Incorporating social interaction into hybrid choice models

Maria Kamargianni · Moshe Ben-Akiva · Amalia Polydoropoulou

Abstract The aim of this paper is to develop a methodological framework for the incorporation of social interaction effects into choice models. The developed method provides insights for modeling the effect of social interaction on the formation of psychological factors (latent variables) and on the decision-making process. The assumption is based on the fact that the way the decision maker anticipates and processes the information regarding the behavior and the choices exhibited in her/his social environment, affects her/his attitudes and perceptions, which in turn affect her/his choices. The proposed method integrates choice models with decision makers' psychological factors and latent social interaction. The model structure is simultaneously estimated providing an improvement over sequential methods as it provides consistent and efficient estimates of the parameters. The methodology is tested within the context of a household aiming to identify the social interaction effects between teenagers and their parents regarding walking-loving behavior and then the effect of this on mode to school choice behavior. The sample consists of 9,714 participants aged from 12 to 18 years old, representing 21 % of the adolescent population of Cyprus. The findings from the case study indicate that if the teenagers anticipate that their parents are walking lovers, then this increases the probability of teenagers to be walking-lovers too and in turn to choose walking to school. Generally, the findings from the application result in: (a) improvements in the explanatory power of choice models,

M. Kamargianni (✉)
UCL Energy Institute, University College London, Central House, 14 Upper Woburn Place,
WC1H 0NN London, UK
e-mail: m.kamargianni@ucl.ac.uk

M. Kamargianni · A. Polydoropoulou
Department of Shipping, Trade and Transport, University of the Aegean, Korai 2a, 82100 Chios,
Greece
e-mail: polydor@aegean.gr

M. Ben-Akiva
Department of Civil and Environmental Engineering, Massachusetts Institute of Technology,
Cambridge, MA 02139, USA
e-mail: mba@mit.edu

(b) latent variables that are statistically significant, and (c) a real-world behavioral representation that includes the social interaction effect.

Introduction

For many years, socioeconomic factors and the attributes of transport alternatives have been the key elements considered in most models used to support stakeholder planning (Shiftan et al. 2008). However, it has been recognized that a complex interaction between several factors takes place whenever an individual makes a choice and as a consequence beliefs, values, emotions, attitudes and other personal characteristics have been incorporated into choice models (Walker and Ben-Akiva 2001; Ben-Akiva et al. 2002).

In addition, it is well-known that individuals' choice behavior is often influenced by the existence, opinions, choices and behaviors of other people (van den Bos et al. 2013; Rose and Hensher 2004; Brock and Durlauf 2001; Manski 1993) or generally by the social environment of the decision maker. In sociology and psychology, there is much empirical evidence confirming the effect of social interaction. In terms of neighborhood influences, Crane (1991) found a relationship between both school dropout and teenage childbearing rates and the occupational composition of a community. Haveman and Wolfe (1994) presented similar findings regarding high school dropout rates. In a different context, Durlauf and Walker (1998) argued that social interaction plays a major role in explaining variations in fertility rates and the adoption of different birth control technologies.

In recent years, the effect of social interaction and social influence on individuals' decision-making has attracted attention in the transportation sector as well. In this context, Paez and Scott (2007) and Wilton et al. (2011) found that the decision of an individual to telecommute is heavily influenced by others deciding to telecommute. Social interaction effects have also been recognized in other settings, such as modal choice decisions (Dugundji and Walker 1921; Goetzke 2008), leisure travel (Axhausen 2005), participation in social activities (Carrasco and Miller 2006), and even illicit parking behavior (Fukuda and Morichi 2007). A guiding philosophy in these surveys is that the incorporation of social interaction variables leads to a more behaviorally realistic representation of the choice process, and consequently to a better explanatory power.

Arguably, the utility of an individual's choice is a function of socioeconomic characteristics and psychological factors (Ben-Akiva et al. 2002). The psychological factors are affected by the choices and behavior exhibited in the social environment of the individual, and also by the way that the individual processes or anticipates this information. McFadden (1997) argued that the most cognitive anomalies in utility theory operate through errors in perception that arise from the way information is stored, retrieved and processed, and that the empirical study of economic behavior would benefit from closer attention to how attitudes and perceptions are formed and how they influence decision-making. Currently, there is still a gap between decision-making in real life, where the influence of the social environment is extensive, and decision-making as measured in the laboratory, which is often done in the absence of any social influences (Weinberg and Pehlivan 2011).

Given the above, the aim of this paper is to develop a modeling framework for the incorporation of social interaction into hybrid choice models (henceforth HCM), based on the previous work of Ben-Akiva et al. (2002a, b, 2012). More specific, the developed method provides insights for modeling the effect of social interaction on the formation of psychological factors (latent variables) and on the decision-making process. The assumption of the method is based on the fact that the way the decision maker anticipates and processes the information regarding the behavior and the choices exhibited in her/his social environment affects her/his attitudes and perceptions, which in turn affect her/his choices. The proposed method integrates choice models with decision makers' psychological factors and latent social interaction. The model structure is simultaneously estimated providing an improvement over sequential methods as it provides consistent and efficient estimates of the parameters.

A mode choice model is developed in order to test the methodology and to provide a conceptual example, sample of equations, and estimation results. The methodology is tested using data from a large-scale transport survey conducted in the Republic of Cyprus in 2012 referring to adolescents. The total sample consists of 9,713 participants aged from 12 to 18 years old, representing 21 % of the total adolescent population of the country. The assumption that is tested is that when the teenager anticipates that their parents are walking-lovers, then this affects positively its own attitudes towards walking increasing the probability of teenager to be a walking-lover and in turn to choose walking for the trip to school.

The remainder of the paper is structured as follows. "State of the art" section reviews and discusses the literature. "Methodology" section describes the proposed methodology, the modeling framework, the associated mathematical formulations and the data required to estimating this model. The application of the methodology is presented in "Model application" section. "Conclusions" section concludes the paper by providing a summary of the findings and suggestions for further research.

State of the art

Process and context of decision-making

Decision-making plays a pivotal role in daily life, comprising a complex process of assessing and weighing the short-term and long-term costs and benefits of competing alternatives. The output of the decision-making process is determined by an interaction between impulsive or emotionally based systems responding to potential rewards and losses, and reflective or cognitive control systems controlling long-term goals (Visser et al. 2011). The description of choice behavior can be given more structure by describing choice behavior as a decision making process involving two dimensions: process and context (Ben-Akiva et al. 2012). Process refers to the steps involved in decision-making, while context refers to the factors affecting the process.

Individuals recognize opportunities and constraints regarding their choices. They collect and process information about the attributes of available options which, together with their attitudes and emotional states, influences their perceptions and beliefs about these options. Decision makers then focus and refine their preferences, targets and needs and form a plan for making the decision (Ben-Akiva 2010). The plan can be thought of as a strategy, set of decision criteria or set of intentions. Different alternatives are evaluated and the decision is made by following the plan. Decision-making is influenced by many factors, such as

gender, age, genotype, and personality, which have been extensively investigated and discussed (van den Bos et al. 2013; Overman 2004; Homberg 2012). Nevertheless, relatively little attention has been paid to the crucial moderating effect of social context on decision-making.

Context refers to factors affecting the process. In real life, decisions are often strongly influenced by the person's social environment and involve direct and indirect social interactions. A valuable way to structure this is through social networks, as they affect the flow and the quality of information (Granovetter 2005). A person's social network may affect their decision-making in numerous ways. In daily life, individuals constantly make decisions based on their personal information and experience, as well as that of others. An individual's decisions may also be indirectly influenced by their social environment, through the effect the latter can have on an individual's emotional/psychological state (van den Bos et al. 2013).

