What drives active transportation choices among the aging population? Comparing a Bayesian belief network and mixed logit modeling approach

What drives active transportation choices among the aging population? Comparing a Bayesian belief network and mixed logit modeling approach

Astrid Kemperman & Harry Timmermans
Eindhoven University of Technology
Urban Planning Group
P.O. Box 513
Eindhoven, 5600 MB
The Netherlands
+31 (0)40 247 3291
a.d.a.m.kemperman@tue.nl

Abstract

As people age, they typically face declining levels of physical ability and mobility. However, walking and bicycling can remain relatively easy ways to be physically active for older adults provided that the built environment facilitates these activities. The aim of this study is to investigate which variables are most effective for urban planners and management to promote the participation in active travel behavior by the older population. To do so we investigate the participation of the aging population in walking and bicycling activities as a function of socio-demographics and physical and social environmental characteristics. Revealed individual travel choice data, including all walking and bicycling trips from a random sample of 4,396 persons in the age category of 60 years and over in the Netherlands were investigated. For each trip, a large number of explanatory variables was available including information on mode choice, purpose of the trip, travel distance, travel duration, weekday, and number of trips per day. In addition, socio-demographics such as gender, age, income, whether the person has a partner, owns a bike, has a driver's license, and car possession were included. The data was fused with social and physical environmental characteristics at the neighborhood level including, urban density level, accessibility of shops, green/recreation areas, and restaurant/cafes, and indicators describing the safety and social cohesion. In this paper we therefore analyze the rich available data set using alternative approaches, a Bayesian belief network and a Mixed logit model. By comparing and integrating the outcomes of these analyses we can support more informed decisions about variable and model selection as well as provide guidance for specific urban planning and management interventions to promote active transportation choices by the elderly.

Introduction

The 65-plus age group is the fastest growing segment worldwide. As people age, they typically face declining levels of physical ability and mobility (e.g., Payne, Mowen & Orsega-Smith, 2002). Today, more than 60% of older adults are inactive (Centers for Disease Control and Prevention, 2008). However, walking and bicycling can remain relatively easy ways to be physically active for older adults, and can be done for multiple purposes (e.g., Handy, Boarnet, Ewing & Killingsworth, 2002; Joseph & Zimring, 2007).

The idea is that our built environment has become inhospitable to physical activity. Towns in some countries are built without sidewalks; suburbs are developed with stores reachable only by car; cities lack parks and recreation areas. A large proportion of people in industrialized countries live in the sprawling and exclusively residential environments that were designed for driving rather than walking, leading people to drive more and walk less, thereby contributing to declining physical activity. Specifically, for the aging population that is in general less mobile and has more limited activity spaces accessible facilities in the direct living environment are important to support their needs and improve quality of life.

The aim of this study is to investigate which variables are most effective for urban planners and management to promote participation in active travel behavior by the older population. We have adopted the human ecological theory that assumes that various forms of human behavior are strongly correlated with various factors at different spatial levels (e.g., Giles-Corti & Donovan, 2002). Ecological models thus suggest that the participation of the aging population in walking and bicycling activities is a function of socio-demographics and physical and social environmental characteristics.

We investigated revealed individual travel choice data including all walking and bicycling trips for one day from a random sample of 4396 respondents in the age category of 65 years and over in the Netherlands. For each trip, a large number of explanatory variables was available including information on mode choice, purpose of the trip, travel distance, travel duration, weekday, and number of trips per day. In addition, socio-demographics such as gender, age, income, whether the person has a partner, owns a bike, has a driver's license, and car possession were included. The data was fused with social and physical environmental characteristics at the neighborhood level including, urban density level, accessibility of shops, green/recreation areas, and restaurant/cafes, and indicators describing the safety and social cohesion.

Mixed logit models (ML) (e.g., Louviere, Hensher & Swait, 2009) have proven to be a useful tool for predicting active transportation choices and assessing policy measures and planning interventions. However, including such a large number of attributes and covariates in the model and finding meaningful interactions with the mode alternatives is a challenging task. Because variables are often highly correlated and the structure of their relationships is typically not clear (e.g., mediating effects, interaction effects, etc.) model variable selection and defining an appropriate structure for explanatory variables typically is difficult.

A Bayesian belief network (BBN) approach can overcome such difficulties by deriving and representing all direct and indirect relations between variables by using a network learning algorithm. The network learning involves two main tasks: first learning the structure of the network and then finding the parameters (Conditional Probability tables) for that structure. However, although BBN is useful in discovering the appropriate data structure among a set of variables it is less well suited to predict outcomes of specific dependent variables.

