the shortest path is chosen and the activities people perform. Furthermore, in line with Parkany et al (2006), congestion has no significant impact on the modelled route choice decisions. In accordance with literature (see e.g. Abdel-Aty and Huang (2004) and Parkany et al (2006)), next to trip-related attributes (trip distance), also socio-demographic variables and geographical differences play an important role.

TABLE 1 Type III Analysis of Effects

Effect	DF	Chi-Square	P-value
Road type	2	0.317	0.853
Purpose	4	2.526	0.640
Distance	1	164.372	<0,001
Congested	1	0.159	0.690
Age	3	9.734	0.021
Sex	1	1.516	0.218
Profession	5	9.849	0.080
Net personal income	5	19.690	0.001
Province	4	17.912	0.001

The parameter estimates of the logistic regression model, presented in Table 2, provide more insight in the factors that explain whether or not the shortest path is taken. To detect potential multicollinearity problems the Variance Inflation Factors (VIFs) are calculated. In general, VIF values exceeding 10 indicate the presence of serious multicollinearity undermining the validity of the results (Marquardt, 1980). Other authors consider this boundary too liberal and suggest that the variance inflation factors should not exceed 4 (Montgomory and Runger, 2003). The VIFs for the model presented in this paper indicate that there is no problem of multicollinearity.

With regard to the trip distance, the parameter estimates indicate that the longer the trip distance is, the less likely one takes the shortest path for this trip. When the trip distance would increase by 1 km, the odds of travelling by the shortest path decreases by 14 % (the odds are multiplied by 0.86 (=exp(-0.150)).

Concerning the effect of the socio-demographic variables, one can observe the clear difference between 65+ and the remaining age categories. When analyzing the results with respect to profession, it is clear that the choice for the shortest path or not is influenced by the profession of the respondent. Regarding the net personal income one could note that income has a decreasing effect on the likelihood of choosing the shortest path. The odds of choosing the shortest path are 57.2% (the odds are multiplied by 0.428 (=exp(-1.005-(-0.157))) lower for the lowest income class when compared to the highest income class. Finally, the parameter estimates also show that interprovincial differences exist. The likelihood of choosing the shortest path is smallest for Limburg and largest for Flemish Brabant.

TABLE 2 Maximum Likelihood Parameter Estimates

Parameter	Est.	St. Er.	VIF
Intercept	1.850	0.302	
Road type			
- Primary (through) road	-0.084	0.154	1.403
- Secondary (distributor) road	-0.060	0.205	1.234
- Local (access) road	0.000	n.a.	n.a.
Purpose			
- Home	0.019	0.200	1.948
- Work	0.000	n.a.	n.a.
- Leisure	-0.139	0.220	1.702
- Shopping	0.214	0.234	1.655
- Other	0.055	0.221	1.724
Distance	-0.150	0.012	1.290
Congested			
- During peak	-0.052	0.130	1.057
- Off-peak	0.000	n.a.	n.a.
Age			
- 18-25	-0.311	0.367	1.880
- 26-40	-0.167	0.162	1.338
- 41-64	0.000	n.a.	n.a.
- 65+	0.998	0.347	1.271
Sex			
- Female	0.194	0.157	1.315
- Male	0.000	n.a.	n.a.
Profession			
- Blue-collar worker	0.325	0.279	1.255
- White-collar worker	0.000	n.a.	n.a.
- Independent	-0.276	0.319	1.120
- Student	0.218	0.482	1.618
- Not professionally active	-0.484	0.204	1.814
- Other	-0.214	0.290	1.217
Net personal income			
- 0-1250 Euro	-1.005	0.246	1.716
- 1250-1750 Euro	0.000	n.a.	n.a.
- 1750-2250 Euro	-0.296	0.189	1.737
- 2250-2750 Euro	-0.082	0.270	1.528
- More than 2750 Euro	-0.157	0.341	1.275
- No answer	-0.601	0.224	1.744
Province			
- Antwerp	0.000	n.a.	n.a.
- East Flanders	-0.176	0.203	1.513
- West Flanders	0.107	0.273	1.270
- Flemish Brabant	0.157	0.191	1.741
- Limburg	-0.608	0.195	1.655

In addition, the relation between the deviation from the shortest path and the corresponding activity-travel behaviour is studied. To estimate this relation a linear regression is developed (Neter et al., 1996). To assess the significance of the various trip-related and non-trip related predictors, a type III analysis of the effects is made, displayed in Table 3.

