Effects of life events and attitudes on vehicle transactions: A dynamic Bayesian network approach

Transportation Research Part C 147 (2023) 103988

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Effects of life events and attitudes on vehicle transactions: A dynamic Bayesian network approach

Yajie Yang *, Soora Rasouli, Feixiong Liao

Urban Planning and Transportation Group, Eindhoven University of Technology, Eindhoven, the Netherlands

A R T I C L E I N F O

Keywords:
Vehicle transaction model
Life events
Attitude
Dynamic Bayesian network (DBN)

A B S T R A C T

Individual and household life events are interdependent and influence mobility-related decisions at different levels over time. This paper developed an integrated dynamic model to capture the interdependences among life events, with a special focus on vehicle transactions. Particular attention was paid to the inclusion of vehicles' characteristics such as the age, fuel type, and size of cars, which are pertinent to emission forecast. A dynamic Bayesian network (DBN), containing individual and household characteristics and latent attitudes toward car ownership and use alongside life events, was employed to study the interdependences. The temporal relationships among life events and lead-lag effects were also captured in the DBN. The longitudinal survey data "the Netherlands Mobility Panel (MPN)" from 2013 to 2018 was used to train and test the DBN. The analysis results confirm the dynamic interdependences between vehicle transactions and other life events and reveal noticeable associations between attitudes and purchase decisions. It is found that several life events (e.g., "Birth of a baby", "Marital status change") have concurrent or varied lag-effects on vehicle transaction decisions. The validation indicates that the proposed DBN approach has a high predictive accuracy of vehicle transaction decisions and other life events.

1. Introduction

Car ownership exerts a large effect on travel demand, and the change of car ownership has a strong influence on travel behavior (Ben-Akiva and Lerman, 1974; Train et al., 1982; Bhat and Guo, 2007; Bhat et al., 2009; Kim et al., 2017; Jain et al., 2021). Compared to other transport modes, the car gives more freedom and flexibility in scheduling and conducting activities and travel under spatial-temporal constraints (Rasouli and Timmermans, 2014; Liao, 2019; Qin and Liao, 2021). The convenience and other values brought by owning a car stimulate car ownership. For example, the number of passenger cars has grown >200 times (8.5 million in 2019) since 1927 (41 thousand) in the Netherlands. This growth has been caused mainly by population growth, the rise of income, and the increasing need of commuting over medium and long distances. Car ownership has a direct effect on traffic congestion, energy consumption, air quality, greenhouse gas emission, etc. (Achtnicht, 2012; Beck et al., 2013; Knez et al., 2014; Serrenho et al., 2017; Fontaras et al., 2017; Mazur et al., 2018; Moro and Lonza, 2018; Gómez Vilchez et al., 2019; Zhao et al., 2022).

Travel demand forecasting models are powerful tools to predict the future travel demand of a targeted population allowing the authorities to devise policies toward a desirable future. The large scale operational travel demand forecasting models typically contain population synthesizer modules to establish the future configuration of the targeted population in terms of various socio-demographics

* Corresponding author.
E-mail addresses: y.yang2@tue.nl (Y. Yang), s.rasouli@tue.nl (S. Rasouli), f.liao@tue.nl (F. Liao).

https://doi.org/10.1016/j.trc.2022.103988
Received 19 November 2021; Received in revised form 5 December 2022; Accepted 8 December 2022
Available online 5 January 2023

and economic attributes. While a potentially valuable application of travel demand forecasting models can be the prediction of emissions, such application requires a dedicated vehicle transaction or car ownership model that contains information about fuel type, emission code, and size of vehicles as part of household characteristics. In the current practice, due to the absence of such dedicated models, in the majority of travel demand forecasting models, emission evaluation, even if a disaggregate travel demand model is employed for the calculation of travel distance, is commonly done in a rather aggregate and cross-sectional manner by associating the share of different vehicle types in the fleet with the overall distance traveled (Degraeve and Proost, 2001; Velders et al., 2013; Khan and Habib, 2021). Although such aggregation is expected to exert biases, due to the scarcity of a comprehensive car holding model, either in isolation or as a part of a population synthesizer, such practice continues to be dominant in research and practice. Therefore, it is important to capture the dynamics and specifics in car ownership, ideally as an integral part of the population synthesizers.

Previous empirical studies on car ownership are mainly cross-sectional. Apart from model specification, all previous studies explored the influencing factors on car ownership, such as socio-demographic variables, built environment characteristics, and life events. The application of these models for emission forecast and energy consumption is restrictive due to two major shortcomings. First, these models lack the necessary components such as size, fuel type, and age of the cars. Second, the majority of these models are cross-sectional and ignore the impacts of life events of preceding years on changes in car ownership decisions in succeeding years. Examples of such occasions include purchasing an additional car after a baby is born or after changing residence or workplace.

In response to these shortcomings, dynamic vehicle ownership models (containing vehicle transactions and vehicle holding) were developed. The vehicle holding models describe the choices of a household in holding a certain number of cars (with specific types) within a time frame (typically one year), while vehicle transaction models explain the decision to dispose, replace, or acquire a car. To some extent, the vehicle holding status is the result of vehicle transactions over time (Mohammadian and Miller, 2002; de Jong and Kitamura, 2009; Khan and Habib, 2021).

To link with disaggregate travel demand forecasting, it is crucial to integrate the vehicle transaction model as part of a dynamic synthetic population. To the best of our knowledge, none of the few comprehensive dynamic population synthesizers have an elaborated car transaction module including necessary components such as age, fuel type, and size, and their interlinks with other life events. Characteristics of the vehicle (age, fuel type and size) are intuitively important for decision to purchase, sell and replace cars. When the owned vehicles get older, increasing maintenance costs and the desire to have new models and can be the impetus to change them. Fuel type can impact the decision as well because of the expenses associated with it (tax, fuel price) as well as environmental concerns or status. Lastly, size of the vehicle and the relation with the household need (household changes in size) is yet another intuitively important variable to consider in modeling vehicle transaction decisions. Moreover, although attitude toward car ownership and use appeared to be a significant determinant in the car ownership decision (Kim and Mokhtarian, 2018), it has not been included in any of the previous population synthesizers mainly due to the scarcity of dynamic attitudinal data. Challenges underlying synthesizing attitude is yet another reason to be cautious of including such latent variables in the modeling of car transaction decisions. That being said, advances in deep learning algorithms and their applications in various fields including travel behavior and social science may promise a future in which attitudinal constructs can be imputed "as unobserved components" from other observable individuals' backgrounds and living experiences.

To contribute to the limited body of knowledge on comprehensive vehicle transaction models, this study developed an integrated dynamic model based on the dynamic Bayesian network (DBN) that integrates socio-demographic variables, life events such as, "Employment mobility", "Residential mobility", "Work location mobility", "Driving license obtain", "Marital status change", "Birth of a baby" with "Buy a car" and "Sell a car". Age, size, and fuel type of vehicles were considered as part of the DBN as well. In this context, the mobility of a life event refers to the change of status of that life event. For example, "Residential mobility" represents the change of residence place. Moreover, "Attitudes toward owning and using a car" was incorporated into the DBN. More precisely, "Attitude toward owning and using a car" measured by 13 indicators was used to cluster the sample and the clusters were used as an additional parent node for buying and selling cars in the DBN. DBN allows incorporating non-linear relation between factors and their lag effects without imposing any pre-defined causation hypothesis, all of which are important in simulation of relation between life events. The complexity of relationships between socio-demographics and multiple life events (of which vehicle transaction is one example) may indeed calls for a modeling approach that does not require a pre-imposed relations but allows the data, to a great extent, to reveal the relations. In addition, some previous studies tested the relations and lag effects in life events (e.g., Fatmi and Habib, 2016) and confirmed that birth of a baby has a lagged effect on the decision of buying a car. DBN is therefore a suitable method with required functionalities.

The remainder of this paper is organized as follows. Section 2 reviews the relevant studies while Section 3 presents the integrated conceptual framework. Section 4 describes the features of data used in this study. The results of latent class analysis and the prediction performance of the DBN are presented in Sections 5 and 6 respectively. The last section concludes the paper and reflects on the limitations of the suggested approach.

2. Literature review

Disaggregate car ownership models, typically based on the application of discrete choice models, received increased attention in the late 1980s with the concurrently increased interest in disaggregate models for travel demand forecasting. The vehicle ownership model can be classified into "cross-sectional vehicle ownership model" and "dynamic vehicle ownership model". The dynamic models aim to represent changes in vehicle ownership, possibly including other details such as age and fuel type over time. A vehicle ownership model, either in the static or dynamic context, can be further decomposed into vehicle holding that models the number of cars and vehicle transaction that models the transactions such as buying a car, selling a car, or replacing a car. Our literature review contains relevant studies concerning vehicle holding models and vehicle transaction models.

2.1. Cross-sectional vehicle ownership models

The cross-section models, particularly based on discrete choice methods, can be classified into the ordered models (Bhat and Guo, 2007; Maltha et al., 2017) and unordered models (Bhat et al., 2009; Potoglou and Kanaroglou, 2008; Potoglou and Susilo, 2008; Oakil et al., 2014). The ordered logit models reflect the propensity to own a certain ordered number of cars in the family (0 car, 1 car, 2 cars, 3 cars, etc.), while the unordered models typically use the random utility maximization theory and associate certain utility to the choice of owning a car with certain characteristics. Some comparisons between the two types of models were conducted and it was concluded that unordered models outperform (Bhat and Pulugurta, 1998; Potoglou and Susilo, 2008). The methods of static Bayesian network (Oakil, 2013) and multiple regressions (Flamm, 2009) were also used to model car-ownership decisions.

Bhat and Koppelman (1993) developed an integrated model considering the interdependencies among employment, income, and car ownership. The ordered response probit model was used to simulate the car-ownership decisions. The education level and employment status of both household heads appeared to be significant determinants of car ownership. Similarly, by using an ordered-response car ownership model, Bhat and Guo (2007) studied how the built environment and socio-demographic variables influence car-ownership decisions. The urbanization level and employment status showed significant impacts on car ownership. In addition, households' commuting time and cost as well as accessibility to public transportation proved to be influential determinants. Income, employment type, and home ownership also appeared as important factors. Potoglou and Kanaroglou (2008), using a multinomial logit model, found that urbanization level of residential place, commuting distance, household composition influence the decision of owning cars.

A few studies also explored the role of attitude toward the environment in general and owning and using of private cars in particular in car ownership decision. Flamm (2009) used sequential multiple regression and found that attitude toward the environment is strongly associated with car ownership. Households with pro-environmental attitudes prefer fewer cars, drive less, and possess cars with less environmentally destructive fuel types.

Belgiawan et al. (2011) investigated the effect of habitual behavior (frequency of car trips per week) and psychological factors on purchasing a new car using binary logistic regression. Travel distance to the workplace and the frequency of use of public transportation in their daily life in Bandung city in Indonesia appeared to be significant determinants. It was also found that the model with attitudinal variables outperforms the one without such variables.

Van Acker et al. (2014) developed a structural equation model to explain the influential factors of car ownership levels (the number of cars owned). They found out that attitude toward different modes (car, walking, cycling, and public transportation), household composition, and the residential neighborhood characteristics (regional accessibility, local accessibility, and the distance from home to local center and regional center) have significant influences on the number of cars owned. Oakil et al. (2016) applied a logistic regression model and found that urbanization level and childbirth were significant factors in explaining the differences in car ownership among young families in The Netherlands.

In terms of the fuel type, Valeri and Cherchi (2016) applied a hybrid choice model and revealed that habitual car users prefer to buy liquefied petroleum gas (LPG) and compressed natural gas over gasoline cars. Following this line of work in identifying the important determinants of people's choices for purchasing electric vehicles, Li and Zhao (2017) identified demographic factors, technical features, the cost, environmental attitude, and government policy as important factors. The role of latent attitude in people's decisions of purchasing EVs appeared in Kim et al. (2016) as well. The importance of pro-environmental attitudes and the symbolic value of the car was highlighted in their study. There are also other studies particularly focusing on the intention to purchase a certain fuel type (Daly et al., 2012; Kowalska-Pyzalska et al., 2021; Wang and Matsumoto, 2021).

2.2. Dynamic vehicle ownership models

The importance of developing dynamic car ownership models has been acknowledged by many researchers (e.g., Dargay and Vythoulkas, 1999; Nolan, 2010). Due to the lack of longitudinal data, dynamic car ownership models were not studied as much as static models.

Dargay (2001) applied a dynamic econometric model to the UK Family Expenditure Surveys datasets from1970 to 1995 to study the changes in the number of cars in the households and the related triggering factors. The results showed that the influence of household income on car ownership is asymmetric and has a certain lag. Besides habitual behavior, income and attitude play important roles in the decision of purchasing or disposing of a vehicle. In a similar study, the author found the urbanization level as a substantial determinant in the vehicle holding model realizing that rural households have less sensitivity to the price of the car because of relatively high expenses to meet their transportation need in the rural area compared to the people living in the urban areas (Dargay, 2002). Anne Nolan (2010) also conducted a dynamic analysis of household car ownership (the number of cars) and identified household income and number of existing cars as the strongest determinates in the car ownership decisions by estimating a dynamic random effects probit model calibrated with data from 1995 to 2001 in Ireland.

To explore the factors motivating households to change the number of cars over time, Oakil et al. (2014) estimated a dynamic mixed logit model to represent the relationship between the number of owned cars, life events, and other socio-demographics. A retrospective questionnaire tracking the past 21 years from 1990 to 2010 in the Netherlands was administrated to capture the temporal interdependencies among car ownership, socio-demographics, and life events. The results showed that changes in employment status, work location, residential location, marital status, as well as birth of children, moving out, death of a household member have significant influences on car-ownership decisions. A similar study conducted by Clark et al. (2016), using the UK Household Longitudinal Study (UKHLS) survey from 2009 to 2011 and a dynamic logistic regression model, revealed that socio-demographics and life events are dominant factors in predicting the change of car ownership (number of cars). More specifically, household composition and driver's license status appeared as the strongest predictors of car ownership change, followed by changes of employment status and income level. Maltha et al. (2017) estimated a series of ordered logistic regression models dynamically to model car ownership (number of cars) over a period from 1985 to 2014 in the Netherlands. The results suggested that household income, size and composition, persons' gender, age, education level, and working status, as well as suburbanization level influence car ownership, with the household income and household size contributing the most. The estimated parameters revealed that the relative importance of the variables changes over time. The importance of family size gradually increased from 30% in 1987 to 35% in 2014. The importance of household income had a quick rise from around 39% in 1987 to 47% in 1991 and then gradually decreased every year to around 27% in 1991. The relative importance of other variables remained stable during this period. In a similar study, estimating a latent class model, Gu et al. (2021), revealed that changes in household car ownership (electric/non-electric car, new or second-hand car considered) is sensitive to the change of household composition and job (also called work) or residential location mobility.