In this paper we therefore analyze the rich available data set using both alternative approaches. By comparing and using the outcomes of these analyses we can support more informed decisions about variable and model selection as well as provide guidance for specific urban planning and management interventions to promote active transportation choices by the elderly.

Data

For this study several types of data were used to measure the various factors that are assumed to influence participation in walking and bicycling activities by the aging population.

Travel Behavior and Socio Demographics

Firstly, individual travel diary data (OVin2010-data, Mobility Research Netherlands) collected by the Ministry of Transport, Public Works, and Water Management in 2009 was used. The aim of the MON data is to describe and predict daily mobility of inhabitants of the Netherlands. A random sample of 78,694 persons in the Netherlands received a letter explaining the goal of the research, a household questionnaire, and a travel diary for a designated day, and a return envelope. In total 43,191 respondents returned the questionnaires and diaries and provided useful data representing a response rate of 55%. For this study, all respondents, a total of 4,396, in the age category of 60 years and over that made at least one trip during the travel diary day were selected. The travel diary included all trips made by the respondent during one day and the number of travel diaries filled out by the respondents was equally divided across all days of the year. For each trip, a large number of explanatory variables was available including information on mode choice, purpose of the trip, travel distance, travel duration, weekday, and number of trips per day.

In addition to the travel diary, individual characteristics were asked in the survey. For this study the following information was included: gender, age, income, whether the person has a partner, owns a bike, has a driver's license, and car possession.

Physical and Social Environment

Secondly, the travel behavior data was fused with social and physical environmental characteristics at the neighborhood level. For each neighborhood, the postal code (with four digits) served as the variable to fuse the various types of data. On average 4,070 residents from on average 1,765 households live in a four digit postal code area. If there were more postal codes in a neighborhood, the postal code with the largest number of addresses was taken.

Data describing physical characteristics at the neighborhood level were derived from Statistics Netherlands (CBS). Specifically, for every neighborhood in the Netherlands, the mean distance in kilometers to specific facilities were calculated:

- Accessibility of green and recreational facilities
- Accessibility of shopping facilities and services
- Accessibility of restaurants and cafes

The degree of urbanization was measured based on the 'surrounding address density', which was the average number of addresses per 500 meter square within a kilometer radius from the address. This indicator has been widely used in the Netherlands (e.g., Maas, Verheij, Groenewegen, Vries & Spreeuwenberg, 2006) and consists of five categories:

- Very strongly urbanized (surrounding address density of 2500 and over per km^2)
- Strongly urbanized (surrounding address density between 1500 and 2500 per km^2)
- Moderately urbanized (surrounding address density between 1000 and 1500 per km^2)
- Little urbanized (surrounding address density between 500 and 1000 per km^2)
- Not urbanized (surrounding address density below 500 per km^2)

Finally, to measure the social environment indicators about safety and social cohesion in a neighborhood were defined based on a survey from a national representative sample (N =

78,071) on housing preferences collected by the Dutch Ministry of Housing, Spatial Planning and the Environment in 2009. These indicators are used by the Ministry for their Major City Policy to develop strong, comprehensive, and safe cities and towns. The indicators were derived from responses to statements about various aspects in the neighborhood using 0 to 10 scales.

Specifically the following indicators were used:

- Safety (noise pollution, trouble caused by younger persons, and trouble caused by people living in the neighborhood). This indicator can score from 0 to 10. The higher the score, the higher the safety.
- Social cohesion (solidarity between residents of the neighborhood, contacts between residents in the neighborhood, and feeling at home in the neighborhood). Also for this indicator the highest score (between 0 to 10) is the most positive.

Table 1 presents an overview of all variables and their levels included in the analyses.

Methodology

Bayesian Belief Network

A Bayesian belief network (BBN) was used to formulate and estimate the relations between the variables that directly and indirectly influence participation in active travel behavior by the aging population. A BBN is composed of a set of variables, connected by links to indicate dependencies, and containing information about the relationships between the variables. BBN is a technique for reasoning under uncertainty and emerged from combined work of Artificial Intelligence, Statistics, Operations Research, and Decision Analysis (Arentze & Timmermans, 2009).

Formally, a BBN is a directed acyclic graph (DAG) and can be written as follows (Arentze & Timmermans, 2009; Heckerman, Mandani & Wellman, 1995; Pearl, 1988):

BBN = (V, E)

where, V is a set of variables X, Y, …
and E is a set of links (X,Y).