Important to underline is the importance of the activity-based segmentation: there is a clear relationship between the deviation of the shortest path and the activities people perform. Next to trip-related attributes (trip distance), also socio-demographic variables and geographical differences play a noticeable role.

TABLE 3 Type III Analysis of Effects

Effect	DF	Chi-Square	P-value
Road type	2	1.314	0.519
Purpose	4	10.493	0.033
Distance	1	943.490	<0,001
Congested	1	0.426	0.514
Age	1	4.816	0.028
Sex	1	0.586	0.444
Profession	5	19.945	0.001
Net personal income	5	6.180	0.289
Province	4	10.219	0.037

The parameter estimates of the logistic regression model, presented in Table 4, provide more insight in the factors that explain the deviation of the shortest path. Again, the Variance Inflation Factors (VIFs) were calculated to detect potential multicollinearity problems. The VIFs calculated for the model presented in this paper indicate that there was no problem of multicollinearity.

With regard to the trip distance, the parameter estimates indicate that the longer the trip distance is, the higher the deviation from the shortest path. When the trip distance would increase by 1 km, the deviation from the shortest path will increase by 112 metres. Concerning the effect of the socio-demographic variables, one can observe that if age increases, the deviation from the shortest path decreases. When analyzing the results with respect to profession, a clear difference between blue-collar workers and the remaining categories is observed. Finally, the parameter estimates also show that interprovincial differences exist. The deviation from the shortest path is smallest for Limburg and largest for Eastern Flanders.

DISCUSSION AND CONCLUSIONS

In this study the relations between route choice decisions (i.e. shortest path or not), activity patterns and other influencing variables have been assessed. When analyzing the relationship between whether or not the shortest path is chosen and the activities people perform, there is no significant influence of the activity-based segmentation. However, when the deviation from the shortest path is related to the activities people perform, a significant relationship is found. Furthermore, in accordance with international literature, next to trip-related attributes (trip distance), also socio-demographic variables and geographical differences play a noticeable role.

In future research, instead of studying the shortest path decision, it might be interesting to study the relationship between the fastest route and the activities of the people. Another potential pathway for further investigating route choice decisions might lie in the roots of more psychological underpinnings. Besides, factors describing the situational context such as weather conditions (Cools et al., 2010a, 2010b) or public holidays (Cools et al. 2007, 2009a, 2010) can also be taken into account. Moreover, future research should extent to other transport modes such as walking, bicycle use, public transport and carpooling.

TABLE 4 Maximum Likelihood Parameter Estimates

Parameter	Est.	St. Er.	VIF
Intercept	0.019	0.331	
Roadtype			
- Primary (through) road	0.146	0.130	1.401
- Secondary (distributor) road	0.105	0.179	1.232
- Local (access) road	0.000	n.a.	n.a.
Purpose			
- Home	0.199	0.164	1.949
- Work	0.000	n.a.	n.a.
- Leisure	0.573	0.182	1.700
- Shopping	0.259	0.196	1.656
- Other	0.198	0.185	1.720
Distance	0.112	0.003	1.285
Congested			
- During peak	0.073	0.111	1.056
- Off-peak	0.000	n.a.	n.a.
Age	-0.014	0.007	2.010
Sex			
- Female	-0.097	0.127	1.250
- Male	0.000	n.a.	n.a.
Profession			
- Blue-collar worker	-0.340	0.227	1.211
- White-collar worker	0.000	n.a.	n.a.
- Independent	0.623	0.265	1.119
- Student	0.318	0.365	1.353
- Not professionally active	0.418	0.194	2.174
- Other	-0.386	0.245	1.239
Net personal income			
- 0-1250 Euro	-0.029	0.212	1.690
- 1250-1750 Euro	9.000	n.a.	n.a.
- 1750-2250 Euro	-0.101	0.151	1.581
- 2250-2750 Euro	-0.229	0.223	1.448
- More than 2750 Euro	-0.419	0.285	1.230
- No answer	0.177	0.179	1.635
Province			
- Antwerp	0.000	n.a.	n.a.
- East Flanders	0.382	0.172	1.511
- West Flanders	-0.051	0.229	1.267
- Flemish Brabant	0.034	0.161	1.664
- Limburg	-0.198	0.161	1.590