Importantly, the modulating role of the social environment is strongly affected by an individual's characteristics and personality as well as those of its group mates (Webster and Ward 2011). Decision makers generally belong to a number of social networks, which may be small or large and include few or many members. At this point, a distinction between tight and loose social networks is worth making. Tight social networks have few members, strong interactions between those members, and high entry and exit costs (Christakis and Fowler 2009). Examples of tight social networks include groups defined by family relationships or close friendships. Tight social networks exhibit strategic interactions, joint constraints, and joint production. Loose social networks have low entry and exit costs. They are larger, and involve weaker interactions between members. There are many examples of loose social networks, such as friends, online networks, neighborhoods, ethnic groups, classrooms, clubs, and professional networks (e.g., close work colleagues). The size of loose social networks implies that the potential for strategic interaction is small (Golub and Jackson 2010).

Concluding, the research community is gradually starting to appreciate the importance of factors such as interactions between decision makers, the actual processes leading to choices, and the role of subjective factors. As a result, new models are emerging that give a more realistic representation of real-world behavior, such as the Brock and Durlauf's (2001) discrete choice models with social interaction, which is in essence a static Nash equilibrium model in which a random utility framework is extended to include an effects of the expected choices of others on individual payoffs. Another model is the strategy adjustment model of Blume (Abou-Zeid and Ben-Akiva 2011), in which binary choice evolve in response to the past behavior of others via a stochastic population process.

Discrete choice models

Discrete choice models (DCM) have played a significant role in transportation modeling. DCM consider demand to be the result of several decisions made by each individual under consideration, where each decision consists of a choice made among a finite set of alternatives (Ben-Akiva and Lerman 1985; Bierlaire 1998). They explain choice behavior simply as a set of preferences ranking all potential outcomes, where the consumer is assumed to choose the most preferred available outcome. Under certain assumptions, consumer preferences can be represented by a utility function such that the choice is the utility maximizing outcome. These models have traditionally presented an individual's choice process as a "black box", in which the inputs are the attributes of available alternatives and the individual's characteristics, and the output is the observed choice

(Ben-Akiva et al. 2002). Behavioral researchers have stressed the importance of the cognitive workings inside the black box in determining choice behavior (Abelson et al. 1985; Olson and Zanna 1993; Kivetz et al. 2004), and a great deal of research has been conducted to uncover cognitive anomalies that appear to violate the basic axioms of utility theory (Gärling et al. 1998; Rabin 1998).

Hybrid choice models

Over the last few decades, researchers have focused on enhancing DCM, and numerous improvements have been made that aim to predict realizations of the choice behavior. These methods are integrated in HCM (Walker and Ben-Akiva 2001). Among the numerous extensions of HCM is the explicit modeling of latent psychological factors such as attitudes and perceptions (latent variables). HCM by combining "hard information" (such as socioeconomic characteristics) with "soft information" on population heterogeneity (such as psychological characteristics), explain better decision makers' behavior and in doing so a substantial part of the unobserved heterogeneity (Ben-Akiva et al. 2002b).

Walker and Ben-Akiva (2001) presented the extended HCM framework, where they estimated mode choice models using revealed and stated preference data, latent perceptions of comfort and taste heterogeneity in the form of random parameters and latent class segmentation. The latent factors provided for a richer behavioral representation of the choice process (although not a significant improvement in the overall fit of the model), while the inclusion of taste heterogeneity improved the explanatory power of the model. Given that the HCM framework is constructed by integrating modular components such as latent variable models, flexible disturbances, etc., its development has been catalyzed by technical developments and growing practical experience with each of the modular components (Ben-Akiva et al. 2002b).

An extension to HCM has been presented by Abou-Zeid and Ben-Akiva (2011) in order to capture the indirect effect of social comparisons on travel choices through its effect on comparative happiness. They argued that social comparisons arise from exchanges of information among individuals and they postulated that the social gap resulting from comparisons is a determinant of "comparative happiness", which in turn affects subsequent behavior. They studied how perceived differences between experienced commute attributes and those communicated by others affect comparative happiness and consequently overall commute satisfaction.

In this paper, we postulate that the information that the decision makers receive from their social environments, represent the social interaction effect and this affects the attitudes and perceptions of the decision maker. Thus, we extend the HCM by incorporating the social interaction effect.

Methodology

Modeling framework

A starting point for the proposed methodology is a combination of a choice model with a latent variable model. That is, the framework of HCM that has been developed to enrich the behavioral realism of DCM by accounting for latent factors such as perceptions and attitudes and employing more flexible error structures. The framework of the HCM has

been applied in various transportation contexts, such as mode choice (Johansson et al. 2006; Polydoropoulou et al. 2013; Abou-Zeid et al. 2011), vehicle purchase (Bolduc et al. 2008) and route choice (Tsirimpa and Polydoropoulou 2007).

Having in mind the HCM in combination with the latest findings in psychology, neuroscience and biobehavioral research, which state that the individual's decisions are indirectly influenced by the social environment, as it affects the individual's psychological state (van den Bos et al. 2013; Homberg 2012), we add one more dimension in the construction of the latent variable model, that of the social environment. The choices or the behavior exhibited in the social environment are filtered by the decision maker, which in turn shapes her/his attitudes towards these choices or behaviors, or as per Anais Nin, "We do not see things as they are, we see them as we are." Thus, the social environment is one more latent variable that represents the social interaction between the decision maker and her/his social environment and it is added into the latent variable regarding the decision maker (Fig. 1).

Figure 1 presents the modeling framework. The rectangular box in the upper right corner represents the social environment of the decision maker. The social environment has its own explanatory variables (S), i.e. socioeconomic characteristics, and the choices or behaviors seen in that environment are measured as psychological indicators (IS), as perceived by the decision maker. These psychological factors (IS) are used to build a latent variable (S^*) regarding the social environment of the decision maker. This latent variable (S^*) is incorporated into the formulation of the latent variable regarding the decision maker (X^*). For the construction of X^*, psychological indicators are used that refer to the decision maker's attitudes and perceptions, while X^* is affected by the explanatory variables X as well. The utility of the choice (U) is affected by the explanatory variables X and the latent variable X^*. The latent variable (X^*) and the explanatory variables (X) directly affect the choice made by the individual, while the social environment indirectly affects the utility of that choice. y represents the choice indicator.

In the social environment box, we can include as many social networks as we want, each one representing a latent variable. For example, we could introduce a latent factor regarding family, another one regarding friends or colleagues etc., or even a latent factor for each individual member of the household.

The integrated model is used to include latent variables regarding the decision maker and her/his social environment in choice models. The methodology incorporates indicators of the latent variables, provided by the responses to survey questions, to aid in the estimation of the model. A simultaneous estimator is used, which results in latent variables that provide the best fit to both the choice and the latent variable indicators.

Specification of the model

In this sub-section it is presented a generic formulation of the model shown in Fig. 1. It is assumed that all latent variables and their indicators are continuous for simplicity. The model consists of structural and measurement equations. The structural equations express latent variables S^* (Eq. 1), X^* (Eq. 2) and utility U (Eq. 3) using the links shown in Fig. 1. Each of these variables is also a function of a random error term. U is a vector, whose dimensionality is equal to the number of alternatives considered (i).
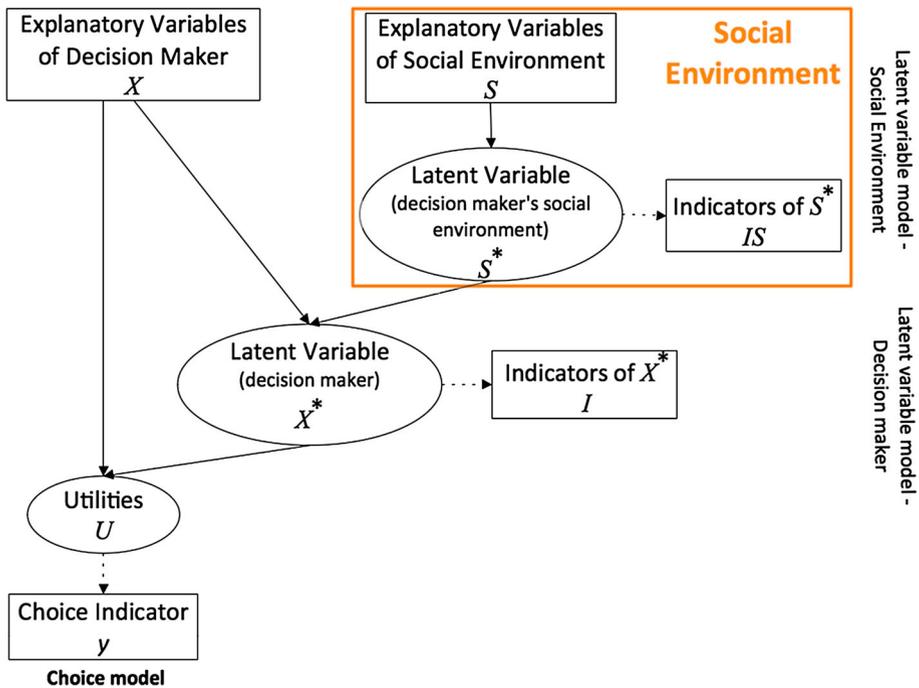
Fig. 1 Modeling framework