When there is a link between X → Y, X is called a parent of Y, and Y a child of X (of course, Y could be the parent of another variable). For each variable a conditional probability table (CP table) is provided, which quantifies how much a variable depends on its parents (if any). The CP tables are referred to as the parameters of the network.

A BBN is estimated from a database by using a network learning algorithm. A BBN deals with discrete variables and therefore continuous variables were categorized. All the variables used in the BBN and their categories are presented in Table 1. The network learning involves two main tasks: first learning the structure of the network and then finding the parameters (CP tables) for that structure. The purpose of the learning algorithm is to identify the connections between the variables. The BBN-learning is based on the three-phase dependency method (Cheng, Bell, & Liu, 2002): (a) drafting the network, (b) thickening the network, and (c) thinning the network.

In the first phase, the mutual information of each pair of variables as a measure of closeness is computed and based on this information a draft network is created. The mutual information between two variables X and Y is defined as:

$$I(X,Y) = \sum_{x,y} P(x,y) \log \frac{P(x,y)}{P(x)P(y)}$$

where P(x) and P(y) are unconditional probabilities of X=x and Y=y and P(x,y) is the joint probability. The mutual information between variables X and Y measures the expected information gained about Y, after observing the value of variable X. In a BNN this means that if two variables are dependent and we know the value of one variable this will give information about the value of the other variable.

In the second phase, connections are added based on tests of conditional independence between pairs of variables (Pearl, 1988). This means that two variables are conditionally independent if their mutual information can be fully explained by indirect relations between the variables in the draft network. Only if mutual information is left the variables are connected. In the third phase, each relation is re-examined and removed if the variables of the draft network, due to implemented changes in the network, now seem conditionally independent. By setting a threshold the number of links in the network can be controlled: a lower threshold results in more links and a higher one in less links (Keuleers, Wets, Arentze & Timmermans, 2001).

For the second task in constructing a BBN, CP tables are estimated based on the same data set using the commonly used expectation-maximization (EM) learning algorithm (Lauritzen, 1995). This algorithm tries to find the conditional probability distributions and works by an iterative process, it starts with a candidate BBN, and uses it to find a better one by doing an expectation (E) step, and this is followed by a maximization step (M). This process is repeated until the log likelihood numbers are no longer improving (according to a tolerance that is specified).

PowerConstructor (Cheng et al., 2002) a software tool that is freely available on the internet, was used to learn the network structure and estimate the CP tables. The resulting network was visualized and compiled using Netica (Norsys Software Corp., 2006).

The researcher can define constraints on the presence of links between variables and pre-define special cases for the network structure a priori based on domain knowledge. For our BBN only the socio-demographic variables, age and gender were defined as parent variables because they, by nature, cannot be influenced by any of the other variables.

After constructing a BBN, it may be applied to a particular case. For example, the effect of changes in the characteristics of the physical environment, such as a higher accessibility of shops in the neighborhood on the participation in active travel behavior by the aging population can be predicted. Thus, for each or some variables values can be entered as a finding. Subsequently, probabilistic changes in other variables can be predicted and changes under certain conditions can be simulated. Every time new findings are entered into the network the CP tables of all variables can be updated based on probabilistic reasoning methods.

Mixed Logit Model

A Mixed logit model (ML) (e.g., Louviere, Hensher & Swait, 2000; Train, 2003) is used to predict active travel mode choice by the aging population as a function of various socio-demographic, environmental and trip related variables. The formal model can be described as follows. An individual *i* chooses among *J* possible transport modes on choice occasion *t*. The utility that individual *i* would obtain from transport mode *j* on choice occasion *t* is:

$$U_{ijt} = \beta_i' X_{ijt} + \varepsilon_{ijt}$$

where,

X_{ijt} is a vector of observed variables

β_i is a vector of coefficients that is unobserved for each *i* and varies randomly over the individuals representing each individuals' preferences, and

ε_{ijt} is an unobserved random term that is distributed iid extreme value, independent of β_i and X_{ijt}.