In a study by Xiong et al. (2018), a high-order hidden Markov model was applied for dynamic car ownership analysis, in which the interdependencies of hidden states could cross multiple periods. The life events also act as triggers to car ownership decisions besides related socio-demographics. Using retrospective life trajectory data and applying DBN, Guo et al. (2019, 2020) demonstrated the interdependencies among life events. The results showed that both "Residential mobility" and "Work location mobility" have positive influences on the change of car ownership.

It has been commonly recognized that the dynamic vehicle ownership model has advantages over the static counterpart in capturing the impact of changes in individual and household characteristics, triggered by life events, on car holding and transaction. However, it is important to realize that life events themselves are also influenced mutually and by social, economic, and environmental determinants. Moreover, there might be a time lag between one life event and another, although one still contributes to the occurrence of the other. An example could be the decision to purchase a car a few months after the birth of a baby in the household. Similarly, it may take some time after changing one's work/residential location to realize the need to have a car or an additional one. It would be therefore ideal to capture the changes over time of many life events with possible lagged effects in an integrated framework.

To the best of our knowledge, there is only one study (Guo et al., 2019) applying DBN to car ownership (which is the focus of the current work). Our work advances their work in various dimensions: 1) While the representation of car ownership in the Guo et al. (2019) was rather crude (either binary "Yes/ No" or "never changed, changed once, changed twice"), some useful detail to the levels such as differentiations between buy, sell and change as well as emission code, fuel and size of the vehicles were added. Such details allows the integration of knowledge about future prediction of car fleet as part of disaggregate travel demand forecasting models and emission forecast. 2) Guo et al. (2019) also considered residential mobility and work location mobility as binary variables of "Yes/ No". In the current study, if the residential and/or work location changes, the network is able to also predict the new location. This addition brings the current study much closer to a comprehensive dynamic simulation framework. 3) The data in their work was collected retrospectively and respondents were encouraged to remember all the life events that happened in the past. Memory biases could undermine the quality of responses. Our data is panel data in which the changes in life events were recorded each year. 4) Our sample size is larger, 1035 individuals (from 737 households) versus 414 individuals and 266 households in their work. 5) Attitude toward car ownership and use was not part of their work. 6) The study area differs. Their study has been conducted in China and our data comes

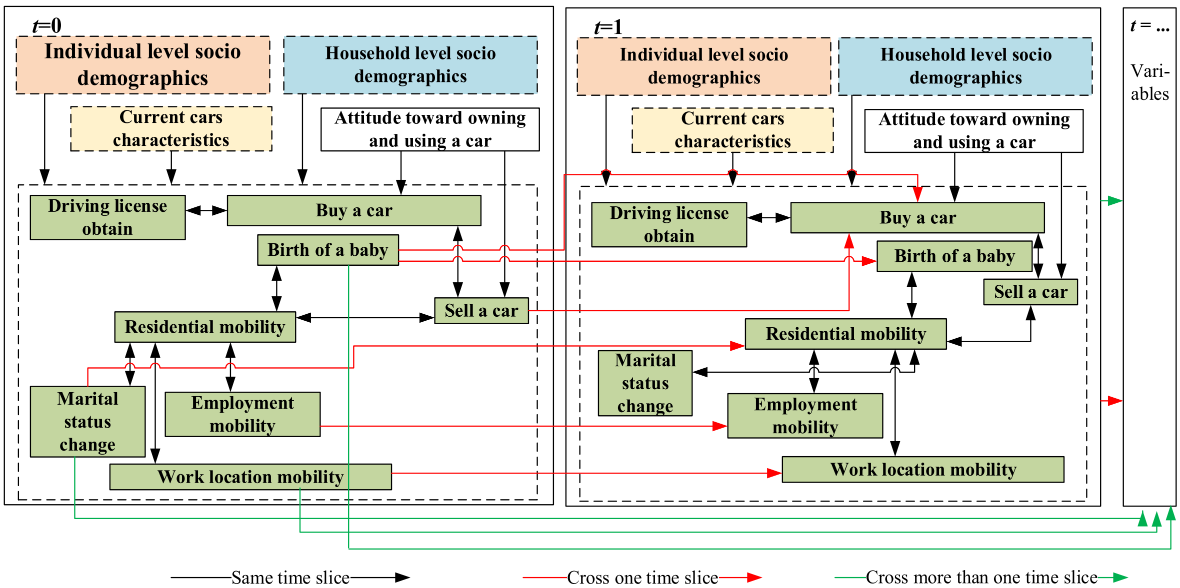

Fig. 1. Conceptual framework of Vehicle transaction model.

from the Dutch population.

3. Conceptual framework and method

This section discusses the conceptual framework for an integrated dynamic vehicle transaction model. Given the advantage in relating factors to each other in a dynamic process, the DBN method was used to integrate all the life events and socio-demographics across several time slices in one network. Section 3.1 explains the conceptual framework and Section 3.2 introduces the DBN method.

3.1. The conceptual framework

Based on previous studies (see Section 2) and available datasets (explained in Section 4), the socio-demographics were hypothesized to be potentially influential for life events while life events can impact each other within the same year or with some temporal lags. Latent attitude was also added to our framework since they were measured related to owning and using of car. Fig. 1 illustrates the framework. The strength and direction of the relations will be decided as part of learning process explained in the next section. The variables in different times slices, (e.g., $t = 0, t = 1, \ldots t = \mathrm{T}$) were encapsulated in separate large rectangles (one year is the unit of time in our study). The influences among variables could occur in the same or across several years. Intuitively, socio-demographics can not be influenced by other variables and incur an impact on other variables within the same year. Life events can have impacts on each other within the same time slice or cross several time slices. These interdependences within the same time slices represent by black bi-directional arrows in Fig. 1 (since the relation could have any of the two directions). The interdependences across several time slices were marked by red arrows in Fig. 1. The green arrows represent the influences of variables in time slice t on variables in further future t' ($t' > t+1$).

3.2. Dynamic Bayesian network (DBN)

In the current study, the DBN method was adopted to capture interdependencies between life events and socio-demographics and possible lagged effects. Below, a concise introduction of the DBN method is provided. For detailed theories and methods, the readers are referred to Heckerman et al. (1995), Pearl (1988), Mihajlovic and Petkovic (2001), and Uusitalo (2007). Bayesian Network (BN) is a directed acyclic graph (DAG). The DAG illustrates the joint probability distributions of all variables in the model. The joint probability distribution of a set of variables in a BN is represented as Eq. (1).

$$P(X_1, X_2, \ldots, X_n) = \prod_{i=1}^{n} P(X_i|\pi(X_i)), i = 1, 2, \ldots, n \tag{1}$$

where X_1, X_2, …, X_n represent variables in the model, each corresponding to one node. $\pi(X_i)$ represents the parent nodes of node X_i in the DAG. $P(X_i|\pi(X_i))$ represents the conditional probability of node X_i given its parents.

A BN represents the relationships among variables at one single time slice. The relationships over time and lead-lag effects can be captured by a dynamic Bayesian network (DBN). The DBN is an extension of the BN to model influences over time (Murphy, 2002). Fig. 2 schematically illustrates a DBN with 3 nodes, "A", "B", and "C" over T time slices. Within the same time slice, "A" influences "B" and "C", and "B" influences "C". Across time slices, "A" has an influence on itself in the following time slices and the next two time slices, "C" has an influence on itself in the following time slice, and "A" has an influence on "B" in the following time slice. The structure and parameters of the DAGs of different time slices can differ.

The joint probability of nodes across time slice(s) is represented as (Roos et al., 2017):

$$P(X^1, X^2, \ldots, X^T) = \prod_{t=1}^{T}\prod_{i=1}^{n} P(X_i^t|\pi(X_i^t)) \tag{2}$$

where $X^t = \{X_1^t, X_2^t, \ldots, X_n^t\}$, X_i^t represents node X_i at time slice t, and $\pi(X_i^t)$ represents parent nodes of X_i^t possibly from time slice 1 to t, $t = 1, 2, \ldots, T$ and $i = 1, 2, \ldots, n$.

Suppose G is a DBN model, $P_{prior}(G)$ represents the prior distribution of the DBN, and $P_{prior}(\Theta|G)$ represents the corresponding parameter set. Given observed data D and the model G, the parameter set Θ needs to be estimated. This estimation needs to maximize

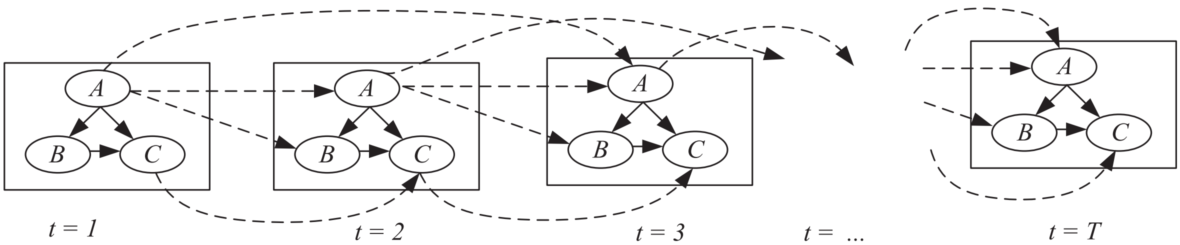

Fig. 2. Sketch of DBN (Mihajlovic and Petkovic, 2001).

Table 1
Descriptive statistics of the respondents' characteristics (percentage %)($N = 1035$, HH = 737).

Items	Levels	2013 (%)	2014 (%)	2015 (%)	2016 (%)	2017 (%)	2017 CBS(%)
Age*	18–24 years	4.06	3.57	3.19	2.51	1.84	10.82
	25–39 years	27.92	26.47	25.80	24.35	23.77	22.96
	40–69 years	56.23	56.91	56.52	56.71	55.65	50.55
	Older than 70 years	11.79	13.04	14.49	16.43	18.74	15.68
Gender*	Male	46.09	46.09	46.09	46.09	46.09	49.22
	Female	53.91	53.91	53.91	53.91	53.91	50.78
Employment status and working hours type*	full time employment	29.66	29.95	28.41	28.41	30.34	
	part time employment	28.31	24.73	24.93	25.31	23.00	
	unemployment	13.53	16.81	16.33	15.65	15.46	
	retired	25.02	24.83	27.25	28.02	29.37	
	student attending school	3.48	3.67	3.09	2.61	1.84	
Marital status*	In 1st marriage or partnership	61.45	62.22	61.93	62.71	62.71	
	In 2nd or 2nd more + marriage or partnership	0.48	0.00	0.39	0.10	0.39	
	single	38.07	37.78	37.68	37.20	36.91	
Highest completed education level*	Primary education level or no education experience	3.77	3.09	2.42	2.42	2.42	
	Secondary education level[a]	56.71	56.62	57.20	57.10	56.81	
	bachelor	26.47	26.76	26.28	26.28	26.18	
	Master	13.04	13.53	14.11	14.20	14.59	
Driver license status*	Yes, I have a drive license	88.79	89.57	89.47	90.14	90.53	80.2
	No, I do not have a driver license	11.21	10.43	10.53	9.86	9.47	19.8
Distance from home to work classification*	Within 10 km	75.27	78.74	79.52	80.10	79.13	
	10–20 km	9.95	7.92	7.63	7.15	7.83	
	20–30 km	4.44	4.83	4.54	4.44	4.44	
	≥30 km	10.34	8.50	8.31	8.31	8.60	
Household composition	singe	50.47	50.20	49.80	48.85	48.44	35.26
	couple	32.02	32.02	31.07	31.48	31.34	26.76
	Couple + children	14.93	15.33	16.42	16.69	17.23	30.79
	Single + children	2.58	2.31	2.44	2.99	2.99	6.69
	others	0.00	0.14	0.27	0.00	0.00	0.49
Number of children below 12	0 child	88.74	88.74	87.79	87.65	87.38	
	1 child	6.38	5.97	6.24	5.83	5.56	
	≥ 2 children	4.89	5.3	5.97	6.51	7.06	
Children 12 to 18	0 child	94.98	95.39	95.25	95.12	94.98	
	1 child	2.85	2.58	2.58	3.26	3.39	
	2 children	1.90	1.76	1.90	1.36	1.22	
	3 or more children	0.27	0.27	0.27	0.27	0.41	
Household Income	below the national benchmark income	28.90	28.63	56.17	50.88	57.12	57.69
	national benchmark income	28.09	28.63	21.85	20.35	21.98	29.88
	higher than the national benchmark income	43.01	42.74	23.74	30.80	22.93	12.44
Number of cars	0	22.25	22.39	21.98	28.77	20.90	
	1	54.27	55.50	56.17	50.88	57.12	
	≥ 2	23.47	22.12	21.85	20.35	21.98	
Age of the current main car	No car	22.25	22.39	21.98	28.77	20.90	
	(0,1990]	0.95	1.22	0.81	0.41	0.68	
	(1990–2000]	11.13	9.23	7.87	5.97	5.43	
	(2000–2004] Euro3	21.17	17.10	15.33	11.80	12.89	
	(2005–2009] Euro4	19.54	18.86	17.77	15.20	14.52	
	(2009–2014] Euro5	24.69	30.12	31.75	26.05	29.58	
	(2014∼) Euro6	0.27	1.09	4.48	11.80	16.01	
Fuel type of the current main car							
	No car	22.25	22.39	21.98	28.77	20.90	
	Gasoline	64.18	63.09	63.36	57.12	64.18	
	Diesel	11.40	11.13	10.85	10.58	10.31	
	LPG & CNG	0.00	0.14	0.14	0.14	1.09	
	Hybrid	0.14	0.41	0.41	0.27	0.14	
	Full electric	2.04	2.85	3.26	3.12	3.39	
Size of the current main car	No car	22.39	24.69	23.74	30.80	22.93	
	S (≤1305)	61.19	59.29	60.11	53.60	60.65	
	M (1305, 1760]	14.79	15.06	15.06	14.25	14.65	
	L (>1760)	1.63	0.95	1.09	1.36	1.76	
Urbanization level of residential place							
	very high-density areas with ≥2500 addresses per km2	24.97	25.24	42.61	42.20	41.79	
	high density areas with 1500–2500 addresses per km2	29.99	29.31	19.00	18.86	18.45	

(continued on next page)

Table 1 (*continued*)

Items	Levels	2013 (%)	2014 (%)	2015 (%)	2016 (%)	2017 (%)	2017 CBS(%)
	moderately high-density areas with 1000–2500 addressers per km2	21.71	22.39	10.04	10.18	10.58	
	low density areas with 500–1000 addresses per km2 and very low-density areas with <500 addresses per km2	23.34	23.07	27.14	27.54	27.95	
Age of mom	18–24 yeas	2.99	1.90	1.22	0.81	0.14	
	25–39 years	22.12	21.57	21.85	19.81	19.95	
	40–69 years	40.43	41.25	40.57	41.52	39.62	
	Older than 70 years	7.33	8.14	9.23	10.31	12.35	
	Do not exists	27.14	27.14	27.14	27.54	27.95	
Highest education level of mom	Primary education level or no education experience	1.36	1.22	1.22	1.22	1.22	
	Secondary education level	43.69	42.61	42.61	42.20	41.79	
	bachelor	18.59	19.40	19.00	18.86	18.45	
	Master	9.23	9.63	10.04	10.18	10.58	
	Does not exists	27.14	27.14	27.14	27.54	27.95	
Work distance of main earner	Within 10 km	73.27	75.58	25.80	24.35	23.77	
	10–20 km	10.04	8.96	56.52	56.71	55.65	
	20–30 km	4.88	5.16	14.49	16.43	18.74	
	≥30 km	11.80	10.31	46.09	46.09	46.09	
Age of main earner	18–24 years	7.19	5.70	53.91	53.91	53.91	
	25-39 years	26.05	25.78	28.41	28.41	30.34	
	40–69 years	53.73	54.27	24.93	25.31	23.00	
	Older than 70 years	13.03	14.25	16.33	15.65	15.46	

[a] Secondary education level main contain; LBO\VBO\VMBO(vocational educational programs), MAVO\HAVO-VWO\VMBO (junior years high school education), MBO (middle-level applied education), HAVO and VWO (senior high school year\university propaedeutic diploma.