Structural model

For the social environment of decision maker:

$$S^* = S\zeta + \eta \quad \eta \sim N(0, \Sigma_\eta) \tag{1}$$

where S^* *is* the latent (unobservable) variable regarding the decision maker's social environment, S are matrices of explanatory observed variables regarding the social environment of the decision maker, ζ is a vector of unknown parameters used to describe the effect of observable variables (S) on the latent variables, η is vector of random disturbance terms and Σ_η designates the covariance of random disturbance terms.

For the decision maker:

$$X^* = X\vartheta + S^*\xi + \omega \quad \omega \sim N(0, \Sigma_\omega) \tag{2}$$

where X^* *is* the latent (unobservable) variable based on the decision maker's attitudes or perceptions, X are the explanatory observed variables (RP) regarding the decision maker, θ is a vector of unknown parameters used to describe the effect of observable variables (X) on the latent variable, ξ is a vector of unknown parameters used to describe the effect of the latent variable based on decision maker's social environment (S^*) on the latent variable regarding the decision maker, ω is vector of random disturbance terms, while Σ_ω designates the covariance of random disturbance terms.

Utility:

$$U = X\beta + X^*\gamma + \varepsilon \quad \varepsilon \sim N(0, \Sigma_\varepsilon) \tag{3}$$

where U is a vector of utilities, β is a vector of observed variables regarding the decision maker, γ is the unknown parameter associated with the latent variable X^*, ε is a vector of random disturbance terms associated with the utility terms, and Σ_ε denotes the covariance of the random disturbance terms.

The availability of indicators I of the latent variable regarding the decision maker and IS of the latent variable regarding the social environment eases the identification of the model and results in more efficient parameter estimates (Abou-Zeid et al. 2011). These indicators can be expressed as a function of the corresponding latent variables and a random error term, as shown in measurement Eqs. (4) and (5). Knowing the distributions of the error terms, the density functions of the indicators can be derived. A latent variable may have more than one indicator, so I and IS are vectors.

Measurement model

For the Social environment of decision maker:

$$IS = a' + \lambda'S^* + \upsilon' \quad \upsilon' \sim \mathrm{N}(0, \Sigma_{\upsilon'}) \tag{4}$$

where IS corresponds to the indicators of the latent variable that is constructed for the social environment of the decision maker (S^*), α' is a vector of parameters that indicate the associations between the responses to the scale, λ' is a vector of unknown parameters that relate the latent variable S^* to the indicators, and υ' is a vector of independent error terms with unitary variance and $\Sigma_{\upsilon'}$ designates the covariance of the random disturbance terms. For the decision maker:

$$I = a + \lambda X^* + \upsilon \quad \upsilon \sim \mathrm{N}(0, \Sigma_\upsilon) \tag{5}$$

where I corresponds to the indicators of the latent variable based on the decision maker's psychological factors (X^*), α is a vector of parameters that indicate the associations between the responses to the scale, λ is a vector of unknown parameters that relate the latent variable X^* to the indicators and υ is a vector of independent error terms, and Σ_υ designates the covariance of the random disturbance terms.
Choice model:

$$y_i = \begin{cases} 1, & if \quad U_i = \max j\{U_j\} \\ 0, & otherwise \end{cases} \tag{6}$$

y_i is a choice indicator, taking the value 1 if alternative i is chosen, and 0 otherwise.

The choice probability for a given observation is:

$$P(y_i|X, X^*; \Sigma\mu) \tag{7}$$

where $\Sigma\mu$ designates all the unknown parameters in the choice model of Eq. (3).

The likelihood function for a given observation is the joint probability of observing the choice and the attitudinal indicators as follows:

$$f(y, I, IS|X, S; \delta) =$$

$$\int_{X^*} \int_{S^*} P(y|X, X^*; \beta, \gamma, \Sigma_\varepsilon) f(I|X, X^*; \lambda, \Sigma_\upsilon)(IS|S, S^*; \lambda', \Sigma_{\upsilon'}) f(X^*|X; \vartheta, \xi, \Sigma_\omega)(S^*|S; \zeta, \Sigma_\eta) dS^* dX^*$$

$$\tag{8}$$

where δ designates the full set of parameters to be estimated ($\delta = \{\beta, \gamma, \lambda, \lambda', \theta, \zeta, \xi, \Sigma_\varepsilon, \Sigma_v, \Sigma_{v'}, \Sigma_\omega\}$). The first term of the integral corresponds to the choice model. The second term corresponds to the measurement equations from the latent variable models (both for the decision maker and the social environment) and the third term corresponds to the structural equations from the latent variable models (both for the decision maker and the social environment). The latent variable is only known to its distribution, and so the joint probability of y, I, IS, X^*, and S^* is integrated over the latent constructs X^* and S^*.

Model application

Survey design

The proposed methodology is tested by using data from a large-scale transport survey that took place in the Republic of Cyprus in 2012 and refers only to teenagers. The questionnaire was designed by transport planners in co-operation with psychologists aiming to capture the fundamentals of travel behavior and the multidimensional nature of transportation problems. The web-questionnaire was forwarded to all Cypriot high schools by the Ministry of Education of Cyprus (for more details about the questionnaire and the data collection see (Kamargianni and Polydoropoulou 2013a; Kamargianni 2014).

Among the other topics that the questionnaire covered, we had a section with questions asking participants to state their level of agreement or disagreement regarding: (1) their attitudes and perceptions towards walking and cycling, (2) their attitudes towards their parents travel behavior and walking, cycling and private vehicle use patterns, and (3) their attitudes towards their friends travel behavior and walking, cycling and private vehicle use patterns. For the purposes of this paper, we use revealed preference data regarding:

- the transport mode that teenagers used for their trip to school;
- the built environment characteristics of the route between home and school;
- the attitudes of teenagers towards walking;
- the attitudes of teenagers towards their parents walking behavior; and
- the socio-economic and demographic characteristics of teenagers and teenagers' household members.

That is, the proposed methodology is tested within the context of a household and the social influence between teenagers and their parents. Despite the prominent role that the caregiver likely plays in the travel decision for elementary school children, teenagers typically want to avoid parental supervision by making trips that are not controlled or supervised (Clifton et al. 2010). Teenagers are mature enough to make their own mode choices when traveling to school (Babey et al. 2009), thus we are able to use the utility maximization theory, as the sample consists of teenagers and not children.

Case study: sample

The sample used for this case study consists of 9,713 participants representing the 21 % of the total adolescent population of the country. The descriptive statistics of the sample are presented in Table 1. 55 % of the participants are females. 58 % are 15 to 18 years old. 16 % of the adolescents walk to school, 35 % use bus, while 49 % are escorted by their parents using a private motorized vehicle. Parental level of education is quite low

(secondary education). The household car ownership is rather high, and none of the students stated that there were no cars in their household. Also, various built environment characteristics were measured for each individual's route from home to school after the completion of the survey.

Table 2 presents the responses to the attitudinal questions regarding teenagers' willingness to walk and regarding their parents walking habits. The answers to the statements *WL1* to *WL8* serve as attitudinal indicators of the latent variable "Walking-lover" (henceforth *WL*). The answers to the statements *PWL1* to *PWL2* are used as indicators of the latent variable *"Parents: Walking-lovers"* (henceforth *PWL*). In this way, *PWL* measures how teenagers anticipate their parents' walking behavior and *WL* measures the teenagers' predisposition to walk. The response scale ranged from 1 to 7, with a response of 1 indicating that the participant completely disagreed with the statement, and 7 indicating that they completely agreed.