The coefficients β_i are allowed to vary across individuals, and this variance induces correlation in utility over transport mode alternatives and choice occasions. Each random parameter β_i follows a normal distribution with an average preference in the population, *b*, and an individual deviation, η_i which represents the individuals' preference relative to the average preference for a particular transport alternative. The utility is:

$$U_{ijt} = b' X_{ijt} + \eta_i' X_{ijt} + \varepsilon_{ijt}$$

The unobserved part of the utility is $\eta_i' X_{ijt} + \varepsilon_{ijt}$, and this term is correlated over mode alternatives and choice occasions. Therefore the mixed logit model does not exhibit the independence from irrelevant alternatives property of standard logit. Hence, general substitution patterns between the various transport modes can be obtained. The explanatory variables are included in the model as interaction terms. In this way, it can be seen whether the mean preferences vary across these variables (Greene, 2007).

The data for estimation were prepared as follows. The dependent variable was the transport mode choice. For each respondent six transport alternatives (car driver, car passenger, public transport, biking, walking, and other) were included in the choice set. For each out-of-home trip made during one day (one day travel diary) a choice set was included in the data estimation file. The number of trips, and thus transport modes chosen per respondent during one day varied from 1 to 15, on average 3.9 trip per respondent. In the estimation data set, dummy variables (1, 0) were used to represent the transport alternatives, and one served as the base (other). The explanatory variables (see Table 1) were all effect-coded (1, -1) except the continuous variables. Simulated maximum likelihood estimation, using 500 Halton draws (Bhat, 2001), was used to estimate the parameters of the choice model.

Results

The aim of this study is twofold, (i) to investigate which variables are most effective for urban planners and management to promote the participation in active travel behavior by the elderly and (ii) to compare and use the outcomes of these Bayesian Belief Network and Mixed Logit model analyses to support more informed decisions about variable and model selection. Therefore, first the profile of the respondents is described followed by the results and comparison of the models estimated.

Profile of Respondents

The profile of the respondents is presented in Table 2. The results show that slightly more men than women participated in the study. About 45% were in the age group of 65-70 years of age, 43% in the 71-80 years group, and 12% percent older than eighty years of age. About three-quarter of the respondents has a partner. The largest group of households was in the medium level of household income, 25% belonged to the high level income group, and about 35% had a low income level. Most of the respondents, 86%, owned a bike, and almost 69% owns one car, while, 12% owns two or more cars, and 77% possessed a driver's license.

Bayesian Belief Network

Table 1 presents the variables that are included in the BBN model analysis. The threshold for establishing links between the variables was set at 1.4, a bit higher than the standard norm of 1.0 (Keuleers et al, 2001), showing only most significant relations. The resulting network is presented in Figure 1. For each variable the probability distributions across the categories of that variable are shown. The arrows represent the relationships between two variables. The relations and variables of interest for this study (for reasons of clarity) are shown in a darker color or grey shade. This does, however, not mean that these relations are stronger. Furthermore, it can be noticed that all variables are included in the network model.

The BBN network shows a complex model and indicates that mode choice is directly related to travel distance, trip purpose, trip duration and distance, the number of trips made, urban density, and the socio-demographics gender, car and bike possession and whether someone has a driver's license. All the other variables are indirectly related to travel mode choice.

First, the network model shows that the degree of urbanization has a direct influence on mode choice and therefore is an explanatory variable for walking and bicycling activities of the aging population. All the other physical and social environmental characteristics do not have a direct influence on travel mode choice; however, they are directly related to the urban density level and indirectly related to the participation in active transportation modes.

The Bayesian belief network model can be used to predict the effect of one variable on one or more other variables in the network. Specific data for one variable can be entered as evidence in the network structure and subsequently the probabilities of the other variables can be updated. The results show that the probability that elderly cycle are quite stable (around 22.0%) over the various urban density levels, however slightly lower levels are found for the highest urbanized neighborhoods (18.7%) compared to less urbanized neighborhoods. In contrast, the share of aging people walking does increase with a decreasing urban density level (from 20.8% to 17.3%); elderly tend to walk more often in higher urbanized neighborhoods compared to lower urbanized neighborhoods. As said all the other social and physical environmental characteristics are related to each other and higher urbanized neighborhoods are characterized by lower social cohesion levels and a lower perceived safety and higher accessibility to all facilities and services.

The relations found between socio-demographic characteristics and participation in active travel indicate first that owning a bike is positively related to biking. Most of the young older generation (65-70 years of age) owns a bike, however this number significantly decreases when people get older. People with no car have highest walking and biking participation levels while the group with two or more cars have the lowest participation levels. Furthermore, aging people that have a driver's license more often travel by car, while people without a license use more active transportation modes. There is a large difference between the percentage of men (90.4%) and women (68.2%) owning a driver's license and thus indirectly in their participation rates in active transportation. However, there is also a direct relation between gender and transport mode use, showing similar results.