[*] Refers to individual level characteristics.

the posterior probability of the model given data D (Mihajlovic and Petkovic, 2001):

$$P_{posterior}(G|D) = \frac{P_{prior}(G)}{P_{prior}(D)}\int_{\Theta} P_{prior}(D|\Theta, G)P_{prior}(\Theta|G)d\Theta \tag{3}$$

The parameter set Θ can be estimated using the maximum likelihood (ML) method employing the Expectation-Maximization (EM) algorithm. The maximum likelihood estimate of Θ_{ML} can then be obtained as below:

$$\Theta_{ML} = arg\ max_{\Theta} log P_{posterior}(D|\Theta) \tag{4}$$

A DBN is defined by its structure and parameters. There are two approaches constructing a DBN. The first is to learn only the parameters of the network while the structure of the network is given, using expert knowledge. However, expert knowledge may not capture weak relationships. Moreover, the expert knowledge is also based on historical data and it would be relevant to test and possibly update that knowledge when new data is available. The second approach is to learn the structure and parameters of the network using available datasets (Heckerman et al., 1995; Pearl, 1988; Sun and Erath, 2015). In this approach, the structure and the parameters of the network are learned simultaneously based on available observations. The disadvantage is that relying purely on datasets for learning the network structure may capture some unreasonable relationships. In our research, these two approaches were combined so that expert knowledge supplements the structure of DBN derived from the data. There are several structure learning algorithms such as "Bayesian Search", "Greedy Thick Thinning", "PC algorithm", "Naïve Bayes", "Augmented Naïve Bayes", "Tree Augmented Naïve Bayes" (Heckerman et al., 1995; Kwisthout, 2011; Mihajlovic and Petkovic, 2001; Murphy, 2002; Suzuki, 2017; Tsamardinos et al., 2006). The "PC algorithm" unlike many others can learn the structure with continuous data and infer the structure by observing the interdependences in data. The "Naïve Bayes", "Augmented Naïve Bayes" and "Tree Augmented Naïve Bayes" mainly focus on capturing the influence of one variable on the others, and thus the learned structure would look like a light bulb emitting many rays, with which the interaction among all variables could not be well obtained. After preliminary tests, the "Bayesian Search" appeared to have slightly higher goodness-of-fit than "Greedy Thick Thinning" in structure learning, using current datasets. Therefore, the Bayesian Search structure learning algorithm is used in the current research. In our network, there exist 8 life events and 21 socio-demographic variables. The time slices span across 5 years (2013–2017). Due to the complexity of our network and the fact that the size of conditional probability grows exponentially with increasing the number of parent nodes, the maximum number of parent nodes has been set to 8 nodes in the structure of BN.

4. Data

The data from the Netherlands Mobility Panel (Mobiliteitspanel Nederland, MPN) administrated by the KiM Netherlands Institute for Transport Policy Analysis was used to train the DBN. The MPN data traces life events as well as travel behavior of a panel of individuals and households in the Netherlands since 2013. The datasets used in the current work cover a period of six years (2013–2018) involving 7760, 13,028, 7827, 9293, 10,676, and 8561 individuals aged over 12 from 3572, 4685, 3125, 3439, 4802, and 4468 households, respectively. Changes in individual and household characteristics, life events at individual and household levels, travel behavior, and attitudes toward owning and using a car are included in this longitudinal survey (Hoogendoorn-Lanser et al., 2015).

After data cleaning, the sample used for the analysis consists of 737 households with complete information across six years. To extract life events such as “Marital status change”, “Birth of a baby”, and so on, the corresponding values in successive years were compared. If changes were observed, the relevant life events were assigned to the preceding year. From this perspective, 2018 acts as the reference year, and the datasets from 2013 to 2017 were used for further analysis.

Socio-demographic characteristics of the sample and their life events are reported in Tables 1 and 2, in the first column, the variables specifically marked with * denote the individual level variables. The individual-level socio-demographics include “Age”, “Gender”, “Employment status and working hours type”, “Marital status”, “Highest completed education level”, “Driver license status” and “Distance from home to work place”. Household-level socio-demographics include “Household composition”, “Number of children below 12 years old”, “Number of children between 12 and 18 years old”, “Household income”, “Number of cars”, “Urbanization level of residential place”, “Age of mom”, “Highest education level of mom”, “Work distance of the main earner” and “Age of the main earner”. The characteristics of the current cars include “Age of the current main car”, “Fuel type of the current main car” and “Size of the current main car”. The individual-level life events include “Employment mobility”, “Work location mobility”, “Driving license obtain” and “Marital status change”. The household-level life events include “Residential mobility”, “Birth of a baby”, “Sell a car” and “Buy a car”. The sample statistics were compared with the population characteristics (whenever available) reflected in the last columns of Tables 1 and 2. The comparison suggests that most of sample socio-demographic characteristics are close to the population with few exceptions. Youngest age group (18–24 years) and couple with children were underpresented in the sample while single households were over presented. The disparity was presumed related to the decision to keep the households which completed the survey for 6 consecutive years. As for the life events, “Birth of a baby”, “Marital status change” and “Driving license obtain” were specified dichotomous, and some other variables like “Age”, “Urban density” etc. have a rather clear classification (followed the classifications of the Netherlands Statistics). For the classifications of the “Work location mobility” and “Residential mobility”, a more sophisticated yet informative approach was adopted. The data (see Section 4 for the detail) provides information about the residential and working location at the subzone level. Since defining the nodes with all possible subzones (within the Netherlands) would lead to the explosion of categories and likely overfitting, the municipalities based on the population size were ranked according to the quantiles of the observed distribution into five classes. Since “municipality” is still rather a crude spatial unit, the subzones were further used for the spatial unit of “Work location mobility” and “Residential mobility”. Subzones within each municipality were then ranked based on population size first (largest to smallest), as “order 4”, “order 3”, “order 2” and “order 1”. The cut-off points were chosen to select the 4-quantiles. This approach leads to ten classes for “Work location mobility” and “Residential mobility”, i.e., 3 (moving to higher, same, and lower municipality order) * 3 (moving to higher, same, and lower subzone order) + 1 (“no moving”). Some levels were grouped due to the very low data points within those categories in our datasets. The final classification of “Work location mobility” after grouping was chosen as “no moving”, “same municipality order, same subzone order”, “higher municipality order, same subzone order”, “lower municipality order, same subzone order” and “other types”. The classification of “Residential mobility” was similar to that of “Work location mobility”.

The classification of the age of vehicles were defined in line with the European emission codes (European Vehicle Market Statistics 2016/17, 2016). Such classification makes the results readily applicable for further application in terms of emissions forecast in combination with disaggregate travel demand forecasting models. In terms of size, the vehicles were classified into Small (S, $\leq$ 1305 kg), Medium-sized (M, 1306 kg - 1760 kg) and Large (L, 1761 kg - 3500 kg) which is in line with European emission standards for light commercial vehicles. “Attitudes toward owning and using a car” (see details in Section 5) was added to the list of determinants as well.

The decisions “Buy a car” and “Sell a car” were used to represent the change of car ownership. “Replace a car” was imputed if selling and buying cars happen in a single year. The characteristics of purchased cars (age, fuel type, and size) and disposed of cars over the five years are presented in Tables 3 and 4 respectively. Buying car (“Buy a car”) occurs in 122, 122, 125, 198 and 181 of cases in the period of 2013–2017, while selling car (“Sell a car”) constitutes 101, 97, 148, 110 and 167 respectively. Based on the descriptive statistics, 101 households sold a car and 122 bought a car in 2013. Interestingly, 88 households sold and bought a car in the same year (replacing car). The age distribution of the cars signals the natural rejuvenation of the fleet reflected by the downward trend in the share of Euro 4 and older and the increase of Euro 5 and Euro 6 in the fleet composition. The proportion of purchasing full-electric vehicles shows a slight upward trend in the five years. In the sample, small and gasoline cars were the ones transacted most often due to the high share of them in the fleet. Electric vehicles, as expected, remained modest in the share of both purchased and disposed of cars with the peak of purchases (11) in 2017. Among the purchased cars, small-sized cars constantly take the highest proportion, followed by the medium-sized ones. The evolution of the proportion of purchased vehicles over time with different emission codes is intriguing. In 2013 and 2014, Euro 5 was leading; Whereas, from 2015 onward, the highest percentage of purchased vehicles fell in the category of Euro 6 emission code. Among the disposed of cars, gasoline cars accounted for the highest percentage, followed by diesel and electric cars. The small-sized cars also took the highest proportion in the disposed of the fleet, followed by medium-sized and large-sized cars, consistently over the 5 years.

Moving to other life events, the percentage of “Work location mobility” ranges from 4.06% to 18.84%, and the percentage of “Residential mobility” ranges from 3.39% to 6.11% across five years. As for the occurrence of other life events, listed in Table 2, the birth of a child (“Birth of a baby”) occurs 1.9% to 2.71% in our dataset while the rates of obtaining a driver’s license (“Driving license obtain”) range from 0.77% to 1.16% in these five years. The proportion of change of marital status (“Marital status change”) is not too high, evidenced by the percentages of getting married ranging from 0.48% to 0.77% and getting divorced ranging from 0.68% to 1.45%. From 2013 to 2017, the proportion of changes in the employment status (“Employment mobility”) ranges from 11.21% to

Table 2
Descriptive statistics of the respondents' life events (N = 1035, HH = 737).

Items	Levels	2013 (%)	2014 (%)	2015(%)	2016 (%)	2017 (%)	2017 CBS(%)
Marital status change	Get married or start a relationship	0.48	0.58	0.77	0.48	0.48	
	Get divorce or relationship ended	0.87	1.45	0.97	0.77	0.68	
	keep original status	98.65	97.97	98.26	98.74	98.84	
Employment mobility*	yes, full time employment	2.90	2.03	2.13	3.19	1.74	
	yes, part time employment	3.77	4.54	3.96	3.77	3.67	
	yes, unemployment	4.25	2.71	2.71	2.90	3.57	
	yes, retired or taken early retirement	1.55	3.19	2.32	3.00	2.90	
	yes, student attending school	0.97	0.29	0.10	0.29	0.19	
	No	86.57	87.25	88.79	86.86	87.92	
Driving license obtain*	Yes	0.97	0.97	1.16	0.77	0.29	
	no	99.03	99.03	98.84	99.23	99.71	
Work location mobility*	no moving	93.82	95.17	95.94	82.32	81.16	
	same municipality order same subzone order	3.57	2.71	2.22	4.25	4.44	
	higher municipality order same subzone order	1.55	0.87	0.87	5.31	8.41	
	lower municipality order same subzone order	1.06	1.26	0.97	8.12	5.99	
	other types move	0.00	0.00	0.00	0.00	0.00	
Residential mobility	no moving	96.61	95.79	93.89	95.39	96.34	
	same municipality order same subzone order	0.81	1.22	0.95	0.68	0.68	
	same municipality order higher subzone order	0.00	0.68	1.36	0.95	0.95	
	same municipality order lower subzone order	0.95	0.14	1.22	1.22	0.68	
	lower municipality order higher subzone order	0.54	0.27	0.54	0.54	0.54	
	other types move	1.09	1.90	2.04	1.22	0.81	
Birth of a baby	No	98.10	97.42	97.29	97.83	97.96	99.01
	Yes	1.90	2.58	2.71	2.17	2.04	0.99

* Refers to individual level characteristics.

Table 3
Number of age/fuel type/size of purchased cars.

	2013	2014	2015	2016	2017
Fuel type of purchased cars					
gasoline car	86	97	92	153	139
Diesel car	24	19	25	29	28
LPG&CNG car	1	0	1	7	2
Hybrid car	2	1	2	1	1
Full electric car	9	5	5	8	11
Size of purchased cars					
S	92	92	88	143	130
M	29	26	32	41	44
L	1	4	5	14	7
Age of purchased cars					
<1990	6	0	0	4	1
1990–2000	9	8	10	14	10
Euro 3	10	14	15	31	16
Euro 4	28	22	14	27	23
Euro 5	55	42	28	63	54
Euro 6	14	36	58	59	77

13.43%. Whenever national aggregate statistics are available, they are also listed in Tables 1 and 2 for comparisons.

5. Identification of latent variables (latent class clustering)

The individual attitude was included in the DBN to check whether attitude toward car ownership in general and private car use in particular significantly impact the decision to purchase or sell cars over time. The latent class clustering method was used to classify individuals according to their attitudes. The statements to measure car use attitude are listed at the bottom of Table 5. All statements were measured on a 5-points Likert scale (i.e., 1: strongly disagree, 2: disagree, 3: neutral, 4: agree, and 5: strongly agree). The statements are listed in Table 5. Noticeable is that the responses (to attitude statements) were not constant over the five years. Table A.1 reports the results of the paired-sample *t*-test for the comparison of responses in different years. It suggests that the largest difference between the item responses occurs between 2014 and 2018, in which the test statistics were significant for the majority of the responses. Explanatory factor analysis (EFA) was conducted to test the reliability and consistencies of the factors underlying the

Table 4
Number of age/fuel type/size of disposed of cars.

	2013	2014	2015	2016	2017
Fuel type of disposed of cars					
gasoline car	80	77	119	80	116
Diesel car	18	17	20	22	37
LPG&CNG car	0	0	0	0	4
Hybrid car	0	1	3	2	0
Full electric car	3	2	6	6	10
Size of disposed of cars					
S	81	73	113	74	107
M	17	11	33	31	52
L	3	2	2	5	8
Age of disposed of cars					
<1990	4	0	1	0	1
1990–2000	22	15	23	16	16
Euro 3	30	23	38	18	34
Euro 4	26	22	26	28	26
Euro 5	19	30	56	35	62
Euro 6	0	7	4	13	28

Table 5
Comparison of models with different numbers of classes.