Model specification

In order to test the methodology, a mode choice model is developed aiming to investigate how the anticipated parental (social environment) walking behavior affects teenagers' (decision makers) attitudes towards walking, and how the latter affects teenagers' mode choice behavior. We hypothesize that when the teenager anticipates that their parents are walking-lovers, then this affects positively its own attitudes towards walking increasing the probability of the teenager to be a walking-lover and in turn to choose walking for the trip to school. The assumption is based on the fact that the payoff to the decision maker of choosing walking is a direct function of his/her attitudes towards walking and an indirect function of her/his attitudes towards the walking habits of her/his social environment (McMillan 2007; Seraj et al. 2011; Bhat et al. 2010; Yoon et al. 2011). That is the effect of social interaction or social influence, which is incorporated in the latent variable regarding the decision maker and then this is included in the choice model. Figure 2 presents the modeling framework of the case study.

Structural model

The utility of choice is a function of socio-economic characteristic, urban environment characteristics and the latent variable *WL*. The deterministic utility of alternative Walk (denoted as *WALK*—Eq. 9) contains the distance from home to school interacted by gender (as previous surveys found that the distance that teenagers walk is affected by gender (McMillan 2005), age the family income, various built environment characteristics and the latent variable *WL*, as we assume that the teenagers who are walking-lovers prefer walking for their trip to school. The utility of *BUS* (Eq. 10) includes the traveled distance from home to school. The utility of being escorted by adults by private motorized vehicles to school (denoted as *CAR*—Eq. 11) is affected by distance interacted with gender, age, the family income and the number of available motorized vehicles in the household divided by the household size. Travel time is captured by distance per mode. Travel cost variables are not used, due to the fact that: (1) students in Cyprus can use the bus without any charge, (2) teenagers do not anticipate clear the travel cost by car as their caregivers pay for it. Availability constraints were inserted in the alternative *WALK*; when distance from home to school is more than 2.1 km, this alternative is not available. Since all the participants' households own at least one motorized vehicle, the alternative *CAR* was available to all. Restrictions on parents' availability to escort their children were not imposed, as this

Table 1 Descriptive statistics of the sample

	Percentage (%)
Teenagers (decision maker)	
Gender	
Female	55
Male	45
Age	
11–14 years old	42
15–18 years old	58
Lived abroad	
Yes	14
No	86
Mode to school (dependent variable)	
Active transport	16
Bus (public transport)	35
Private vehicle (escorted by parents)	49
Parents (social environment)	
Educational level of father	
Low (secondary education)	64
Medium (university)	25
High (masters & PhD)	11
Educational level of mother	
Low	60
Medium	29
High	11
Household's characteristics	
Income (euros per month)	
Less than 2000€	25
2001€ to 4000€	27
More than 4000€	20
Not available	28
Driving license	
One of the parents holds a driving license	23
Both parents hold a driving license	77
Car ownership	2.6
Household size	4.8
Built environment characteristics for the route home-school	
Aesthetics: existence of green (trees and flowers)	38
Existence of cross walks	67
Sidewalks	55
Distance traveled on foot	Max = 2.1 km, Min = 0.002 km, Mean = 0.75 km
Distance traveled by bus	Max = 34.6 km, Min = 1.1 km, Mean = 11.3 km
Distance traveled by private motorized vehicle	Max = 24.1 km, Min = 0.02 km, Mean = 3.8 km

Table 2 Indicators of latent variables

|		Mean	SD
Indicators of walking-lover			
I_{WL1}	I am willing to walk to school, in order to be fit	4.2	2.280
I_{WL2}	I am willing to walk to school, as it is the most cost-effective mode	3.8	2.154
I_{WL3}	I am willing to walk, in order to protect the environment	4.3	2.002
I_{WL4}	I consciously make an effort to walk instead of being escorted	4.6	2.001
I_{WL5}	I prefer walking rather than being escorted for my short-distance trips	3.8	2.192
I_{WL6}	I really enjoy walking	3.9	2.183
I_{WL7}	I am not willing to walk to school, because it is time consuming	4.0	2.233
I_{WL8}	I am not willing to walk to school alone	3.7	2.178
Indicators of anticipated "parents: walking-lovers"			
SI_{PWL1}	My mother walks for her short-distance trips	4.0	1.868
SI_{PWL2}	My mother prefers walking than using the car	2.5	2.020
SI_{PWL3}	My father walks for his short-distance trips	3.2	2.039
SI_{PWL4}	My father prefers walking than using the car	2.4	2.020
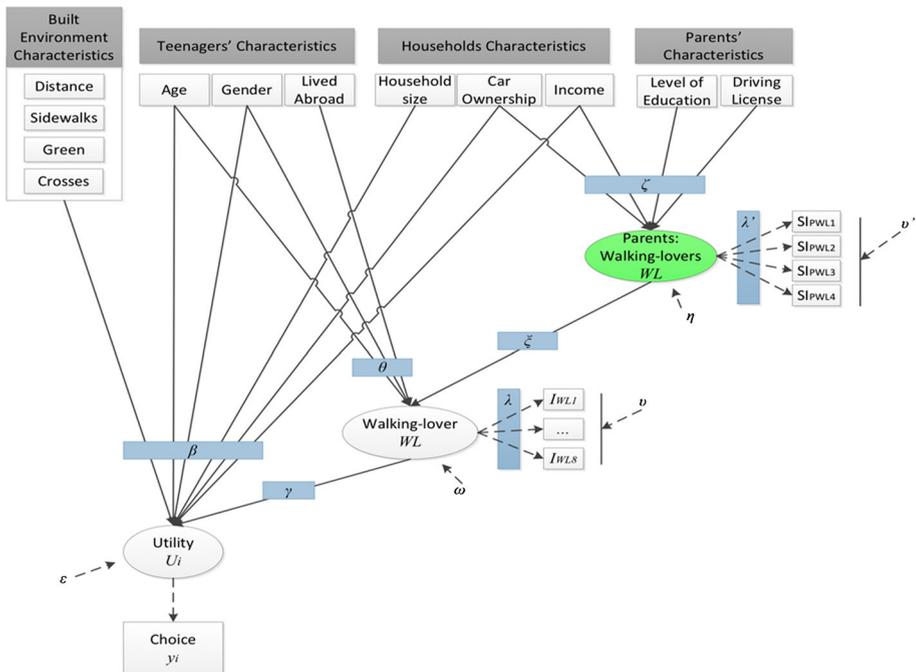
Fig. 2 Modeling framework for teenagers (decision makers) and social interaction effect of their parents (social environment)

variable was not available in the dataset. However, even if the parents were not available, the option would still be available to all the participants as someone else could escort them (e.g. the parents of a fellow student).

$$U_{WALK} = (\beta_{D1} + \beta_{G1} * FEMALE) * DIST2km + \beta_{A1} * AGE1114 + \beta_{INC1} * INCOME$$
$$+ \beta_{G1} * GREEN + \beta_{C1} * CROSS + \gamma * WL + \varepsilon_{WALK} \tag{9}$$

$$U_{BUS} = \beta_{BUS} + \beta_{D2} * DIST5km + \varepsilon_{BUS} \tag{10}$$

$$U_{CAR} = \beta_{CAR} + (\beta_{D3} + \beta_{G3} * FEMALE) * DIST25km + \beta_{A3} * AGE1114$$
$$+ \beta_{INC3} * INCOME + \beta_{CHH} * (CARHH/HHSIZE) + \varepsilon_{CAR} \tag{11}$$

where: *FEMALE* takes the value 1 if the participant is female, 0 otherwise; *AGE1114* takes the value 1 if the participant is from 11 to 14 years old, 0 otherwise; *INCOME* denotes the monthly family income in Euros, continuous variable; *DIST2 km* takes the value 1 if the traveled distance between home and school is up to 2.0 km, 0 otherwise; *DIST25 km* takes the value 1 if the traveled distance between home and school is between 2.0 and 5.0 km, 0 otherwise; *DIST5 km* takes the value 1 if the traveled distance between home and school is more than 5.0 km, 0 otherwise; *CARHH* is a continuous variable representing the number of cars in household; *HHSIZE* is a continuous variable denoting the number of household members; $\varepsilon_{WALK,}$ $\varepsilon_{BUS},$ ε_{CAR} are vectors of error terms.