The network shows that travel behavior has a direct relationship with other trip characteristics such as travel duration, distance, and trip motive. The relations between travel time and travel distance with travel mode show similar results. Within a distance of 1 kilometer from their home or 5 minutes traveling the aging generation typically tend to walk. In between 1-2 kilometers they bike more often. For longer distances with a longer duration they tend to travel as car driver or passenger. However, also for the smaller distances there is still a significant amount of people that travel by car. Finally, the purpose of the trip in relation to the participation in active travel behavior. The elderly use by far most active travel modes (17.8 biking and 34.2 walking) for touring purposes. The respondents use their bike significantly more for shopping and for recreational motives. However, traveling by car as either a driver is the most important transport mode for shopping (42%) and for social visits (44.9%).

Mixed Logit Model

Including a large number of variables in the model and finding meaningful interactions with the mode alternatives is a challenging task because the variables are often highly correlated and there is a limitation on the number of variables that can be included for the model estimation. Therefore, first separate models were estimated including only a selection of the variables: a model including all socio-demographics, a model with the social and physical environmental characteristics, and a model including the trip characteristics. Secondly, a model is estimated including the variables that showed significant effects in the separate models. For this model the results and statistics are presented in Table 3.

The loglikelihood of the estimated model is -9369.3 (with 125 degrees of freedom) against a value of –23572.4 for the null model. Rho-square is .60 suggesting a good fit to the data. The constants and standard deviations for the transport mode alternatives, car driver, car passenger, public transport, biking and walking (other served as the base alternative) are all significant at the 1% level. Specifically, public transport is the least preferred transport mode alternative. The other modes, specifically walking, show some significant variation in preferences. The estimated unobserved covariances between the transport mode alternatives are also all significant at the 1% level, and almost all positive, suggesting that the likelihood that one transport mode alternative is chosen increases the likelihood that the other transport mode is chosen as well. Only for the combination of traveling by public transport and walking this parameter is negative, meaning that the choice of public transport as transport mode decreases that likelihood that someone will walk as well.

The non-random parameters, the effect of the independent variables on the transport mode choice are also shown in Table 3. First, the parameters estimated for the socio-demographic variables are discussed. A negative parameter is found for living without a partner and traveling as a car passenger, meaning there is a positive effect for people having a partner and traveling as a car passenger. Not surprisingly, when traveling by car as a couple one will drive and the other will travel as a car passenger. Furthermore, males significantly travel less as a car passenger, by public transport, or bike or walk than females do. The younger older generation travels more by public transport than the older, 80+ years generation. Furthermore, respondents owning a car or a bike travel more by all transport modes (only no effect for public transport), indicating that people with more transport possibilities in general travel more and are more active. This is possibly also related to income level, the aging population with a low income level has negative parameters for all transport mode alternatives.

Secondly, the trip characteristics show that for shopping and services trips respondents are more likely to use their car or bike than they tend to walk. On the other hand touring trips are typically walking trips. Also during the weekend respondents tend to walk more than during a weekday. The negative parameter for travel distance for walking indicates that the respondents walk less as the distance increases. The significant duration parameters show that traveling by car leads to shorter trips and walking or biking to longer trip duration.

Finally, the parameters estimated for the social and physical environmental characteristics only show significant effects for urban density levels. The aging population travels more by public transport and walks more in very strongly urbanized neighborhoods while they use both transport modes less in lower urbanized neighborhoods.

Comparing the Modeling Approaches

The results from the Bayesian Belief Network shown in Figure 1 and the Mixed Logit model results presented in Table 3 in general provide similar results, the same variables (travel distance, trip purpose, trip duration and distance, the number of trips made, urban density, gender, car and bike possession and whether someone has a driver's license) show similar effects. However, the mixed logit model suggests more variables to have a direct effect on transportation mode choice, including also age, whether someone has a partner, income, and week day. But, of course in the mixed logit model it was only tested whether variables have a direct effect. In the BBN used for this study no restrictions on the relations between the variables were set and a more complex data structure was found. This was for example clearly shown in the strong relations between all the social and physical environmental characteristics that were all presented in the urban density level.

The parameters in the BBN are the conditional probability tables which quantify how much a variable depends on its parents. By entering specific data for one variable as evidence in the network the updated probabilities for the other variables can be found. Thus the effect of one variable on other variables can be found, however, it is less well suited to clearly predict outcomes of the effect of one specific variable on one or more other variables. The parameters of the ML model provide direct quantitative information of the effect of one variable on the de pendent variable; in this study the mode choice alternatives.