	Classes	Log-likelihood	BIC	NPar	df	*p*-value	Percentage of minimum class
Model1	1	−63,625.1	127,673.6	52	3389	0.000	100
Model2	2	−61,315.5	123,258.0	77	3364	0.000	33.59
Model3	3	−59,539.8	119,910.3	102	3339	0.000	25.4
Model4	4	−58,773.4	118,581.1	127	3314	0.000	16.13
Model5	5	−58,009.8	117,257.4	152	3289	0.000	11.03
Model6	6	−57,434.1	116,309.6	177	3264	0.000	9.61
Model7	7	−57,020.5	115,686.1	202	3239	0.000	8.72
Model8	8	−56,706.9	115,262.5	227	3214	0.000	7.98
Model9	9	−56,420.7	114,893.6	252	3189	0.000	5.16
Model10	10	−56,219.5	114,694.7	277	3164	0.000	5.42
Model11	11	−56,022.8	114,504.9	302	3139	0.000	4.63
Model12	12	−55,870.9	114,404.7	327	3114	0.000	3.69
Model13	13	−55,697.8	114,262.1	352	3089	0.000	3.44
Model14	14	−55,568.3	114,206.7	377	3064	0.000	3.38
Model15	15	−55,413.9	114,101.4	402	3039	0.000	3.34

measured items (reported in Table A.2). Eigenvalue which shows variance explained by that particular factor out of the total variance was used to determine the number of factors. Those with values >1 were used in the subsequent analysis. That led to the selection of four factors. Next, latent class clustering was conducted to extract latent clusters for inclusion in the DBN. While for the detailed theory and method, the reader is referred to Goodman (1974), McCutcheon (1987), a concise explanation of the method is presented below (Vermunt, 2010).

Let K presents the number of statements ($k = 1, 2, \ldots, K$), while the response of subject s ($s = 1, 2, \ldots, S$) on item k is denoted by Y_{sk}, the full response vector by Y_s, and the discrete latent classes by m ($m = 1, 2, \ldots, M$). Latent class is denoted by the covariates q ($q = 1, 2, \ldots, Q$), while Z_{sq} represents one of q covariates, related to subject s. Z_s represents the full vector of covariates.

$$P(Y_s|Z_s) = \sum_{m=1}^{M} P(m|Z_s)P(Y_s|m) \tag{5}$$

$$P(m|Z_s) = \frac{exp\left(\gamma_{0m} + \sum_{q=1}^{Q} \gamma_{qm} Z_{sq}\right)}{\sum_{m=1}^{M} exp\left(\gamma_{0m'} + \sum_{q=1}^{Q} \gamma_{qm} Z_{sq}\right)} \tag{6}$$

The parameter vector γ contains γ_{0m}, γ_{qm}, and $\gamma_{0m'}$, and the multinomial parameters defining $P(Y_s \mid m)$ are estimated by maximizing the log-likelihood function.

$$\sum_{s=1}^{S} logP(Y_s|Z_s) = \sum_{s=1}^{S} \sum_{m=1}^{M} P(m|Z_s)P(Y_s|m) \tag{7}$$

To determine the appropriate number of classes, models with different numbers of classes (from 1 to 15) were estimated. The models' fitness including their log-likelihood and Bayesian Information Criterion (BIC) are listed in Table 5. "NPar" is the number of parameters in each model, and "percentage of minimum class" represents the size of the smallest class in the corresponding model. Although the BIC decreases as the number of classes increases, the appropriate number of classes was eventually specified by setting the minimum class size as 15% (Ton et al., 2020). With this restriction, the model with four classes was selected.

Estimation results are listed in Table 6. Cluster 1 is the largest (43.87%) and labeled as "Financially limited and not thrilled users". People in cluster 1 would limit their car drive mainly because of the costs and environmental awareness. They only drive when it is necessary and postpone buying a new car due to financial constraints. They do not think owning a car is associated with one's sense of style. They do not acknowledge the advantages brought by driving a car compared to other transport modes either. Being a female, aged between 18 and 24, owning no car, belonging to the lowest household income level (below 27,000 €/year), and having a low education (primary) level increase the probability of belonging to cluster 1.

Cluster 2 is labeled as "Pragmatic car users". People in cluster 2 do not limit their car use because of the associated costs; they do not postpone purchasing a new car because of financial reasons. They neither acknowledge the social status and advantages brought by owning a car nor enjoy the experiences of using cars; they are not much dependent on using cars in their daily lives and only use the car when necessary. They consciously try to drive less because of environmental concerns. Being a female, aged 40 and more, belonging to the high-income group (>40,000 €/year), owning no car, and being highly educated (bachelor or above) increase the probability of belonging to this cluster.

Cluster 3 is labeled as "Middle-class car freaks". People in cluster 3 likely afford the costs of car use, and they would not drive car less because of concerns over environmental impact. They disagree with the idea that one's car says a lot about his/her style and status. They simply enjoy driving and acknowledge the advantages brought by cars compared to other transport modes. They have a strong reliance on car. Being a male, age between 25 and 39 years old, having two or more cars, belonging to the average income class (>27,000 €/year and <40,000 €/year) increase the likelihood of belonging to cluster 3. Being highly educated decreases the probability of belonging to this cluster.

Cluster 4 is labeled as "Finically restricted car lovers". People in cluster 4 drive less because of the cost and to some extent environmental concerns. They postpone buying a new car because of their financial limitations. They believe that owning a car in general and the type of the owned car in particular contribute to their social status. They believe travelling by car is comfortable, relaxing, and pleasant. They agree with the opinion that driving has more advantages compared to other modes of transport, and they strongly rely on car use in daily life. Being a male, aged between 25 and 39 or older than 70, belonging to the lowest income group (below 27,000 €/year), and having a secondary level education diploma increase the probability of being part of cluster 4.

These clusters were further used in the DBN as "Attitudes toward owning and using a car". The changes over years are marked if respondents move from one class to another.

6. DBN model learning, interpretation, and validation

In the DBN model, the vehicle transaction is captured by "Buy a car" and "Sell a car", while replacing a car is calculated indirectly when two incidents happen in the same year. Section 6.1 describes the detailed model formulation and estimation. Section 6.2 illustrates the learned parameters of the DBN and how "Buy a car" and "Sell a car" are influenced by other life events and socio-demographics. Section 6.3 reports the performance of the model by displaying the prediction accuracy of each life event with a special focus on the accuracy of age/fuel type/size of the purchased car.

6.1. DBN structure

Software "GeNIe"[1] was used to learn the structure and parameters of the DBN. The initial network structure was learned and illustrated in Fig. 3.

From Fig. 3, it is realized that "Sell a car" and "Birth of a baby" in year t influence "Buy a car" in year $t + 1$. It suggests that the purchase of cars has a significant association with "Birth of a baby" in the households and "Sell a car" although with one year delay. "Sell a car" has a concurrent impact on "Buy a car" as well. The concurrent and 1 year lagged influence from "Sell a car" on "Buy a car" have never been captured before. "Marital status change" has an impact on "Buy a car" four years later, while obtaining a driving license has a concurrent influence on "Buy a car", suggesting that people mainly buy a car in the same year they obtain the driving license. Some of the above relations were confirmed in previous studies (Clark et al., 2016; Oakil et al., 2016) except that the lag effect between "Marital status change" and "Buy a car" had not been reported. The "Number of children age 12 to 18" also showed effects on "Buy a car", which is a rather exceptional observation compared to the literature as most researchers reported the event of "Birth of a baby" as an influencing factor to purchase a car other than the "Number of children age 12 to 18" in the household. It is understandable realizing that a driving license can be obtained from the age of 17 in the Netherlands. The "Size of the current main car" has an influence on "Buy a car" in the same year, which has not been considered in prior research. Different from the previous research (Flamm,

[1] Thanks for the support of GeNIe software: https://support.bayesfusion.com/docs/GeNIe/

Table 6
The parameters of latent class - 4 clusters.

	Financially limited and not thrilled users	Pragmatic car users	Middle-class car freaks	Finically restricted car lovers
I_1 Due to high costs, I drive less with the car than I actually want to	0.6189	−0.6718	−0.9167	0.9696
I_2 Due to costs, it is difficult for me to own a car	0.7986	−1.55	−1.5385	2.2899
I_3 Due to costs, I opt to travel by public transport and bicycle instead of by car	0.6391	−0.0082	−1.0171	0.3863
I_4 With the environment in mind, in the past year I have consciously tried to drive a car less	0.1914	0.3735	−0.6376	0.0726
I_5 I only use a car if it is really necessary	0.055	0.5805	−0.5914	−0.0441
I_6 My current financial situation is a reason to postpone the purchase of a (new) car	0.3512	−1.103	−0.6833	1.4351
I_7 A car says a lot about someone's personal taste / sense of style	−0.4124	−0.8734	−0.6418	1.9276
I_8 A car says a lot about a person's status in society	0.1481	−0.3987	−0.1387	0.3893
I_9 I find travelling by car to be comfortable	−1.494	−1.1408	0.9566	1.6782
I_{10} Travelling by car is pleasurable	−1.1561	−1.1279	1.1921	1.092
I_{11} I find travelling by car to be relaxing	−0.756	−0.8433	0.7441	0.8551
I_{12} If I have to go somewhere, I nearly always go by car	−0.5061	−0.8041	0.4356	0.8746
I_{13} Driving a car offers many advantages compared to the use of other transport modes	−1.0195	−1.002	0.7185	1.303
Model for Clusters				
Cluster Size	0.4387	0.2082	0.1918	0.1613
Covariates				
Gender				
Male	−0.0994	−0.0507	0.0906	0.0595
Female	0.0994	0.0507	−0.0906	−0.0595
Number of cars				
0 car	0.7824	0.1816	−0.4355	−0.5286
1 car	−0.3319	0.1013	0.0377	0.1929
$\geq$ 2 cars	−0.4506	−0.2829	0.3978	0.3357
Household Income				
below the national benchmark income (<27,000 euros/year)	0.3283	−0.4483	−0.2159	0.3358
national benchmark income (27,000–40,000 euros/year)	0.0316	0.0192	0.1559	−0.2067
higher than the national benchmark income (>40,000 euros)	−0.3600	0.4291	0.0600	−0.1291
Age				
18–24	0.9646	−0.3892	−0.4570	−0.1184
25–39	−0.5822	−0.1915	0.6886	0.0851
40–69	−0.3529	0.2955	0.1108	−0.0534
$\geq$ 70	−0.0295	0.2852	−0.3425	0.0867
Highest completed education level				
Below primary education level	0.5168	−0.2742	0.1171	−0.3597
Secondary education level	−0.0490	−0.2772	0.1039	0.2223
bachelor	−0.3181	0.1306	0.0634	0.1241
Master	−0.1497	0.4208	−0.2844	0.0133

2009; Kim and Mokhtarian, 2018; Maat and Kasraian, 2020), in our study "Attitudes towards owning and using a car" was not captured in the initial structure learning of DBN, but was later added to the network as part of the expert knowledge.

"Number of cars", "Fuel type of the current main car", "Size of the current main car", "Age of the current main car" and "Residential mobility" appeared as significant links for the "Sell a car" decision after the structure learning. These relationships have rarely been studied, particularly the impact that "Residential mobility" can have on "Selling car". Researchers have proposed that "Residential mobility" has a significant influence on "Buy a car" (Gu et al., 2021; Guo et al., 2020b).

Moving to other life events, "Birth of a baby" has an influence on itself in the following 1, 2, and 3 years. This learned temporal relationship is consistent with previous research indicating that the birth interval plays an important role and should be considered in modeling dynamic synthesizing at the disaggregate level (Knodel and Hermalin, 1984). "Marital status change" appears to have an impact on "Birth of a baby" 2 years later. Although studies (e.g., Hong, 2006) reported the education level of mother as an important determinant of giving birth, our network did not pick this relation, which was later added as the expert knowledge. "Residential mobility" has an impact on "Birth of a baby" in the same year and 1 year later. This means people are more likely to change the place of residence in the same year or the year before they have a child. As for "Marital status change", DBN shows impacts from the number of

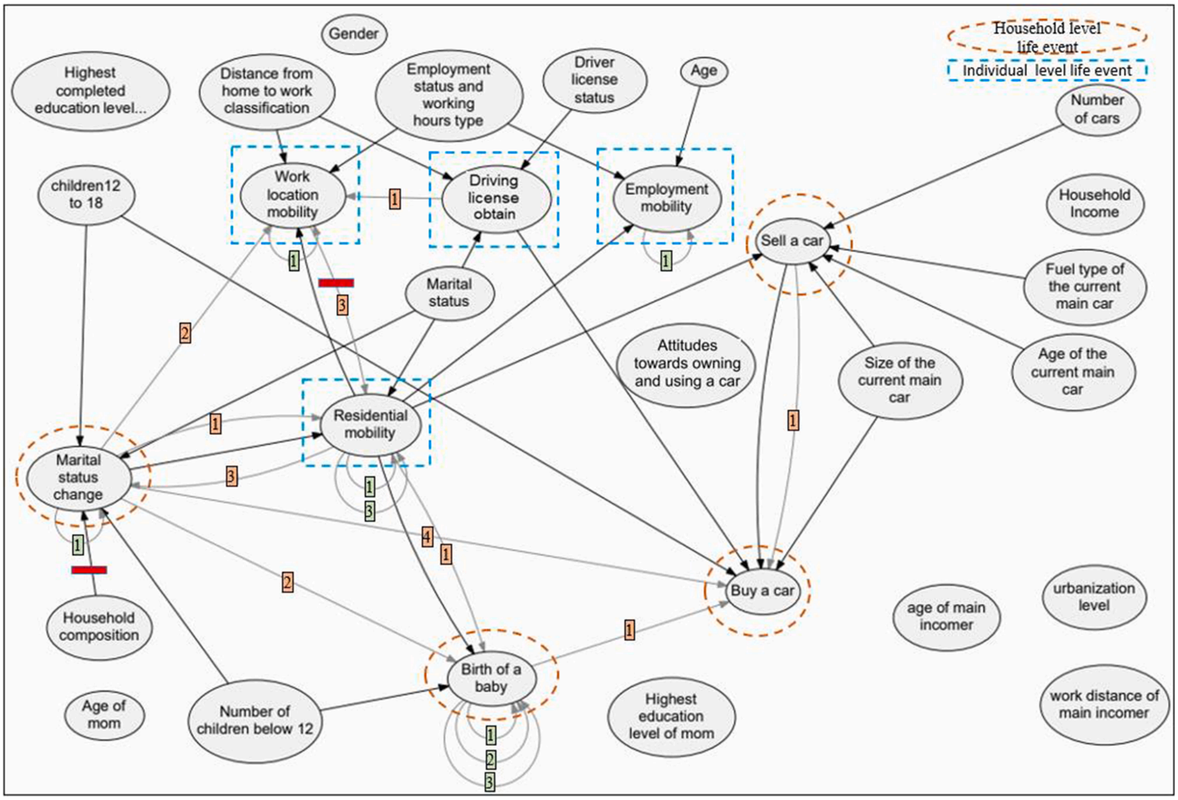

Fig. 3. Structure learning of DBN (log(p) = −74,988.15).
(Considering 1 year as the time interval, the number (1, 2, 3, …) in the green textbox means that the variable in year t has an influence on itself after "1, 2, 3, …" year(s). The number in the orange textbox (1, 2, 3, …) implies that the variable at the arrow tail in year t has an influence on the variable at the arrow head after "1, 2, 3, …" year(s). The arrows without any numbers suggest the impacts within the same year. The large red "minus" sign refers to the links that have been removed (from the literature review).)

children in different age groups in the same year, while residential mobility influences "Marital status change" 3 years later. These relationships are consistent with previous work (Saadeh et al., 2013). "Marital status change" influences "Residential mobility" in the same year and 1 year later. "Work location mobility" impacts "Residential mobility" 3 years later. Fig. 3 also suggests that "Residential mobility" in year *t* has an influence on itself in year $t+1$ and year $t+3$, likely pointing to the fact that people do not change home within such a small time interval; "Employment mobility", "Work location mobility" and "Marital status change" in year *t* influence on themselves in year $t+1$. "Distance from home to work" and "Marital status" show significant association with the "Driving license obtain".