The attitudes that teenagers have regarding their parents walking behavior are modeled as a function of their parents' socioeconomic and demographic characteristics, as shown in Eq. (12). Part of the explanatory variables that are used for the teenagers (decision makers) are common with the explanatory variables used for the structural model of parents, since they share the same household socioeconomic characteristics. The structural equation links parents' characteristics with the latent variable *PWL* through a linear regression equation based on their parents level of education, family income and number of available motorized vehicles in the household. At this point, it is worthwhile to mention that the proposed methodology could be applied without having a structural model for the social environment, but with having only the measurement model regarding the social environment, an issue that would be further presented in future work.

$$PWL = \zeta_{PWL} + \zeta_{CARHH} * (CARHH/NLICENSE)$$
$$+ (\zeta_{INC} + \zeta_{EDH} * EDUHIGH + \zeta_{EDL} * EDULOW) * INCOME * +\eta_{PWL} \tag{12}$$

where: *NLICENSE* represents the number of driving license holders; it takes the value 1 when one of the parents holds a driving license and 2 when both parents have a driving license. All the households have at east one parent that holds a driving license, thus this variable does not take the value 0. *EDULOW* takes the value 1 when the educational level of both parents is high, 0 otherwise; *EDUHIGH* takes the value 1 when the educational level of both parents is low, 0 otherwise; η_{PWL} is random error term.

The attitudes of teenagers regarding walking are modeled as a function of socioeconomic characteristics and the latent variable "Parents: Walking-lovers" (Eq. 13). The structural equation links teenagers' characteristics with the latent variable *WL* through a linear regression equation based on gender, age, if they lived in another country integrated by household income and the latent variable *PWL*. We multiply the variable *ABROAD* by income, due to the fact that in Cyprus there are both a lot of economic immigrants, but also a lot of wealthy foreign residents.

$$WL = \theta_{WL} + \theta_{GWL} * FEMALE + \theta_{AWL} * AGE1114$$
$$+ (\theta_{ABWL} + \theta_{INCWL} * INCOME) * ABROAD + \xi_{WL} * PWL + \omega_{WL} \tag{13}$$

where: *ABROAD* takes the value 1 if the teenager was living in a different country in the past, 0 otherwise; ω_{WL} is random error term.

Measurement model

The choice between the alternatives is assumed to be based on utility maximization and can be expressed as follows:

$$y = \begin{cases} 1 & \text{if} \quad U_i \geq U_j \; \forall j \neq i \\ 0 & \text{otherwise} \end{cases}, \quad i = WALK, BUS, CAR \tag{14}$$

where y_i is the choice indicator, 1 if alternative is chosen, 0 otherwise.

Four measures are used as indicators of the latent variables "Parents: Walking-lovers" as shown in Eq. (15) to (18). Equation (15) is normalized by setting the intercept term to 0 and the coefficient of attitude to 1. The indicators are specified as continuous variables for simplicity. At this point, it is worthwhile to mention that the proposed methodology could be implied without having a structural model for the social environment latent variable, but with having only the indicators regarding the social environment (measurement model).

$$SI_{PWL1} = a'_1 + \lambda'_1 * PWL + \upsilon'_1; \quad \alpha'_1 = 0, \quad \lambda'_1 = 1 \tag{15}$$

$$SI_{PWL2} = a'_2 + \lambda'_2 * PWL + \upsilon'_2 \tag{16}$$

$$SI_{PWL3} = a'_3 + \lambda'_3 * PWL + \upsilon'_3 \tag{17}$$

$$SI_{PWL4} = a'_4 + \lambda'_4 * PWL + \upsilon'_4 \tag{18}$$

where: SI_{PWL1}, SI_{PWL2}, SI_{PWL3}, SI_{PWL4} are responses to the attitudinal questions regarding parents (Table 2), $\upsilon'_1, \upsilon'_2, \upsilon'_3, \upsilon'_4$ are random error terms with unitary variance from $\upsilon'_1 \sim N(0, \sigma^2_{\upsilon'_{PWL1}})$ to $\upsilon'_4 \sim N(0, \sigma^2_{\upsilon'_{PWL4}})$ and $\alpha'_1, \alpha'_2, \alpha'_3, \alpha'_4, \lambda'_1, \lambda'_2, \lambda'_3, \lambda'_4$ are parameters.

Eight measures are used as indicators of "Walking-lover" (Eq. 19 to 26). Equation (19) is normalized by setting the intercept term to 0 and the coefficient of attitude to 1.

$$I_{WL1} = a_1 + \lambda_1 * WL + \upsilon_1; \quad \alpha_1 = 0, \lambda_1 = 1 \tag{19}$$

$$I_{WL2} = a_2 + \lambda_2 * WL + \upsilon_2 \tag{20}$$

$$I_{WL3} = a_3 + \lambda_3 * WL + \upsilon_3 \tag{21}$$

$$I_{WL4} = a_4 + \lambda_4 * WL + \upsilon_4 \tag{22}$$

$$I_{WL5} = a_5 + \lambda_5 * WL + \upsilon_5 \tag{23}$$

$$I_{WL6} = a_6 + \lambda_6 * WL + \upsilon_6 \tag{24}$$

$$I_{WL7} = a_7 + \lambda_7 * WL + \upsilon_7 \tag{25}$$

$$I_{WL8} = a_8 + \lambda_8 * WL + \upsilon_8 \tag{26}$$

where: I_{WL1}, I_{WL2}, I_{WL3}, I_{WL4}, I_{WL5}, I_{WL6}, I_{WL7}, I_{WL8} are the responses of teenagers to the attitudinal questions regarding their own behavior (Table 2), υ_1, υ_2, υ_3, υ_4, υ_5, υ_6, υ_7, υ_8 are random error terms with unitary variance from $\upsilon_1 \sim N(0, \sigma^2_{\upsilon_{WL1}})$ to $\upsilon_8 \sim N(0, \sigma^2_{\upsilon_{WL8}})$ and α_1, α_2, α_3, α_4, α_5, α_6, α_7, α_8, λ_1, λ_2, λ_3, λ_4, λ_5, λ_6, λ_7, λ_8 are parameters.

The likelihood of a given observation is the joint probability of observing the choice, the eight indicators of the attitude 'Walking-lover' and four indicators of the attitude "Parents: Walking-lovers", as shown in Eq. (27).

$$f(y, I_{WL}, IS_{PWL}|X, S; \delta)$$
$$= \int\limits_{WL} \int\limits_{PWL} P(y|X, X^*; \beta, \gamma, \Sigma\varepsilon)f(I_{WL}|X, WL; \lambda, \Sigma\upsilon)(IS_{PWL}|S, PWL; \lambda', \Sigma\upsilon')f(WL|X; \vartheta, \xi, \Sigma\omega)$$
$$(PWL|S; \zeta, \Sigma\eta)dPWL\,dWL$$

$$(27)$$

where δ designates the full set of parameters to estimate ($\delta = \{\beta, \gamma, \lambda, \lambda', \theta, \zeta, \xi, \Sigma\varepsilon, \Sigma\upsilon, \Sigma\upsilon', \Sigma\omega\}$).

Model estimation results

Mode choice model

This section presents and discusses the estimation results of the choice model (see Table 3). We first estimated an MNL model served as base model. Afterwards in the MNL we added the latent variable *WL* without including the social interaction latent variable *PWL. Finally,* we estimated again the MNL model with the latent variable WL including in its structural model the *PWL* latent variable as a component. The models were estimated using the Pythonbiogeme software (Bierlaire and Fetiarson 2009). The number of draws was set to 1,000.