To conclude, in studies were there is a large set of possible explanatory variables variable selection and defining an appropriate structure for the explanatory variables typically is difficult. Variables are often highly correlated and the structure of their relationships is typically not clear. Although a BBN is less well suited to predict outcomes of specific dependent variables, it is a useful model in discovering the appropriate data structure among a set of variables. Subsequently, the ML model is a more useful tool for predicting active transportation choices and assessing policy measures and planning interventions.

Conclusion and Discussion

In this study the relation between the participation in active travel behavior for various purposes by the aging population, individual and household socio-demographics, and objective and subjective measurements of the built environment was explored. Detailed individual travel data, including all walking and bicycling trips from a random sample of 4,396 persons in the age category of 60 years and over in the Netherlands were investigated. Two alternative approaches, a Bayesian Belief Network and Mixed logit model were estimated to investigate which variables are most effective for urban planners and management to promote the participation in active travel behavior by the older population. The outcomes of both approaches are compared and integrated. In line with the ecological model, the models estimated show a set of relations between the various factors and participation in active travel behavior.

On average the aging population in the Netherlands uses for 48% of their trips active transportation modes (walking or biking), while 47% of their trips are made by car (driver or passenger). They walk and bike specifically for touring trips, but the bike is also used for shopping and recreational activities. Specifically for shorter trips, in the neighborhood, walking and biking are popular transport modes. However, also for the smaller distances there is still a significant amount of people that travel by car. These trips are of interest when promoting participation in active transportation, because these small trips are just right for walking and cycling for most adults. Also taking into account that a large percentage of the aging population still owns a bike. Although, this percentage decreases when people age.

The results of the ML model suggested that aging people with more transport possibilities, a car and/or a bike, are also more active and make more trips. This might also be related to people's health condition. When people get older they in general are faced with health related changes and therefore they are less mobile and have more limited activity spaces. People's health or whether they have a physical handicap were not included as explanatory variables in the models. In future research it might be of interest to include these variables as well to test their effect on participation in active transportation.

The subjective and objective environmental characteristics of the neighborhood are indirectly, through the urban density level of the neighborhood, related to the participation in active travel behavior by the aging population. Higher urbanized neighborhoods are characterized by more accessible facilities and this leads to more walking trips.

Finally, when comparing the two methods, a BBN and ML model, it can be concluded that in studies were there is a large set of variables and the variables are often highly correlated and the structure of their relationships is not known a BBN is a useful method to find the appropriate structure. On the other hand, a BBN is less well suited to predict outcomes of specific dependent variables. The ML model is a more useful tool for predicting choices and assessing policy measures and planning interventions.

References

Arentze, T.A., & Timmermans, H.J.P. (2009). Regimes in social-cultural events-driven activity sequences: modeling approach and empirical application, *Transportation Research A, 43,* 311-322.

Bhat, C.R. (2001). Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research*, 35B (7), 677-695.

Centers for Disease Control and Prevention (2008). *Physical Activity for Everyone*. www.cdc.gov/nccdphp/dnpa/physical/everyone/recommendations/older_adults.htm. Accessed June 18, 2008.

Cheng, J., Bell, D., & Liu, W. (2002). Learning Bayesian networks from data: An information-theory based approach. *Artificial Intelligence. 137*, 43-90.

Giles-Corti, B., & Donovan, R. J. (2002). The relative influence of individual, social and physical environment determinants of physical activity. *Social Science & Medicine, 54* (12), 1793-812.

Greene, W. H. (2007). *NLOGIT version 4.0: Reference guide*. New York: Econometric Software, Inc.

Handy, S.L., Boarnet, M.G., Ewing, R., & Killingsworth, R.E, (2002). How the built environment affects physical activity, Views from urban planning. *American Journal of Preventive Medicine, 23* (2S), 64-73.

Heckerman, D., Mandani, A., & Wellman, M.P. (1995). Real-world applications of Bayesian networks. *Communications of ACM*, *38*, 24-26.

Joseph, A., C. Zimring. (2007), Where active older adults walk, *Environment and Behavior*, Vol. 39, No 1, 2007, pp. 75-105.

Keuleers, B., Wets, G., Arentze, T., & Timmermans, H.J.P. (2001). Association rules in identification of spatial-temporal patterns in multi-day activity diary data. *Transportation Research Record, 1752*, 32-37.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and application*. Cambridge, UK: Cambridge University Press.