REFERENCES

Abdel-Aty, M. A., and Y. Huang (2004). Exploratory Spatial Analysis of Expressway Ramps and Its Effects on Route Choice. *Journal of Transportation Engineering*, Vol. 130, No. 1, pp. 104-112.

Agresti, A. (2002). *Categorical Data Analysis*. Second ed. Hoboken, NJ: Wiley.

Bayarma, A., Kitamura, R., and Y. Susilo (2007). Recurrence of Daily Travel Patterns: Stochastic Process Approach to Multiday Travel Behavior. *Transportation Research Record: Journal of the Transportation Research Board*, *2021*(-1), 55-63. doi: 10.3141/2021-07

Bellemans, T., Kochan, B., Janssens, D., Wets, G. and H. Timmermans (2008). Field Evaluation of Personal Digital Assistant Enabled by Global Positioning System: Impact on Quality of Activity and Diary Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2049, Transportation Research Board of the National Academies, Washington, D.C., pp. 136-143.

Cools, M., Moons, E., Bellemans, T., Janssens, D. and G. Wets (2009). Surveying activity-travel behavior in Flanders: Assessing the impact of the survey design. *Proceedings of the BIVEC-GIBET Transport Research Day*, VUBPress, Brussels, pp. 727-741.

Cools, M., Moons, E., Creemers, L. and G. Wets (2010a) Changes in Travel Behavior in Response to Weather Conditions: Do Type of Weather and Trip Purpose Matter? *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2157, pp. 22-28.

Cools, M., Moons, E. and G. Wets. (2007) Investigating Effect of Holidays on Daily Traffic Counts: Time Series Approach. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2019, pp. 22-31.

Cools, M., Moons, E. and G. Wets. (2009a) Investigating the Variability in Daily Traffic Counts 16 through Use of ARIMAX and SARIMAX Models: Assessing the Effect of Holidays on Two Site Locations. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2136, 2009, pp. 57-66.

Cools, M., Moons, E. and G. Wets. (2010) Assessing the Impact of Public Holidays on Travel Time Expenditure: Differentiation by Trip Motive. *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2157, pp. 29-37.

Cools, M., Moons, E. and G. Wets (2010b) Assessing the Impact of Weather on Traffic Intensity. *Weather, Climate, and Society*, Vol. 2, No. 1, 2010, pp. 60-68.

Davidson, W., Donnelly, R., Vovsha, P., Freedman, J., Ruegg, S., Hicks, J., Castiglione, J. and R. Picado (2007). Synthesis of first practices and operational research approaches in activity-based travel demand modeling. *Transportation Research Part A*, vol. 41, pp. 454-488.

de Palma, A. and N. Picard (2005). Route choice decision under travel time uncertainty. *Transportation Research Part A: Policy and Practice*, Vol. 39, No. 4, pp. 295-324.

Dowling, R.G. and S.B. Colman (1995). Effects of increased highway capacity: Results of household travel behaviour survey. *Transportation Research Record* 1493, pp. 143-150.

Ettema, D.F., Timmermans, H.J.P. and L. van Veghel (1997). Effect of data collection methods in travel and activity research. European Institute for Retailing and Services Studies, Eindhoven University of Technology, Eindhoven, The Netherlands.

Goldenbeld, C., Drolenga, J., and A. Smits (2007). *Routekeuze van automobilisten. Resultaten van een vragenlijstonderzoek* (No. R-2006-33) (p. 116).

Jan, O., Horowitz, A. and Z.-R. Peng (2000). Using Global Positioning System Data to Understand Variations in Path Choice. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1725, Transportation Research Board of the National Academies, Washington, D.C., pp. 37-44.