It is worth noting that some of the life events (car transactions, residential mobility) are household decisions. In double-head households, the person with the higher income is considered as the main earner and the other head as partner. In the exploration phase of structure learning, the characteristics of both heads (main earner and partner) had been included (e.g., working distance of the main earner, work distance of the partner, employment mobility of main earner, employment mobility of partner on residential move and buying / selling cars.) The constructed network however suggested that the characteristics of the partner were not influential on the household-level life events such as residential mobility and buy / sell car and thus did not appear in the network and there was not included in the parameter learning. The assumption of the main earner influencing household decisions has been applied in some previous research as well (Browning et al., 1994; Posel, 2001).

For the individual life events such as work location mobility and employment mobility, the characteristics of each individual are obviously included as they are individual level life event. That being said, it is important to note that, for vehicle transaction, the "main car" was included in our network (if there are more than one) and it could very well be that in many cases the main car is mainly used by the main earner. As shown in Fig. 4, for "Marital status change", "Number of children between 12 and 18" and "Number of children below 12" were included as the parent nodes.

After the initial structure learning, the adjustments of the DBN structure based on expert knowledge were done. According to the expert knowledge (from previous researches, iterative preliminary tests, and our conceptual framework), if some expected links were missing, these potential links would be added to the DBN. The link was being kept if the fitness of the model improves and otherwise being removed. More specifically, the following links were added as expert knowledge: "Age", "Gender", "urbanization level of residential place" as the parent nodes of "Driving license obtain" since these factors show significant influence on "driving license obtain" decision in some research (Hjorthol, 2016); “Gender”, “Highest completed education level” and “Household composition” as the parent nodes of “Employment mobility”, previous research also indicated these influences (van Ham et al., 2001); “Age” as the parent node of “Marital status change”, “Highest education level of mom” and “Age of mom” as the parent nodes of “Birth of a baby” (e.g., Jain, 1981; Knodel and Hermalin, 1984; Weinberger, 1987; Testa, 2014; Awad and Yussof, 2017); “Age of main earner” and “work distance of main earner” as the parent nodes of “Residential mobility” (Clark and Withers, 2002). “Attitudes towards owning and using a car” as the parent node of “Sell a car” and “Buy a car” (Flamm, 2009; Kim and Mokhtarian, 2018; Maat and Kasraian, 2020). In addition, the link from “Work location mobility” to “Residential mobility” three years later was removed since the action increases the goodness-of-fit and decreases the complexity of the network. The link from “Household composition” to “Marital status change” was also removed since the parent node of “Marital status change” contains “Marital status”, “Number of children below 12” and “Number of children age 12 to 18”. To some extent, “Household composition” is collinear with these three nodes. Although “Household income” is an important socio-demographic factor, unlike previous studies, it did not show any paramount relationships with other variables in the structure learning. During adding the expert knowledge, “Household income” was tested as the parent node of some other variables, but did not improve the network performance. That led to “Household income” becoming an isolated node and did not appear in Fig. 4. The final DBN structure has $\log(p) = -72{,}249.58$, which has been improved significantly compared to the structure without expert knowledge ($\log(p) = -74{,}988.15$). The final DBN structure is shown in Fig. 4. A larger view of each life event with its parent and child nodes is shown in the appendix (Fig. A.2 to Fig. A.9).

With this learned and optimized DBN structure, an optimization algorithm (EM) was applied to learn the parameters.

6.2. *Interpretation of the network for “Buy a car” and “Sell a car”*

To better understand the extent of the impacts of each variable on decisions to “Buy a car” and “Sell a car”, we instantiated certain levels of parent nodes for these two life events and inferred the posterior probability distribution of “Buy a car” and “Sell a car”. Figs. 5 - 9 report the posterior probability distribution of “Buy a car” based on certain inferences. Fig. 5 shows the probability of buying various Euro-standard, fuel type and size of cars (the top part of the figure) when different attitude components (lower part of the figure) were instantiated. The first column corresponds to the observed data while the second column refers to the circumstance when everyone is “Financially limited and not thrilled users”. The remaining columns can be interpreted similarly.

As shown in Fig. 5, belonging to the “Financially limited and not thrilled users” class (indicating by 100% in front of this class) leads

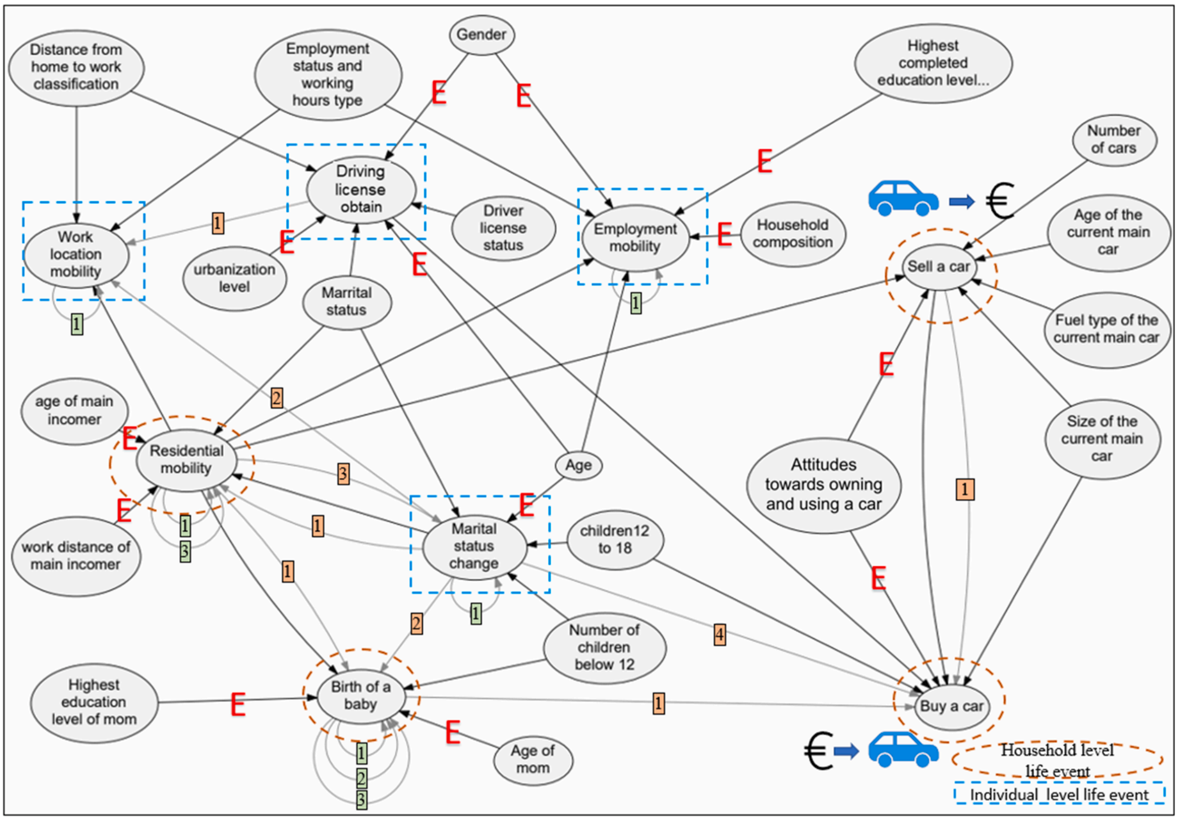

Fig. 4. Modified structure learning of DBN including expert knowledge ($\log(p) = -72{,}249.58$).
(Large red “E” specifies the links that were added as expert knowledge from the literature).

to a higher probability of not buying a car (57% compared to 51% in the sample). Belonging to the "Middle-class car freaks" class increases the probability of buying car (44% "Not buying" versus 51% in the sample) in general and electric (4% instead of 3%) and gasoline cars (26% instead of 23%) in particular. This particular class is also associated with a higher probability of buying small-sized cars. As for the emission code, while "Middle-class car freaks" and "Finically restricted car lovers" are more likely to buy Euro 6 (the best emission code) (6%) than the other two classes of people (5%). Class "Pragmatic car users" has a higher probability of buying Euro 5 (18%). On the contrary "Financially limited and not thrilled users" and "Finically restricted car lovers" groups are associated with lower probabilities of choosing electric, small-sized, and Euro 5 cars (compared to other groups). "Pragmatic car users" are more likely to choose diesel cars than the other three classes.

Comparing the share of "No buying" before and after instantiating the "Number of children between 12 and 18 years old", shown in Fig. 6, it is concluded that increasing number of children in this age group would increase the probability of buying a car (evidenced by decreasing the probability of "Not buying" from 53% when there is no children to 22% when there is 2 or more. Unlike the "Attitudes towards owning and using a car", which has a noticeable influence on the probability of buying an electric car, the number of children, as expected, does not play a significant role here. That is evidenced by stable probability of buying electric cars (2% - 3%) regardless of number of children. Instead, number of children has a strong association with the size of the cars purchased by households. In that sense, having more children increases the probability of buying cars of all sizes with the most increase for large cars (23%). The probability of choosing a small-sized car decreases from 32% to 25% when the number of children aged 12 to 18 increases from 1 to 2 in the household. The relation between the number of children and the emission code of purchased car, despite not as noticeable as for the size of cars, suggests that belonging to family groups with one child aged 12 to 18 is associated with a higher probability of purchasing Euro 4 (15%) (compared to the base (10%)). Families with two children aged between 12 and 18 have a noticeable higher probability of purchasing cars older than 1990 (14%) than other families.

As shown in Fig. 7, the structure learning of DBN recognized that "Sell a car" has an influence on "Buy a car" in the same year and the following year. If there is no car sold in that same year, as expected, the probability of not buying a car is much higher (81%) compared to only 10% when a car is sold in the same year. In addition, when no car is sold in the same year, the shares of purchased medium and large cars are the same (5%), which are half of the probability of purchasing small cars (10%), while when a car is sold in the same year, the probability of purchasing medium-sized purchased cars is >2 times of the large cars(27% vs. 13%). The share of purchased small cars in this case is almost two times that of medium cars (50% vs. 27%) and almost four times of large cars (13%). Changes in the number of cars purchased with different Euro codes when a car is sold in the same year (compared to when no car is sold) are noticeable. If the households do not sell cars in the same year as they purchase cars, the purchased cars across various Euro codes have very similar probabilities (2% - 4%). On the contrary, when households sell cars in the same year, Euro 5, as the emission code of purchased cars, becomes far ahead of the other emission codes (27%), followed by Euro 4 (19%). As shown in Fig. 4, "Buy a car" is influenced by "Sell a car" in the same year and in previous year; therefore, "Sell a car" decision in year T and T-1 were instantiated. The last column instantiates the situation when cars are sold in the previous year but not in the same year. Interesting is the comparison with the situation when selling and buying happens in the same year (column 3). While selling a car in the same year increases the probability of buying a car substantially (reflected in 10% of "no buying"), this is not the case if the car is sold the year before evidenced

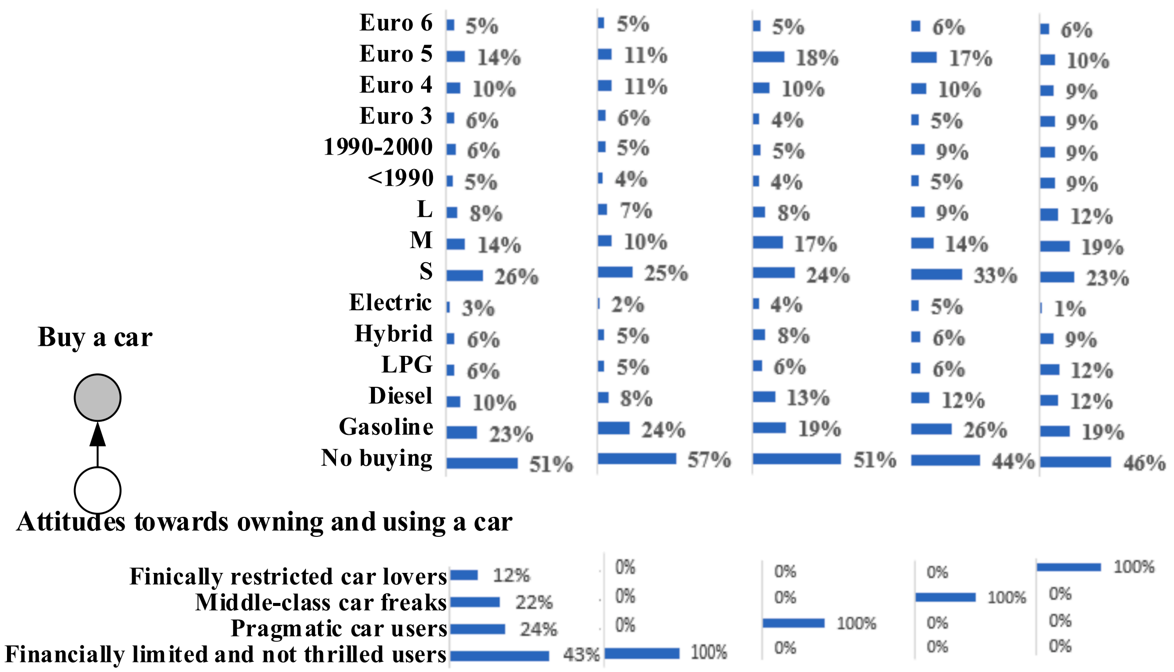

Fig. 5. Posterior probability distribution instantiating "Attitudes towards owning and using a car".