Maas, J., Verheij, R.A., Groenewegen, P.P., Vries, de S., & Spreeuwenberg, P. (2006). Green space, urbanity, and health: how strong is the relation? *Journal of Epidemiological Community Health*, *60*, 587-592.

Norsys Software Corp. (2006). Netica 3.17 for MS Windows. Vancouver, Canada.

Payne, L. L., A. J Mowen, and E. Orsega-Smith (2002). An examination of park preferences and behaviors among urban residents: The role of residential location, race and age. *Leisure Sciences*, 24, 181-198.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*, San Francisco, CA: Morgan Kaufman.

Train, K.E. (2003). *Discrete Choice Methods with Simulation*. Cambridge University Press, UK.

Table 1 Variables used in the analyses

Variable	*Categories*
Socio demographics	
-age	65-70 years, 71-80 years, 80+ years
-gender	male, female
-bike owner	yes, no
-car possession	no car, 1 car, 2 or more cars
-partner	yes, no
-driver's license	yes, no
-income	low, medium-high
Travel behavior	
-mode	car driver, car passenger, public transport, biking, walking, other
-purpose	shopping, social visits, recreational, touring, other
-duration*	0-5 min, 5-10 min, 10-15 min, 15-30 min, >30 min
-distance*	0-1 km, 1-2 km, 2-4 km, 4-10 km, >10 km
-weekday	weekend, weekday
-number of trips*	1-2, 3-4, 5-8, >=9
Physical environment	
-urban density	very strongly urbanized, strongly urbanized, moderately urbanized, little urbanized, not urbanized
-accessibility of green/ recreational facilities*	very good, good, ok, bad, very bad
-accessibility of shops/services*	very good, good, ok, bad, very bad
-accessibility of restaurants/ cafes*	very good, good, ok, bad, very bad
Social environment	
-social cohesion*	very negative, negative, neutral, positive, very positive
-safety*	very negative, negative, neutral, positive, very positive

* used as a continuous variable in the mixed logit model

Table 2 Sample characteristics (*N*=4,396)

Variables	*Levels*	*%*
Gender	Male	51.1
	Female	48.9
Age	65 ≤ 70 years	45.1
	71 ≤ 80 years	43.3
	80+ years	11.5
Partner	Yes	71.1
	No	28.9
Income level	Low	35.4
	Medium	39.2
	High	25.4
Car owner	No car	18.1
	1 car	69.6
	2 or more cars	12.3
Bike owner	Yes	85.8
	No	14.1
Driver's license	Yes	76.9
	No	23.1

Table 3 Mixed logit model of active transportation choice of the aging population

	Car driver	Car passenger	Public transport	Biking	Walking
Random parameters					
Constants	4.29**	4.55**	-6.92**	2.94**	11.86**
Standard deviation	5.38**	8.74**	7.90**	7.00**	8.63**
Unobserved covariances:					
car passenger	6.57**				
public transport	4.22**	4.22**			
biking	3.77**	3.76**	2.98**		
walking	6.49**	2.31**	-2.28**	4.66**	
Non-random parameters					
Socio demographics:					
Partner, single	.61	-1.29**	-.04	.04	.02
Gender, male	-.44	-3.29**	-1.57**	-1.59**	-1.71**
Age, 65-70	.09	-.55	2.13**	.67	-.01
Age, 71-80	-.23	-.35	-.34	.04	-.28
Income, low	-.43	-1.11**	-.23	-.89*	-.89*
Income, medium	.83*	1.17**	.98	1.10**	.91*
Car, no car	-6.93**	-3.70**	.82	-1.17	-.206*
Car, 1 car	2.78**	1.89**	.38	1.15**	1.23**
Driver's license, yes	3.39**	-.84	-.01	-.02	.005
Bike owner, yes	1.44**	1.75**	.26	7.05**	2.21**
Travel behavior:					
Week day	-.61*	-1.10**	-.10	-.37	-1.19**
Motive, shopping/services	1.08**	.61*	.66	.86**	-1.34**
Motive, social visits	.56	.10	.14	-.62	-1.66**
Motive, recreational	.37	.69*	.15	.02	-1.58**
Motive, touring	-2.86**	-1.46**	-1.81	.30	6.06**
Travel distance, hectometers	.000	.003*	.002	-.036**	-.17**
Travel duration, minutes	-.018**	-.028**	.018	.029**	.088**
Physical environment:					
Urb. density, very strongly urb.	-.16	-.33	4.02**	.00	1.46*
Urban density, strongly urbanized	-.75	-1.03*	1.66*	.-.76	-.63
Urban density, moderately urb.	.62	.95	.19	.81	.50
Urban density, little urbanized	-.36	-.43	-2.85**	-.76	-1.30*
Model statistics	LL(0) -23572.4	LL(B) -9369.3	Rho square .60		