Overall, the estimated values of the parameters are in agreement with prior expectations. All the variables used for the estimation of the choice model are statistically significant at the 95 % level. The constants in the model capture the preference of teenagers to private motorized vehicles and bus for their trip to school.

Adolescent females prefer being escorted by their parents by car rather than walking to school. Teenagers aged from 11 to 14 years old also prefer being escorted by car, while older teenagers from 15 to 18 years old prefer walking. This result reflects the fact that teenagers tend to conduct more independent (unsupervised) trips while they reach the age of 18. As monthly household income increases, the probability of teenagers being escorted to school increases too, while the probability of walking decreases. Also, as the ratio of the number of available private vehicles in the household and household members increases, the probability of teenagers to choose to be driven to school by their parents increases too.

Regarding the characteristics of the built environment, the existence of wide sidewalks at least at the ½ of the route from home to school encourages the choice of walk. The absence of crosswalks at least at the 1/2 of the route between home and school decreases the probability of choosing walking. As far as the aesthetics of the route between home and school, the existence of trees and flowers favors significantly the choice of walk. Distance plays the most significant role in mode to school choice, a fact that other surveys have verified as well (McDonald 2008; Schlossberg et al. 2006). Walking is preferred when the walking distance between home and school is less than 2.0 km. In the utility of *WALK*, distance is interacted with gender and the results indicate that even the distance is less than 2.0 km, females do not prefer walking. If the distance from home to school is more than 5.0 km, then bus is preferred. For distances between 2.0 and 5.0 km, teenagers prefer being escorted by private motorized vehicles.

Unsurprisingly, the incorporation of the latent variable "Walking-lover" (*WL*) enhances the explanatory power of the choice model. The latent variable enters significantly into the utility of walk, while it is the most statistically significant variable. Thus, the latent variable

Table 3 Estimation results—choice model

	Base model		HCM		HCM with social interaction	
	Coef.	t stat	Coef.	t stat	Coef.	t stat
β_{BUS}	1.07	5.53	3.72	14.94	3.75	15.04
β_{CAR}	2.47	12.27	5.18	20.21	5.21	20.30
Variables specific to WALK						
Female	−0.30	−3.51	−0.28	−2.94	−0.28	−2.98
Distance between home and school: less than 2.0 km	3.53	20.46	3.67	20.34	3.67	20.34
Age 11–14 years old	−0.19	−2.27	−0.41	−4.36	−0.41	−4.40
Income (continuous)	−0.09	−1.97	−0.43	−1.98	−0.43	−1.97
Existence of green (trees/flowers)	0.434	5.35	0.42	4.78	0.42	4.72
No crosswalks at least at the 1/2 of the route between home and school	−0.60	−7.86	−0.65	−7.68	−0.65	−7.62
Wide sidewalks at least at the 1/2 of the route between home and school	0.29	4.04	0.27	3.39	0.27	3.41
Walking-lover (latent variable)	–	–	0.58	20.67	0.89	23.94
Variables specific to BUS						
Distance between home and school: more than 5.0 km	5.19	10.83	5.12	10.64	5.12	10.64
Variables specific to CAR						
Female	0.11	1.97	0.098	1.98	0.098	1.98
Age 11 to 14 years old	0.25	4.29	0.25	4.26	0.33	8.94
Distance between home and school: 2.0 to 5.0 km	1.97	4.12	1.91	3.97	1.91	3.97
Income (continuous)	0.04	4.38	0.04	4.21	0.04	4.22
Number of private vehicles in the household/Household size (continuous)	1.18	12.24	1.08	10.97	1.08	10.96

encourages the choice of walk to school through a positive impact in the choice of this alternative.

Structural and measurement latent variable model estimation results

Table 4 presents the estimation results of the structural and the measurement models of the latent variable models. All variables used in the structural models are statistically significant at the 95 % level. From the structural model, we can conclude that girls are less "Walking-lovers" than boys. The age between 11 and 14 years old has a negative impact on walking-loving behavior. The participants who were living in a different country in the past seem to be "Walking-lovers". But when this variable is interacted with the income, the results indicate that even they were living abroad in the past, they are not "Walking-lovers". This reflects the fact that the wealthy immigrants in Cyprus do not have positive attitudes towards walking. The majority of the wealthy immigrants were living in Russia before they came to Cyprus, while the less wealthy were living either in Asia–Pacific countries or in countries of North Africa (Egypt, Morocco).

The incorporation of the latent variable "Parents: Walking-lovers" into the latent variable "Walking-lovers" enhanced even more the explanatory power of the model. This component is the most statistically significant variable in the structural equation of *WL*, indicating the strong influence that parents have on the development of their children's attitudes towards walking. The results indicate that when teenagers anticipate that their parents love to walk, then this fact affects positively their attitude regarding walking.

The structural model of the latent variable *PWL* offers significant information about the characteristics of parents that favor a walking-lover behavior. With respect to the educational attainment of the parents, higher levels of education for both parents are associated with greater levels of "Walking-lover" behavior. In contrast, low educational levels for both parents work against a walking-love behavior. However, the variable regarding high educational level of parents is interacted with income, which estimate value is -0.15. This means that as income increases, the positive sign of high educational level is reversed. Generally, Cypriots are highly car-oriented and owning and using a car is a sign of wealth.

Regarding the measurement model of the latent variable *WL*, several indicators were considered, which linked the latent variable of psychometric "walking-lover" behavior to the responses to the attitudinal qualitative survey questions. The coefficient of the first indicator (I_{WL1}) was normalized to 1. The α parameters that indicate the associations between the responses to the scale items and the psychometric scale all have the expected signs. However, the α_2 parameter is negative, indicating that teenagers do not consider travel cost as one of the most important transport mode attributes (see I_{WL2}, Table 2). Here, we can see that a more positive attitude to walk will lead to respondents being more in agreement with the statement that they prefer walking rather than being escorted. Additionally, the effect of the latent variable *WL* on the indicator about environmental protection is positive, reflecting the idea that environmentally conscious teenagers perceived the idea of walking more positively because this is one of the most environmental friendly transport modes.

For the measurement model of the latent variable *PWL* were used indicators that linked the latent variable to the responses to attitudinal qualitative survey questions regarding the walking behavior of the participants' parents. The coefficient of the indicator SI_{PWL1} was normalized to 1. The results indicate that the latent variable *PWL* has a positive effect on the indicators regarding the preference of walking instead of using car, reflecting the idea that walking-lover parents prefer green transport modes.

Table 4 Estimation results—structural model

	HCM		HCM with social interaction	
	Coef.	t stat	Coef.	t stat
Structural model				
Decision maker/teenager—walking-lover				
θ_{WL}	4.00	114.85	1.80	9.48
Female	−2.35	−6.35	−0.13	−3.69
Age 11–14 years old	−0.28	−7.57	−0.33	−8.94
Abroad	0.64	4.23	0.63	4.56
Income	−0.32	−3.95	−0.33	−3.98
PWL	–	–	0.58	11.72
σ_{WL}	1.68	78.10	1.66	77.36
Social environment—parents: walking-lovers				
ζ_{PWL}	–	–	3.58	119.34
Number of private vehicles in the household/ number of persons with driving license in the household	–	–	0.024	4.70
Income	–	–	−0.15	−1.96
Both parents high educational level	–	–	0.04	4.06
Both parents low educational level	–	–	−0.23	−2.87
σ_{PWL}	–	–	0.50	20.31
Measurement model				
α_1	0	–	0	–
α_2	−0.19	−3.48	−0.20	−3.60
α_3	0.09	1.48	0.08	1.35
α_4	0.22	3.74	0.21	3.60
α_5	0.42	6.97	0.43	7.04
α_6	2.27	38.10	2.28	38.26
α_7	4.93	81.17	4.93	81.11
α_8	3.91	74.87	3.93	75.07
λ_1	1	–	1	–
λ_2	0.95	72.99	0.95	72.99
λ_3	0.96	65.29	0.96	65.35
λ_4	0.96	67.28	0.96	67.34
λ_5	0.91	64.02	0.91	63.89
λ_6	0.38	27.48	0.37	27.25
λ_7	−0.19	−13.81	−0.19	−13.79
λ_8	−0.18	−14.97	−0.17	−14.68
σ_1	1.56	106.61	1.56	106.88
σ_2	1.42	102.84	1.41	102.69
σ_3	1.33	98.76	1.33	98.75
σ_4	1.27	95.81	1.27	95.73
σ_5	1.58	119.14	1.58	119.39
σ_6	2.07	137.19	2.07	137.23
σ_7	2.19	138.86	2.19	138.86