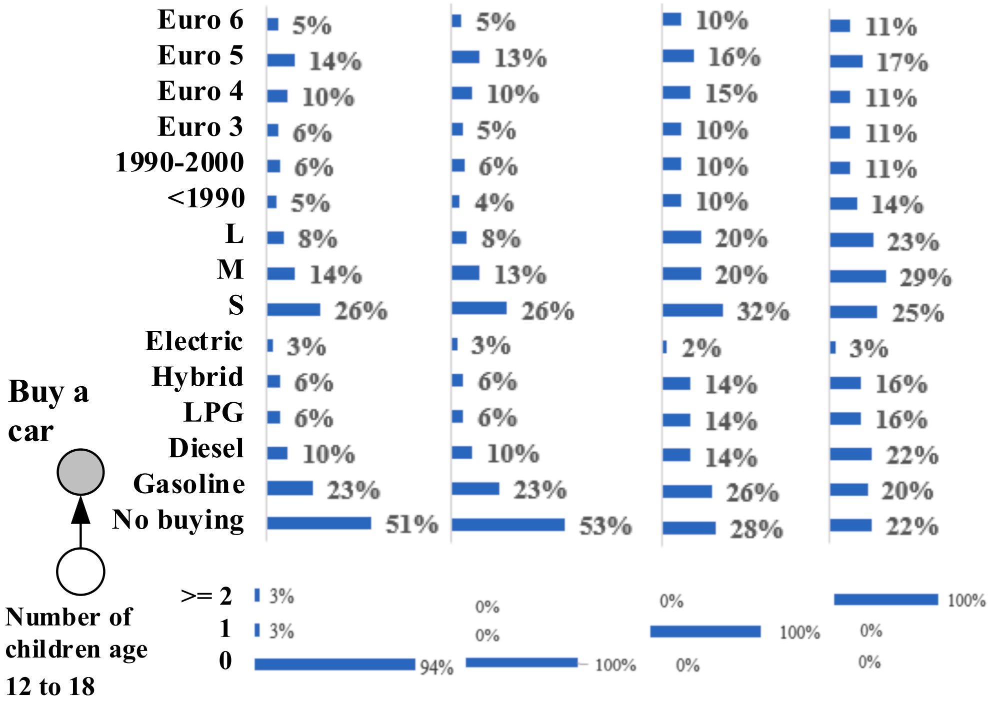

Fig. 6. Posterior distribution initiating "Number of children age 12 to 18".

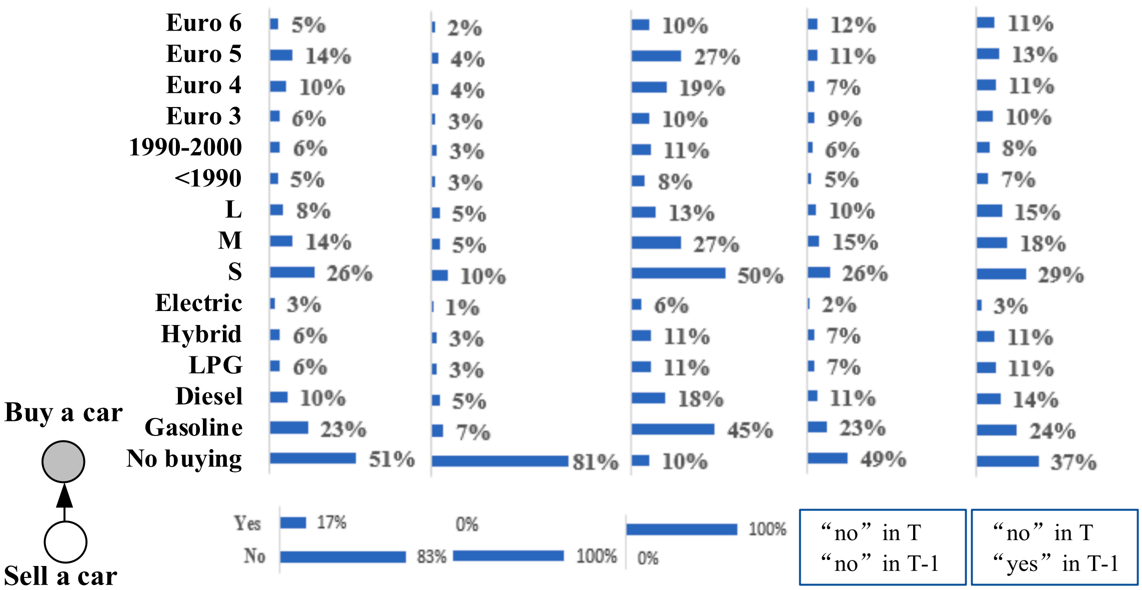

Fig. 7. Posterior distribution initiating "Sell a car".

by 37% of "No buying" in this case. That further means it is much more likely people replace their car by selling and buying within the same year. When people buy a car one year after selling their cars, although still Euro 5 constitutes the highest share of the purchased cars (13%), Euro 6 (11%) competes closely with Euro 4 (11%) for the second place. As for the size of purchased cars, if replacing a car happens with one year delay, the share of purchased small and medium-sized cars decreases substantially (29% and 18%) relative to the situation when households sell and buy a car in the same year (50% and 27%). The share of purchased large cars marginally increases (15% vs. 13%).

The influence of "Birth of a baby" in year t on "Buy a car" in year $t+1$ was captured during the structure learning as illustrated in Fig. 4, suggesting a lagged effect of one year. Fig. 8 displays the results of instantiating these life events. If households welcome their

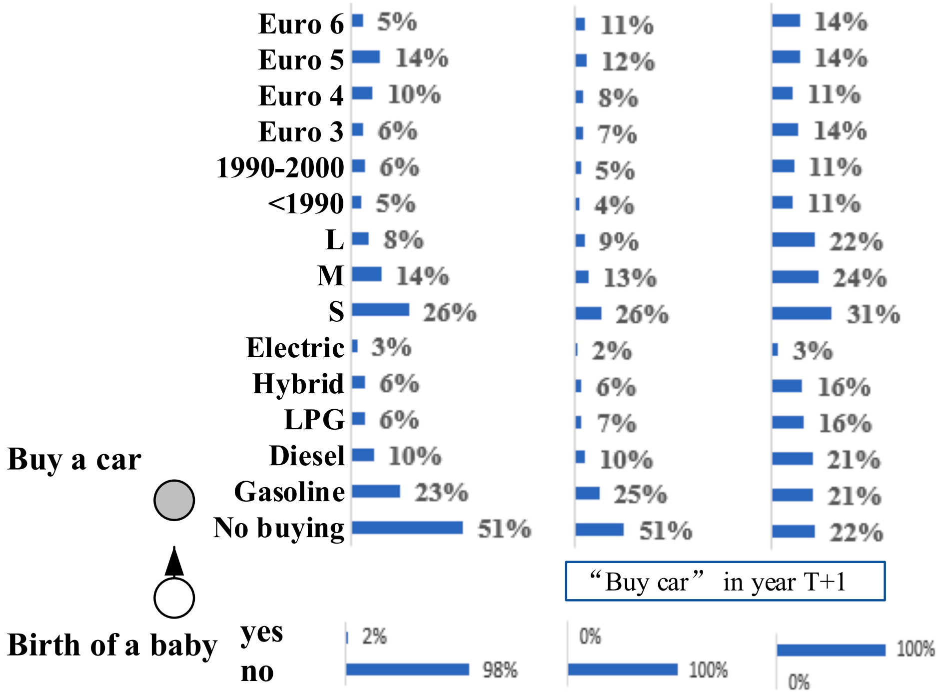

Fig. 8. Posterior distribution initiating "Birth of a baby".

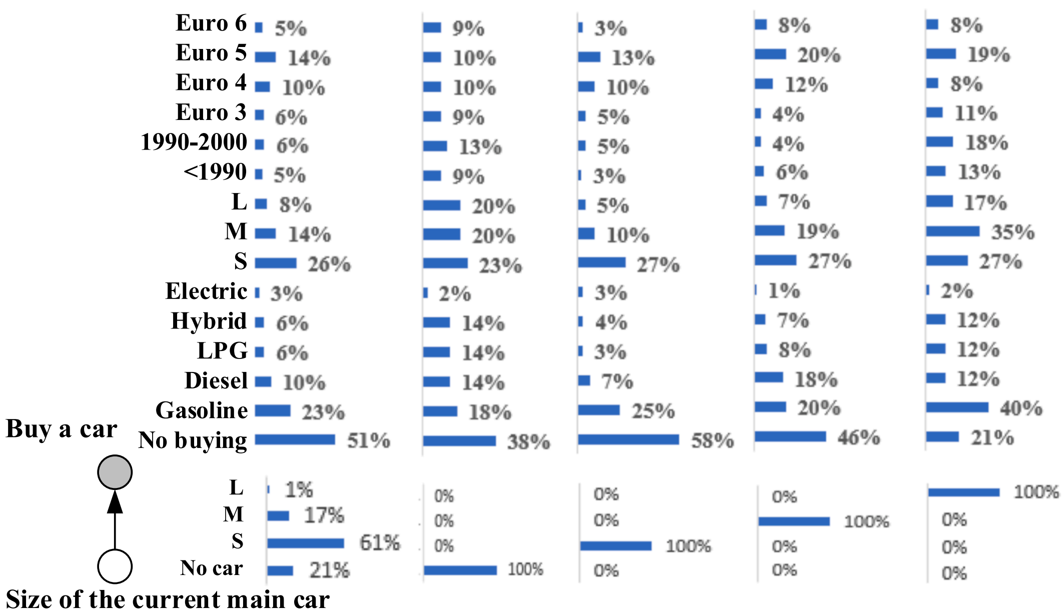

Fig. 9. Posterior distribution initiating "Size of the current main car".

babies in the previous year, they are more likely to purchase a car reflected in a much lower (22%) share of no incident of purchasing cars than the situation when no baby was born (51%). The probability of buying a large-sized car increases more than the other size classes when a baby is born (22%). When a car is purchased without having a newborn baby (column 2), it is observed that Euro 5 and 6 are the emission codes with higher share in the purchased fleet (12% and 11%). However, when cars are purchased a year after a baby is born, the share of purchased cars across emission codes is more even. That may point to the fact that in the latter case, purchasing a car is a necessity and thus financial restriction (which could be more intense after a baby is born) governs the choice of vehicle in terms of price (Emission code has a high correlation with price as it is an indication of the age of the car).

The "Size of the current main car" has a concurrent influence on "Buy a car". Fig. 9 gives detailed information about the magnitude

Table 7
The conditional probability of "Sell a car" after the parent nodes are instantiated.

		Number of cars		Residential mobility (city order + subzone order)						
Sell a car	***Original***	1 car	$\geq$ 2 cars	no moving	same & same	same &higher	same &lower	lower & higher	other move	
No	***58%***	61%	59%	59%	52%	51%	51%	50%	53%	
Yes	***42%***	39%	41%	41%	48%	49%	49%	50%	47%	
		Attitudes toward owning and using a car				Fuel type of the current main car				
Sell a car	***Original***	Financially limited and not thrilled users	Pragmatic car users	White-collar car freaks	Finically restricted car lovers	Gasoline	Diesel	LPG & CNG	Hybrid	Full electric
No	***58%***	58%	59%	60%	57%	62%	56%	50%	50%	53%
Yes	***42%***	42%	41%	40%	43%	38%	44%	50%	50%	46%
		Age of the current main car						Size of the current main car		
Sell a car	***Original***	(0,1990]	(1990–2000]	Euro 3	Euro 4	Euro 5	Euro 6	S	M	L
No	***58%***	54%	55%	60%	61%	62%	62%	62%	57%	50%
Yes	***42%***	46%	45%	40%	39%	38%	38%	38%	43%	50%

of this association. Interestingly, households currently without cars are not the ones with the highest probability of buying a car (62%) but owning currently a large car would increase the likelihood of purchasing a car most (79%). If households who have not had any cars, purchase cars, the share of various emission codes in their purchased fleet is quite uniform (9% - 10%), while those who already own small or medium-sized cars would purchase mainly vehicles from Euro 4 (10% and 12%) and Euro 5 (13% and 20%). In terms of the choice for the size of purchased cars, the households who did not previously own a car would mainly purchase small cars (23%), followed by the other two sizes with the same probability (20%). On the contrary, if households own a small or medium-sized car, they would mainly purchase small cars (27%) and very rarely opt for large-size cars. Lastly, households who currently own large cars mainly purchase medium-sized cars (35%) followed by small-sized cars (27%).

Table 7 shows how and to what extent the parent nodes associated with "Sell a car" influence this particular decision. The second column shows the original probability distribution of "Sell a car" without any evidence settings (before instantiating). The original probability distribution was obtained by learning the parameters of the DBN with data. From Table 7, it is realized that households with two or more cars have a slightly higher probability to sell a car than households with only one car. When households change their residential place, the probability of selling a car in the same year is higher than when no moving takes place. The likelihood of selling cars is highest when the residential place is moved to a lower-order city but a higher-order subzone. The "Financially restricted" groups either car lovers or not really thrilled by cars have the highest probability to sell cars. As for the "Fuel type of the current main car", being a parent node for "Sell a car", owning a gasoline car is associated with the lowest probability of selling cars. Instantiating "Age of the current main car" as yet another parent node of "Sell a car" reveals that, as expected, the older the cars, the higher the probability of selling them is. Lastly, the small-sized cars are least likely to be disposed, while the large-sized cars are most likely to be sold.

6.3. Prediction accuracy of "Buy a car", "Sell a car", and other life events

To test the performance of the DBN model, data were divided into training (70%) and test (30%) sets. The training dataset was used to learn the parameters of the network, and the test dataset was used to make the inference for each life event and test the accuracy of prediction. The life events to be predicted include "Marital status change", "Birth of a baby", "Employment mobility", "Driving license obtain", "Work location mobility", "Residential mobility", "Sell a car", and "Buy a car". The prediction accuracy of size, fuel type, and age of the car were examined and reported as well. The life event decisions were simulated using the Monto Carlo method based on the probability distribution obtained from the inference process.

To show the performance, "Accuracy", "Recall", "Precision" and "F1-score" were calculated (Stehman, 1997; Balfe and Smyth, 2005; Joubert and De Waal, 2020). The "Accuracy" shows the proportion of correct predictions. The "Recall" is measured as the proportion of correctly predicted positive cases relative to the number of all true positives while the "Precision" represents the proportion of correctly predicted positive occurrences relative to all predicted positive. The "F1 – score" is the weighted harmonic mean of the "Recall" and "Precision". Eqs. (8–11) show the calculations for Accuracy, Recall, Recall, and F1-score. *True positive (TP)* is the number of individuals (or households depending on the type of event) for which the event occurs in the data and also is simulated as such in DBN. *False-positive (FP)* represents the number of cases in which the simulation predicts an event occurs but is not observed in the data. *True negative (TN)* represents the number of individuals/households for whom a certain event neither happened in the sample nor appeared in the simulation results. The *False negative (FN)* represents cases for which the event occurred in the data but did not appear in the simulation results. To evaluate the overall performance of DBN, the weighted average was used to consider the prediction performance of each life event.

$$Accurancy = \frac{TP + TN}{\sum test\ datasets} \tag{8}$$

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$F1 - score = \frac{2^*(Recall^*Precision)}{Recall + Precision} \tag{11}$$

where x_{ap} represents the number of individuals in the test dataset who had a status $a \in C$ while the predicted status for those same individuals was $p \in C$. $C = \{1, 2, \ldots, c\}$ is the set of states for each life event.