****** significant at 1% level, * significant at 5% level

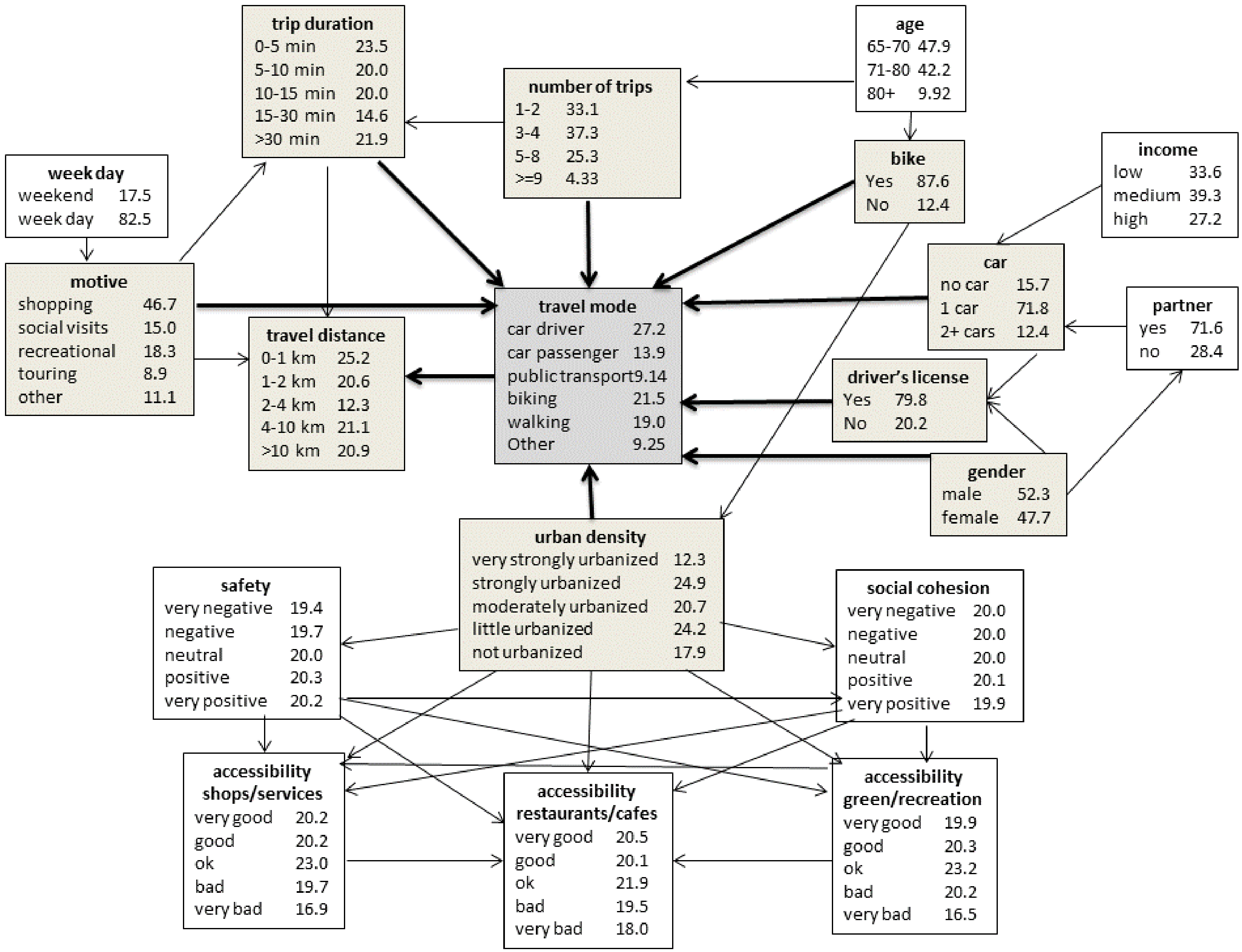

Figure 1 The Bayesian Belief network model of active travel behavior of the aging population