Table 4 continued

	HCM		HCM with social interaction	
	Coef.	t stat	Coef.	t stat
σ_8	1.86	138.79	1.86	138.81
α'_1	–	–	0	–
α'_2	–	–	−7.74	−14.89
α'_3	–	–	−2.41	−8.59
α'_4	–	–	−7.74	−14.72
λ'_1	–	–	1	–
λ'_2	–	–	2.8	19.91
λ'_3	–	–	1.46	19.27
λ'_4	–	–	2.87	19.17
σ'_1	–	–	1.88	133.92
σ'_2	–	–	1.41	72.26
σ'_3	–	–	1.81	127.43
σ'_4			1.35	64.11

Conclusions

This paper presented a general methodology and framework for including social interaction effect into HCM (Walker and Ben-Akiva 2001; Ben-Akiva et al. 2002). Based on the findings in psychology and neuroscience research that the individual's decisions are indirectly influenced by the social environment, as it affects the individual's psychological state (van den Bos et al. 2013; Homberg 2012), the developed method provides insights for modeling the effect of social interaction on the formation of psychological factors (latent variables) and on the decision-making process. Thus, the social environment is a latent variable that represents social interaction with the decision maker and it is included as a component to the latent variable regarding the decision maker, which in turn is included directly in the choice model.

The proposed methodology requires the estimation of an integrated multi-equation model consisting of a discrete choice model, the latent variable model's structural and measurement equations regarding the decision maker and the latent variable model's structural and measurement equations regarding the social environment. The model structure is simultaneously estimated providing an improvement over sequential methods as it provides consistent and efficient estimates of the parameters. Maximum likelihood techniques are used to estimate the integrated model, in which the likelihood function for the integrated model includes complex multi-dimensional integrals (one integral per latent construct).

The methodology is tested within the context of a household aiming to identify the social interaction effects between teenagers and their parents regarding walking-loving behavior and then the effect of this on the mode to school choice behavior. The sample consists of 9,714 participants aged from 12 to 18 years old, representing 21 % of the total adolescent population of Cyprus, while only revealed preference data are used. The findings from the case study indicate that if the teenagers perceive that their parents -as walking lovers, then this increases the probability of teenagers to be walking-lovers too. The latent variable "Parents: Walking-lovers" is the most statistically significant variable

in the formulation of the latent variable "Walking-lover" that refers to the decision maker. Then the latter latent variable is incorporated directly in the utility of the alternative walking and through a positive sign affects significantly the choice of walking to school. Thus, the findings from the case study are that implementation of the integrated choice, latent variable and latent social interaction model framework results in: (a) improvements in the explanatory power of choice models, (b) latent variables that are statistically significant, and (c) a more real-world behavioral representation that includes the social interaction effect. Other variables that affect the mode to school choice behavior are distance, income, age, gender, vehicle ownership, household size and various built environment characteristics, that are consistent with the findings of other mode school choice behavior surveys (McDonald 2008; Clifton et al. 2010).

The data required for implying this methodology is easy to be collected. The main requirement is to include in the questionnaire attitudinal questions regarding the travel behavior of the social environment of the participant in the survey. Then these attitudinal questions are used for the development of latent variables regarding the social environment of the participant. The methodology provides the ability to researchers to specify as many latent variables for the social environment as they want. For example, different latent variables could be used for parents, siblings, friends, colleagues etc., each one representing a different social network. Moreover, this could provide insights about which social network affects more the behavior of the decision maker. Furthermore, the proposed modeling framework could be implied not only in transportation sector, but also to other sectors that study choice behavior (i.e. marketing).

Concluding, further research includes the investigation of social interaction between the decision maker and other social networks such as friends and colleagues. Moreover, due to the fact that it is difficult for the researchers to know the socio-economic characteristics of all the members of the social environment of the decision-maker, a modeling framework will be presented that could model the effect of social interaction based only on the attitudes of the decision maker regarding her/his social environment without being necessary to include the socio-economic characteristics of the social environment. Further research will also investigate other latent variables, such car-loving behaviors. Finally, future work will also present the goodness-of-fit tests for the proposed methodology.

References

Abelson, R.P., Levy, A.: Decision making and decision theory. In: Lindzey, G., Aronson, E. (eds.) Handbook of social psychology, 3rd edn. vol. 1, pp. 231–309, Random House, New York (1985)

Abou-Zeid, M., Ben-Akiva, M.: The effect of social comparisons on commute well-being. Transp. Res. Part A **45**(4), 345–361 (2011)

Abou-Zeid, M., Ben-Akiva, M., Bierlaire, M., Choudhury, C.F., and Hess, S.: Attitudes and value of time heterogeneity. Presented at the 90th Annual Meeting of the Transportation Research Board, Washington DC (2011)

Axhausen, K.W.: Activity spaces, biographies, social networks and their welfare gains and externalities: some hypotheses and empirical results. Presented at the PROCESSUS Colloquium, Toronto, Canada (2005)

Babey, S., Hastert, T., Huang, W., Brown, E.R.: Sociodemographic, family, and environmental factors associated with active commuting to school among US adolescents. J. Public Health Policy **30**, 203–220 (2009)

Ben-Akiva, M., Lerman, S.R.: Discrete Choice Analysis: Theory and Application to Travel Demand. MIT Press, Cambridge (1985)

Ben-Akiva, M.: Planning and action in a model of choice. In: Hess, S. & Daly, A. (eds.) Choice Modelling: The State-of-the-Art and the State-of-Practice. Proceedings from the Inaugural International Choice Modelling Conference, pp. 19–34. Bingley, UK (2010)

Ben-Akiva, M., Walker, J., Bernardino, A.T., Gopinath, D.A., Morikawa, T., Polydoropoulou, A.: Integration of choice and latent variable models. In: Mahmassani, H.S. (ed.) Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges, pp. 431–470. Elsevier, Amsterdam (2002a)

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D.S., Daly, A., de Palma, A., Gopinath, D., Karlstrom, A., Munizaga, M.A.: Hybrid choice models: progress and challenges. Mark. Lett. **13**(3), 163–175 (2002b)

Ben-Akiva, M., de Palma, A., McFadden, D., Abou-Zeid, M., Chiappori, P.A., de Lapparent, M., Durlauf, S.N., Fosgerau, M., Fukuda, D., Hess, S., Manski, C., Pakes, A., Picard, N., Walker, J.: Process and context in choice models. Mark. Lett. **23**, 439–456 (2012)

Bhat, C., Sener, I.N., Eluru, N.: A flexible spatially dependent discrete choice model: formulation and application to teenagers' weekday recreational activity participation. Transp. Res. Part B **44**, 903–921 (2010)

Bierlaire, M., Fetiarson, M.: Estimation of discrete choice models: extending BIOGEME. Presented at 9th Swiss Transport Research Conference, Ascona, Switzerland (2009)

Bierlaire, M.: Discrete choice models. In M. Labbé, G. Laporte, K. Tanczos and P. Toint (eds.) Operations Research and Decision Aid Methodologies in Traffic and Transportation Management, Vol. 166, NATO ASI Series, Series F: Computer and Systems Sciences, Springer, Verlag, pp. 203–227 (1998)

Bolduc, D., Boucher, N., Alvarez-Daziano, R.: Hybrid choice modeling of new technologies for car choice in Canada. Transportation Research Record: Journal of the Transportation Research, No. 0361–1981, pp. 63–71. Transportation Research Board of the National Academies, Washington DC (2008)