One hundred Monte Carlo draws were used for the inferences and calculating the performance indices. The prediction indices including “Accuracy”, “Precision”, “Recall”, and “F1 – Score” of life events are reported in Table 8. Regardless of the type of index, as expected, the prediction performance deteriorates moving forward in time reflected by the lower values of indicators for the later years. The reason is that the predicted life events in the preceding year serve as the input for the following years in the simulation. This will accumulate the input error for the succeeding years. Some life events have this deterioration less than the others. For instance, the “Driving license obtain” hardly loses its performance in terms of “Accuracy” and “Recall” over the 5 predicted years. “Sell a car” and then “Marital status change” come second and third in terms of small diminish in their prediction performances. On the other hand, “Work location mobility”, “Buy a car” and “Residential mobility” have larger deterioration of accuracy across the years. “Precision” and “F1-Score” of different life events have a slightly different order in terms of declining performance over time. While “Driving license obtain” still takes the lead in preserving its prediction performance over the course of 5 years, “Marital status change” comes second followed by “Sell a car”.

While the above compares the degradation of prediction performance over time, when it comes to the average quality of prediction, the best index values are associated with “Driving license obtain” with values in the magnitude of 93 to 98%. The second and third best predicted life events are (for all indicators) “Marital status change” and “Birth of a baby”. “Residential mobility” and “Work location mobility” with the classification we devised (order of municipality and order of subzones) appear to perform well (precision of 93% and 94% respectively). “Buy a car” comes last in the list in terms of performance (precision of 59%) because of the large number of states (76 classifications) involved in the definition of this life event. The small number of data points in some of these states is mainly responsible for the lower performance of this life event. “Sell a car” performs better with an average precision of 72%.

To further examine the predictive performance of prediction for age/fuel type/size of purchased cars, Table 9 reports the share of correct prediction for each of these aspects separately. The size and fuel of purchased cars have similar prediction performance (63%), which is slightly better than the age of the car (61%).

Levenshtein distance (LD) (Levenshtein, 1966; Joubert and De Waal, 2020), using the multi-dimensional sequential alignment method, was also calculated to identify the model prediction at the individual level. LD shows to what extent the prediction of each of the following eight life events, “Birth of a baby”, “Residential mobility”, “Sell a car”, “Buy a car”, “Marital status change”, “Work location mobility”, “Employment mobility”, “Driving license obtain” for each individual, over five years (2013–2017) is consistent with the observed life events at individual level. LD calculates the smallest number of operations (substitution, insertion and deletion) needed to align two multi-dimensional sequences. In this case, each year is treated as a dimension and eight life events, represented by 1 (occurred) and 0 (not occurred) construct the elements of each dimension. To normalize the LD value, the calculated Levenstein distances have been divided by the maximum value, which is the maximum number of operations required to equalize two entirely dissimilar sequences. The cost of deletion and insertion was set to 1 while the cost of substitution was set to 2. After normalizing the value of Levenshtein distance, the average normalized Levenshtein distance (for all individuals) has been calculated as 0.23. This value suggests that the average difference between the observed life events at the individual level over 5 years and the predicted events is 0.23 of such differences when the observed and the predicted values are entirely dissimilar.

To compare the precision of DBN with some other alternative models, the Random Forest (RF) and the Bayesian Belief Network (BBN) are applied to predict “Sell a car” and “Buy a car”. To create random forest (RF), the number of trees is set to 100 while the splitting criterion is set to “entropy”. The min_samples_split is set to 2. In constructing BBN, the criteria in terms of learning algorithm and maximum parent nodes have been set the same as the ones in DBN. These three methods all use the same proportion of 70% of the data for model training and 30% for model testing. DBN enabled us to learn the lagged effect from the data and thus to use certain variables of T-1 in prediction of events in T. We used the learned structure from DBN and applied the same specification to estimate RF. More precisely, for “Buy a car” additional life events of “Sell a car” and “Birth of a baby” in year T-1 are added as additional variables. The average prediction precision is 54.7% for RF while 57.6% for BBN and 59.3% for DBN. For “Selling a car”, the specification in RF has been set identical to the ones learned by structural learning of DBN. The precision for selling car using the DBN is 72.4%, which is higher than those using the RF (68.4%) and the BBN (69.8%). The comparison results show better performance of the DBN method. It is important to note that, as stated in Section 6.1, in creation of DBN, unreasonable causations (e.g., life events influence social demographic such as age and gender) have been blocked and experts’ knowledge based on the literature review has been added to the network (e.g., “Gender”, “Highest completed education level” and “Household composition” as the parent nodes of “Employment mobility”, etc.). Such a treatment makes DBN beyond a blindly unsupervised method and can explain better performance of it compared to RF. In addition, equally important is that DBN should not be judged against other methods only on the basis of the prediction accuracy but also additional insights it can provide concerning lead and lag effects.

7. Conclusions and future work

In this paper, the DBN approach was implemented for modeling vehicle transactions considering life events and latent attitudes toward car ownership and use. The DBN captured the dynamic interdependencies among life events and socio-demographic variables. The results show that all life events were predicted with high accuracy. The detailed aspects of purchased cars such as age/fuel type/size could also be predicted with a precision of higher than 60%. Incorporating the dynamic attitude in the decision to purchase a car, as an integrated part of the DBN, revealed its noticeable associations with the purchase decision, fuel type, and age of the purchased cars. In addition, the inclusion of a much less explored determinant of purchasing a car (i.e., the number of children between 12 and 18 in the household, selling car in the same year or a year before, size of the current car) in our DBN suggests that the decision of buying a car in general and the size of the purchased car in particular is very much impacted by these variables.

Although attitudinal components are generally scarce and not typically included in data collection, DBN is expected to hold its power and outperformance even without attitudinal factors in the network. Moreover, the topic of synthesizing attitudes using deep learning methods is on the research agenda and expected to progress in coming years.

Although the DBN approach has value in considering interdependencies among many socio-demographics, life events, and attitudes, this study has some limitations. The current sample size (1035 individuals, 737 households) for predicting 76 categories of “Buy a car” (combination of age, fuel type, and size) may have been limited, reflected in lower prediction performance compared to the other life events.

In addition, the spatial resolution in the sample was at the level of traffic analysis zone. Having data with higher resolution (such as 4-digit postcode area in the Netherlands) would allow the incorporation of some other factors (nodes in the DBN) such as the accessibility to various transport modes, parking facilities, and so on. These issues will be addressed in our future work.

To test whether the sample size, especially for some life events with low occurrence, is sufficient, we tested the stability of the network structure learning and parameter learning. To that end, multiple randomly selected 70% from the sample was used to form subsamples. For each subsample, we ran the network structure learning and parameter learning and compared the similarities of the outcomes with those associated with the full sample. For the similarity in network structure, we calculated the percentage of overlapping links; for the similarity in parameter learning, we calculated the Jenson Shannon Divergence (JSD) between the learned CPTs. Although there are no strict thresholds for defining high similarities, we found that when the subsample size is 70% of the full sample, the average similarity of structure is 80.29%, indicating relatively high stability. The JSD is 0.02 bits (the range of JSD is [0,1] and a higher value means a larger divergence). The result also indicates 70% of samples are acceptable. Since we only have a small sample for the current study, these tests also motivate us to secure large samples to perform follow-up analysis in future work.

This is in line with the previous findings. Kontkanen et al. (1997) claimed that Bayesian models can achieve good prediction accuracy even with rather small sample sizes. Lähdesmäki and Shmulevich (2008) concluded that the performance would not be perfect even though infinite sample size is used. They also tested the change in performance by increasing the sample size from 50 to 250 and the performance did not increase significantly. Zou and Feng (2009) mentioned that the Granger causality approach and DBN as two common approaches in dealing with multi-dimensional temporal data to explore the network structure and concluded that DBN performs better especially when the data size is small. For instance, Guo et al. (2019, 2020a) used only 414 respondents and 266

Table 8
The weighted accuracy, precision, recall, and F1-score of predicted life events.

		Birth of a baby	Residential mobility	Sell a car	Buy a car	Marital status change	Work location mobility	Employment mobility	Driving license obtain
Accuracy (%)	2013	91.09	87.73	71.63	56.97	93.3	83.77	57.3	93.01
	2014	86.12	83.18	71.37	56.71	92.48	77.33	51.92	92.98
	2015	83.83	80.39	66.94	52.57	91.85	67.93	49.53	92.12
	2016	81.8	77.41	71.03	44.51	88.29	62.04	47.45	92.02
	2017	79.83	74.6	64.86	40.75	86.39	59.5	45.17	93.25
	Average	**84.53**	**80.66**	**69.17**	**50.30**	**90.46**	**70.11**	**50.27**	**92.68**
Precision (%)	2013	95.01	92.8	77.84	68.62	97.51	95.76	76.88	98.38
	2014	92.57	91.3	77.2	67.78	97.7	97.35	75.07	98.05
	2015	93.56	91.64	67.14	62.62	98.53	96.28	78.87	97.59
	2016	93.47	94.26	74.27	46.56	97.38	90.31	76.61	97.46
	2017	94.52	93.87	65.65	50.81	99.25	89.85	76.63	100
	Average	**93.83**	**92.77**	**72.42**	**59.28**	**98.07**	**93.91**	**76.81**	**98.30**
Recall (%)	2013	91.09	87.73	71.63	56.97	93.3	83.77	57.3	93.01
	2014	86.12	83.18	71.37	56.71	92.48	77.33	51.92	92.98
	2015	83.83	80.39	66.94	52.57	91.85	67.93	49.53	92.12
	2016	81.8	77.41	71.03	44.51	88.29	62.04	47.45	92.02
	2017	79.83	74.6	64.86	40.75	86.39	59.5	45.17	93.25
	Average	**84.53**	**80.66**	**69.17**	**50.30**	**90.46**	**70.11**	**50.27**	**92.68**
F1-Score (%)	2013	92.88	90.18	74.3	62.22	95.34	89.31	65.19	95.48
	2014	88.98	87	73.91	61.74	94.99	86.13	60.9	95.28
	2015	88.01	85.56	67.04	57.08	95.05	79.52	60.14	94.6
	2016	86.87	84.95	72.47	45.36	92.53	73.16	57.78	94.52
	2017	86.2	83.04	65.23	45.21	92.28	71.14	56.17	96.51
	Average	**88.59**	**86.15**	**70.59**	**54.32**	**94.04**	**79.85**	**60.04**	**95.28**

Table 9
Accuracy, Precision, Recall, and F1-Score for Age/Fuel/Size of the purchased car(%).

	Accuracy(%)			Precision(%)			Recall(%)			F1-Score(%)		
	Age	Fuel	Size	Age	Fuel	Size	Age	Fuel	Size	Age	Fuel	Size
2013	58.21	59.5	60.27	69.23	70.16	70.48	58.21	59.5	60.27	63.07	64.16	64.67
2014	57.8	59.45	59.81	68.38	69.88	69.64	57.8	59.45	59.81	62.55	64.11	64.07
2015	54.13	55.53	56.19	63.89	64.71	64.75	54.13	55.53	56.19	58.49	59.69	59.97
2016	47.21	50.16	50.78	48.86	54.66	52.96	47.21	50.16	50.78	47.94	51.95	51.73
2017	43.47	46.65	47.01	53.09	56.4	55.91	43.47	46.65	47.01	47.68	50.91	50.83
Average	**52.16**	**54.26**	**54.81**	**60.69**	**63.16**	**62.75**	**52.16**	**54.26**	**54.81**	**55.95**	**58.16**	**58.25**

households for modeling dependencies among mobility decisions and key life events using the DBN. In our study, the sample size is more than double that.

Lastly, to create such a dynamic network, longitudinal data on various life events over multiple years from a sufficiently large panel is required. Such data in general is scarce. It is therefore of high importance to put on the research agenda the transferability of the constructed network for one area to other areas with similar and different social, and spatial contexts.

CRediT authorship contribution statement

Yajie Yang: Conceptualization, Methodology, Data curation, Writing – original draft. **Soora Rasouli:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing, Supervision. **Feixiong Liao:** Conceptualization, Methodology, Writing – review & editing, Supervision.

Data availability

The authors do not have permission to share data.

Acknowledgment

The first author of this research was supported by the China Scholarship Council (CSC). We appreciate several anonymous reviewers and editor for their comments and support.

Appendix

1. Paired Samples Test for attitude variables and corresponding EFA

Table A.1
Paired samples test about these thirteen variables.

	2014–2016		2014–2018		2016–2018	
	t-value	Sig(2-tailed)	t	Sig(2-tailed)	t-value	Sig(2-tailed)
I_1	3.279	0.001	3.890	0.000	1.200	0.230
I_2	1.684	0.093	1.815	0.070	0.445	0.656
I_3	1.912	0.056	2.083	0.038	0.455	0.649
I_4	2.502	0.013	10.440	0.000	8.388	0.000
I_5	1.367	0.172	19.854	0.000	19.072	0.000
I_6	5.050	0.000	13.459	0.000	8.944	0.000
I_7	1.614	0.107	−4.586	0.000	−5.396	0.000
I_8	0.867	0.386	−6.594	0.000	−7.430	0.000
I_9	−1.401	0.162	−3.802	0.000	−2.475	0.013
I_{10}	−1.386	0.166	−1.211	0.226	0.175	0.861
I_{11}	−3.863	0.000	−6.868	0.000	−3.114	0.002
I_{12}	−1.519	0.129	−4.652	0.000	−3.770	0.000
I_{13}	−0.761	0.447	−16.335	0.000	−15.759	0.000

Table A.2
Result of exploratory factor analysis of "Attitudes towards owning and using a car".