Li, H., Guensler, R. and J. Ogle (2005). Analysis of Morning Commute Route Choice Patterns Using Global Positioning System-Based Vehicle Activity Data. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1926, Transportation Research Board of the National Academies, Washington, D.C., pp. 162-170.

Marquardt, D. W. (1980). You should standardize the predictor variables in your regression models. *Journal of the American Statistical Association*, Vol. 75, No. 369, pp. 74–103.

Montgomery, D. C., and G. C. Runger (2003). *Applied Statistics and Probability for Engineers*, Fourth Ed. John Wiley and Sons, New York.

Neter, J., Kutner, M.H., Wasserman, W. and C.J. Nachtsheim (1996). *Applied Linear Statistical Models*. Fourth edn. Burr Ridge, IL: McGraw-Hill/Irwin.

Papinski, D., Scott, D.M., and S.T. Doherty (2009). Exploring the route choice decision-making process: A comparison of planned and observed routes obtained using person-based GPS. *Transportation Research Part F: Traffic Psychology and Behaviour*, *12*(4), 347-358.

Parkany, E., Du, J., Aultman-Hall, L. and R. Gallagher (2006). Modeling Stated and Revealed Route Choice: Consideration of Consistency, Diversion, and Attitudinal Variables. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1985, Transportation Research Board of the National Academies, Washington, D.C., pp. 29-39.

Ramming, M.S. (2002). *Network Knowledge and Route Choice* (Submitted to the Department of Civil and Environmental Engineering in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Transportation). Massachusetts Institute of Technology. Retrieved November 12, 2009, from http://web.mit.edu/its/papers/ramming_phd_final.pdf

Scheiner, J. (2009). Social inequalities in travel behaviour: trip distances in the context of residential self-selection and lifestyles. *Journal of Transport Geography*, *In Press.*

Schuessler, N., Balmer, M. and K.W. Axhausen (2010). Route Choice Sets for Very High-Resolution Data. *Proceedings of the 89th Annual Meeting of the Transportation Research Board*. CD-ROM. Transportation Research Board of the National Academies, Washington, D.C.

Sun, A. Sööt, S., Yang, L. and E. Christopher (1995). Household travel survey nonresponse estimates: The Chicago experience. *Transportation Research Record* 1493, pp. 170-178.

Weijermars, W., Gitelman, V., Papadimitriou, E. and C. Lima de Azevedo (2008). Safety performance indicators for the road network.

Wolf, J., Hallmark, S., Oliveira, M., Guensler, R. and W. Sarasua (1999). Accuracy Issues with Route Choice Data Collection by Using Global Positioning System. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1660, Transportation Research Board of the National Academies, Washington, D.C., pp. 66-74.

Zhang, L. and D. Levinson (2008). Determinants of Route Choice and Value of Traveler Information: A Field Experiment. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 086, pp. 81-92.

BIOGRAPHY

Katrien Ramaekers graduated as Master of Business Economics at the Limburg University Centre in 2002. In 2007, she obtained her Ph.D. in Applied Economic Sciences at Hasselt University. Currently, she is a post-doctoral researcher at Hasselt University and is working on the modelling of freight transport. She is a member of the Transportation Research Institute (IMOB) of Hasselt University.

Mario Cools obtained the degree of master in applied economic sciences at Antwerp University in 2004. In 2005 he graduated at Hasselt University as a master of science in applied statistics. Since August 2005 he has been working as a PhD candidate at the Transportation Research Institute of Hasselt University. In November 2009, he obtained his Ph.D. and now he is working as a post-doctoral researcher at the Transportation Research Institute (IMOB) of Hasselt University.

Sofie Reumers graduated as a master in transportation sciences at the Transportation Research Institute (IMOB) of Hasselt University in 2010.

Geert Wets received a degree as commercial engineer in business informatics from the Catholic University of Leuven in 1991 and a PhD from Eindhoven University of Technology in 1998. Currently, he is a full professor at Hasselt University where he is director of the Transportation Research Institute (IMOB).