Brock, W., Durlauf, S.: Discrete choice with social interactions. Rev. Econ. Stud. **68**, 235–260 (2001)

Carrasco, J., Miller, E.: Exploring the propensity to perform social activities: a social network approach. Transportation **33**, 463–480 (2006)

Christakis, N., Fowler, J.: Connected: The Surprising Power of our Social Networks and How They Shape our Lives. Hachette Digital Inc. , New York (2009)

Clifton, J., Akar, G., Smith, A.L., Voorhees, C.C.: Gender differences in adolescent travel to school: exploring the links with physical activity and health. Women Issues Transp. Conf. Proc. 46 **2**, 203–222 (2010)

Crane, J.: The epidemic theory of ghettos and neighborhood effects on dropping out and teenage child-bearing. Am. J. Sociol. **96**, 1226–1259 (1991)

Dugundji, E., Walker, J.: Discrete choice with social and spatial network interdependencies: an empirical example using mixed generalized extreme value models with field and panel effects. Transp. Res. Rec. **2005**, 70–78 (1921)

Durlauf, S., Walker, J.: Social Interactions and Fertility Transitions. University of Wisconsin at Madison, Mimeo (1998)

Fukuda, D., Morichi, S.: Incorporating aggregate behavior in an individual's discrete choice. An application to analyzing illegal bicycle parking behavior. Transp. Res. Part A **41**, 313–325 (2007)

Gärling, T., Gillholm, R., Gärling, A.: Reintroducing Attitude theory in travel behavior research: the validity of an interactive interview procedure to predict car use. Transportation **25**, 147–167 (1998)

Goetzke, F.: Network effects in public transit use: evidence from a spatially autoregressive mode choice model. Urban Stud. **45**, 407–417 (2008)

Golub, B., Jackson, M.: Naive learning in social networks and the wisdom of crowds. Am. Econ. J. **2**(1), 112–149 (2010)

Granovetter, M.: The impact of social structure on economic outcomes. J. Econ. Perspect. **19**(1), 33–50 (2005)

Haveman, R., Wolfe, B.: Succeeding Generations. Russell Sage Foundation, New York (1994)

Homberg, J.R.: Serotonin and decision making processes. Neurosci. Biobehav. Rev. **36**, 218–336 (2012)

Johansson, M.V., Heldt, T., Johansson, P.: The effects of attitudes and personality traits on mode choice. Transp. Res. Part A **40**(6), 507–525 (2006)

Kamargianni, M.: Development of hybrid models of teenager's travel behavior to school and to after-school activities. PhD thesis, University of the Aegean (2014)

Kamargianni, M., Polydoropoulou, A.: Does social networking substitute for or stimulate teenagers' travel? Findings from a latent class model. Presented at the 3rd International Choice Modeling Conference (ICMC), Sydney, Australia (2013a)

Kamargianni, M., Polydoropoulou, A.: Hybrid choice model to investigate effects of teenagers' attitudes toward walking and cycling on mode choice behavior. Journal of the Transportation Research Record, Vol.2382-Travel Behavior(1), 151–161(2013b)

Kivetz, R., Netzer, O., Srinivasan, V.: Alternative models for capturing the compromise effect. J. Mark. Res., **41**(3), 237–257 (2004)

Manski, C.: Identification of endogenous social effects: the reflection problem. Rev. Econ. Stud. **60**, 531–542 (1993)

McDonald, N.C.: Children's mode choice for the school trip: the role of distance and school location in walking to school. Transportation **35**(1), 23–35 (2008)

McFadden, D.: Rationality for economists. Presented at the NSF Symposium on Eliciting Preferences, Berkeley, California (1997)

McMillan, T.E.: Urban form and a child's trip to school: the current literature and a framework for future research. J. Plan. Lit. **19**, 440–456 (2005)

McMillan, T.E.: The relative influence of urban form on a child's travel mode to distance. Transp. Res. Part A **41**, 69–79 (2007)

Olson, J.M., Zanna, M.: Attitudes and attitude change. Annu. Rev. Psychol. **44**, 117–154 (1993)

Overman, W.H.: Sex differences in early childhood, adolescence, and adulthood on cognitive tasks that rely on orbital prefrontal cortex. Brain Cognit. **55**, 134–147 (2004)

Páez, A., Scott, D.M.: Social influence on travel behavior: a simulation example of the decision to tele-commute. Environ. Plan. A **39**, 647–665 (2007)

Polydoropoulou, A., Kamargianni, M., Tsirimpa, A.: Car use addiction versus ecological consciousness: which prevails in mode choice behavior for young people? In: Roorda, M., Miller, E. (eds.) Travel Behaviour Research: Current Foundation, Future Prospects. International Association for Travel Behavior Research (IATBR), 322–340 (2013)

Rabin, M.: Psychology and economics. J. Econ. Lit. **36**, 11–23 (1998)

Rose, J.M., Hensher, D.A.: Handling individual specific availability of alternatives in stated choice experiments. Presented at the 7th International Conference on Travel Survey Methods, Los Suenos, Costa Rica (2004)

Schlossberg, M., Greene, J., Paulsen, P., Johnson, B., Parker, B.: School trips: effects of urban form and distance on travel mode. J. Am. Plan. Assoc. **72**, 337–346 (2006)

Seraj, S., Sidharthan, R., Bhat, C., Pendyala, R., Goulias, K.: Parental attitudes towards children walking and bicycling to school: a multivariate ordered response analysis. Presented at 91st Annual Meeting of the Transportation Research Board, Washington DC (2011)

Shiftan, Y., Outwater, M.L., Zhou, Y.: Transit market research using structural equation. Modeling and attitudinal market segmentation. Transp. Policy **15**, 186–195 (2008)

Tsirimpa, A., Polydoropoulou, A.: Development of a mixed multi-nomial logit model to capture the impact of information systems on travelers' switching behavior. J. Intell. Transp. Systems: Technol. **11**(2), 79–89 (2007)

van den Bos, R., Jolles, W.J., Homberg, J.R.: Social modulation of decision-making: a cross-species review. Front Hum Neurosci **7**(301), 1–16 (2013)

Visser, L., Homberg, J.R., Mitsogiannis, M., Zeeb, F.D., Rivalan, M., Fitoussi, A.: Rodent versions of the Iowa gambling task: opportunities and challenges for the understanding of decision-making. Front. Neurosci. **5**, 109–122 (2011)

Walker, J.L., Ben-Akiva, M.: Generalized random utility model. Math. Soc. Sci. **43**(3), 303–343 (2001)

Webster, M.M., Ward, A.J.W.: Personality and social context. Biol. Rev. **86**, 759–773 (2011)

Weinberg, B., Pehlivan, E.: Social spending: managing the social media mix. Bus. Horiz. **54**(3), 275–282 (2011)

Wilton, R.D., Páez, A., Scott, D.M.: Why do you care what other people think? A qualitative investigation of social influence and telecommuting. Transp. Res. Part A **45**(4), 269–282 (2011)

Yoon, S.Y., Doudnikoff, M., Goulias, K.: Spatial analysis of the propensity to escort children to school in Southern California. Presented at 90th Annual Meeting of the Transportation Research Board, Washington DC (2011)

Maria Kamargianni holds a PhD in Travel Behavior Modeling and Decision Sciences. She is a Research Associate in Energy and Transport at the UCL Energy Institute, University College of London (UCL). Her areas of research include travel behavior modeling and generational differences, social interaction effects on travel behavior, school transportation, market research and econometrics.

Moshe Ben-Akiva is the Edmund K. Turner Professor of Civil and Environmental Engineering at the Massachusetts Institute of Technology (MIT), and Director of the MIT Intelligent Transportation Systems (ITS) Lab. He holds a Ph.D. degree in Transportation Systems from MIT and 4 honorary degrees and coauthored the textbook *Discrete Choice Analysis*.

Amalia Polydoropoulou is a Professor at the Department of Shipping, Trade, and Transport at the University of the Aegean (Greece) and the Director of the Transportation and Decision Making Laboratory (TransDeM). Her areas of research include transportation systems analysis, transportation demand modeling, market research and econometrics.