	Rotated Component (Factor loading)				Communalities	Eigenvalue	Variance explained(%)	Crobach's alpha
	Factor 1	Factor 2	Factor 3	Factor 4				
I_1	0.012	**0.753**	0.259	0.074	0.64	2.379	18.297	0.777
I_2	−0.058	**0.866**	−0.021	0.157	0.778			
I_3	−0.204	**0.523**	0.496	0.06	0.565			
I_4	0.086	**0.848**	−0.1	0.023	0.738			
I_5	−0.008	0.122	**0.767**	0.021	0.604	1.997	15.364	0.787
I_6	−0.001	0.049	**0.764**	−0.104	0.597			
I_7	0.126	0.181	−0.098	**0.837**	0.758	1.572	12.093	0.698
I_8	−0.023	0.044	−0.053	**0.884**	0.787			
I_9	**0.815**	−0.052	−0.027	−0.01	0.668			
I_{10}	**0.877**	−0.019	−0.031	0.008	0.77			
I_{11}	**0.833**	0	−0.034	0.034	0.696	2.654	20.417	0.787
I_{12}	**0.399**	0.111	−0.6	0.175	0.563			
I_{13}	**0.547**	0.031	−0.354	0.115	0.439			

Note: the number marked in bold represent corresponding variables has higher factor loading value on corresponding factor.

2. Sub figure of the DBN model

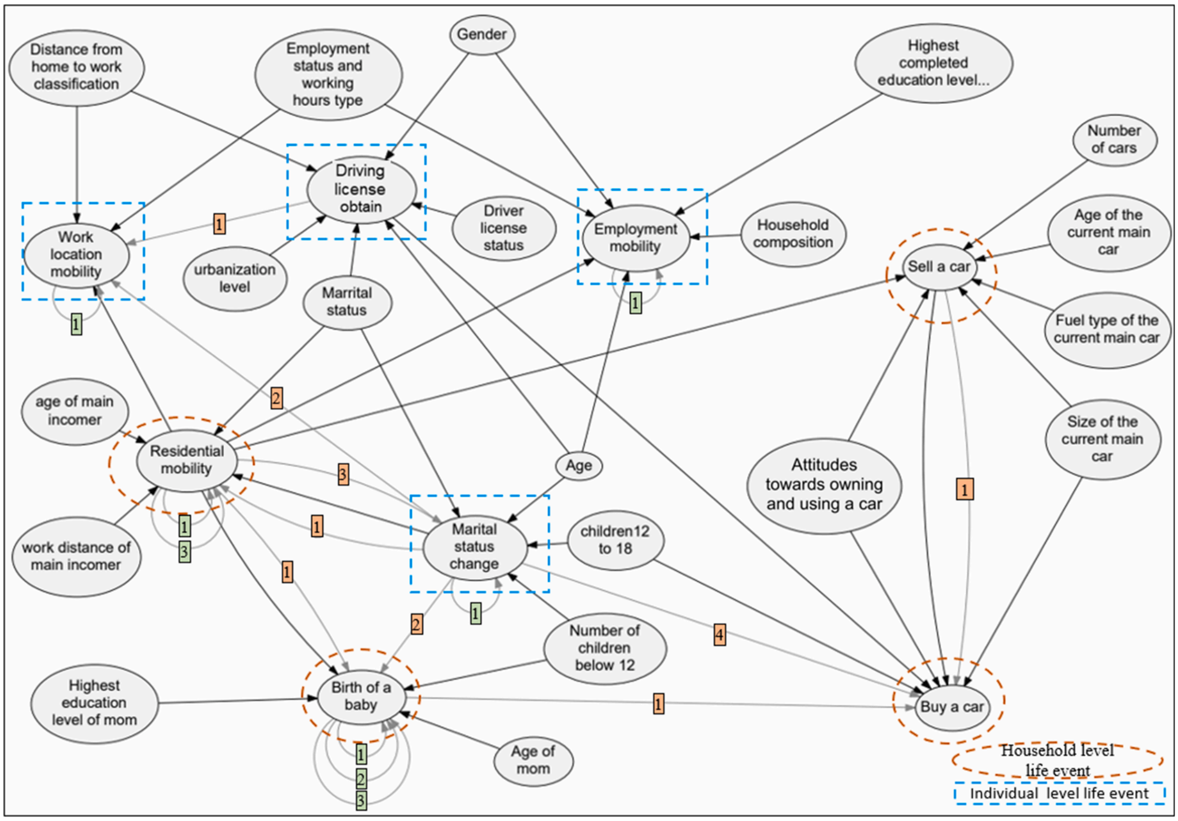

Fig. A.1. Full DBN.

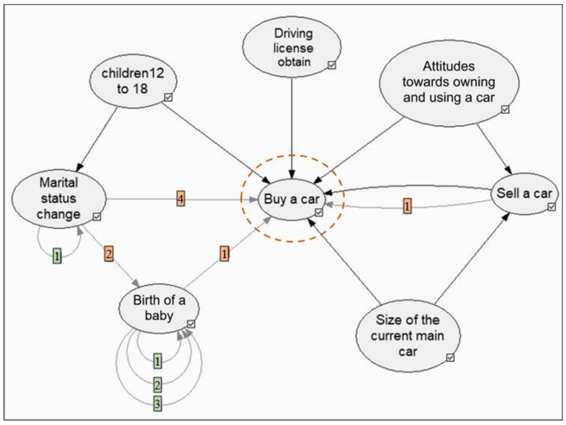

Fig. A.2. “Buy a car” in DBN.

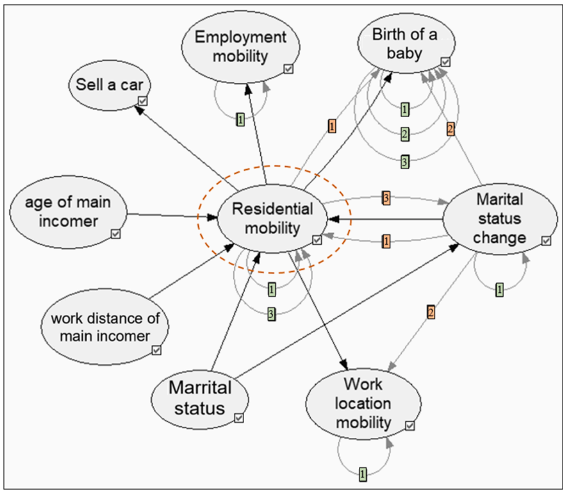

Fig. A.3. “Residential mobility” in DBN.

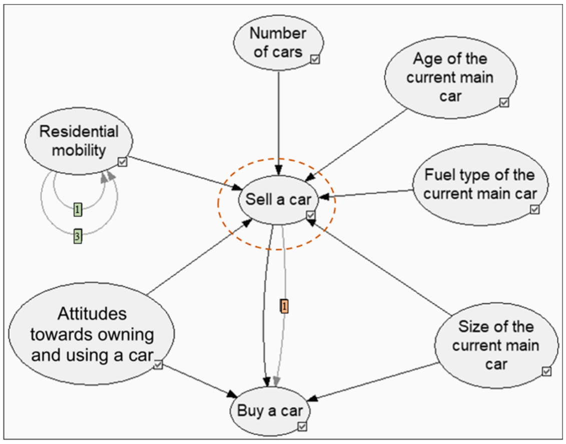

Fig. A.4. "Sell a car" in DBN.

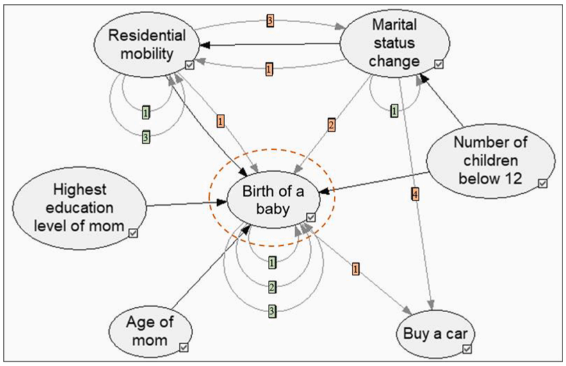

Fig. A.5. "Birth of a baby" in DBN.

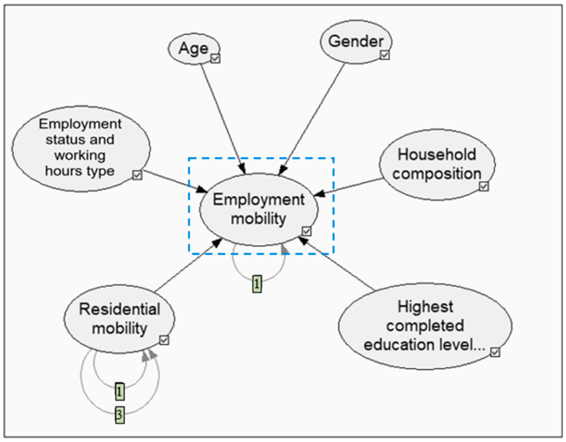

Fig. A.6. "Employment mobility" in DBN.

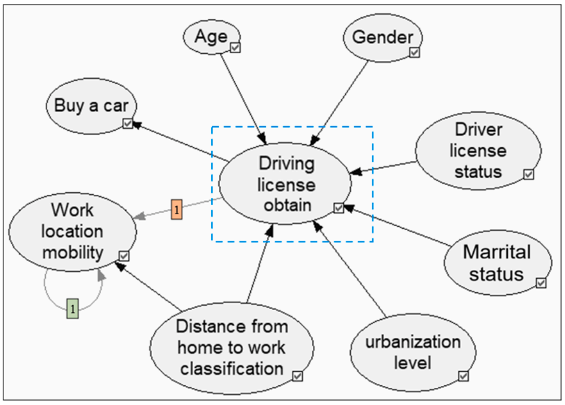

Fig. A.7. “Driving license obtain” in DBN.

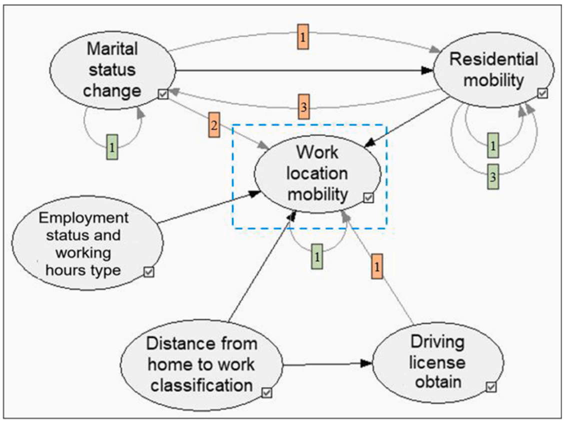

Fig. A.8. “Work location mobility” in DBN.

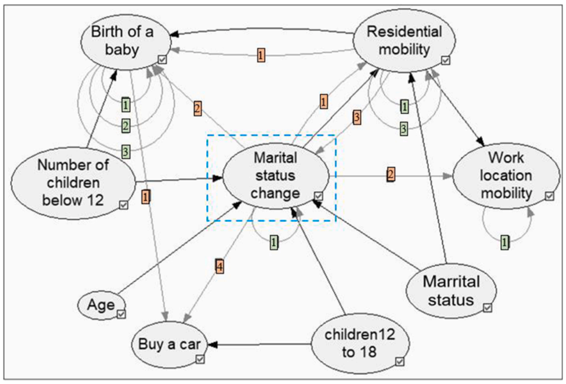

Fig. A.9. “Marital status change” in DBN.

References

Achtnicht, M., 2012. German car buyers’ willingness to pay to reduce CO 2 emissions. Clim. Chang. 113 (3), 679–697.

Awad, A., Yussof, I., 2017. Factors affecting fertility-new evidence from Malaysia. Bull. Geogr. Socio-econom. Ser. 36, 7–20.

Balfe, E., Smyth, B., 2005. An analysis of query similarity in collaborative web search. In: European Conference on Information Retrieval. Springer, Berlin, Heidelberg, pp. 330–344.

Beck, M.J., Rose, J.M., Hensher, D.A., 2013. Environmental attitudes and emissions charging: an example of policy implications for vehicle choice. Transp. Res. A Policy Pract. 50, 171–182.

Belgiawan, P.F., Schmöcker, J.-D.D., FUJII, S., 2011. Psychological determinants for Car ownership decisions. In: Proceedings of the 16th International Conference of Hong Kong Society for Transportation Studies (HKSTS), Hong Kong, pp. 1–7.

Ben-Akiva, M., Lerman, S.R., 1974. Some estimation results of a simultaneous model of auto ownership and mode choice to work. Transportation 3 (4), 357–376.
Bhat, C.R., Guo, J.Y., 2007. A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. Transp. Res. B Methodol. 41 (5), 506–526.
Bhat, C.R., Koppelman, F.S., 1993. An endogenous switching simultaneous equation system of employment, income, and car ownership. Transp. Res. A Policy Pract. 27 (6), 447–459.
Bhat, C.R., Pulugurta, V., 1998. A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions. Transp. Res. B Methodol. 32 (1), 61–75.
Bhat, C.R., Sen, S., Eluru, N., 2009. The impact of demographics, built environment attributes, vehicle characteristics, and gasoline prices on household vehicle holdings and use. Transp. Res. B Methodol. 43 (1), 1–18.
Browning, M., Bourguignon, F., Chiappori, P.A., Lechene, V., 1994. Income and outcomes: a structural model of intrahousehold allocation. J. Polit. Econ. 102 (6), 1067–1096.
Clark, W.A., Withers, S.D., 2002. Disentangling the interaction of migration, mobility, and labor-force participation. Environ. Plan. A 34 (5), 923–945.
Clark, B., Chatterjee, K., Melia, S., 2016. Changes in level of household car ownership: the role of life events and spatial context. Transportation 43 (4), 565–599.
Daly, A., Hess, S., Patruni, B., Potoglou, D., Rohr, C., 2012. Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. Transportation 39 (2), 267–297.
Dargay, J.M., 2001. The effect of income on car ownership: evidence of asymmetry. Transp. Res. A Policy Pract. 35 (9), 807–821.
Dargay, J.M., 2002. Determinants of car ownership in rural and urban areas: a pseudo-panel analysis. Transp. Res. Part E: Logist. Transp. Rev. 38 (5), 351–366.
Dargay, J.M., Vythoulkas, P.C., 1999. Estimation of a dynamic car ownership model: a pseudo-panel approach. J. Transp. Econom. Pol. 287–301.
de Jong, G.C., Kitamura, R., 2009. A review of household dynamic vehicle ownership models: holdings models versus transactions models. Transportation 36 (6), 733–743.
Degraeve, Z., Proost, S., 2001. Cost-Efficiency Methodology for the Selection of New car Emission Standards in Europe. KU Leuven CES; Leuven Working Papers, pp. 1–27.
European Vehicle Market Statistics 2016/17. (2016).
Fatmi, M.R., Habib, M.A., 2016. Longitudinal vehicle transaction model: assessment of lead and lagged effects of longer-term changes and life-cycle events. Transp. Res. Rec. 2566 (1), 11–21.
Flamm, B., 2009. The impacts of environmental knowledge and attitudes on vehicle ownership and use. Transp. Res. Part D: Transp. Environ. 14 (4), 272–279.
Fontaras, G., Zacharof, N.G., Ciuffo, B., 2017. Fuel consumption and CO2 emissions from passenger cars in Europe – laboratory versus real-world emissions. Prog. Energy Combust. Sci. 60, 97–131.
Goodman, L.A., 1974. The analysis of Systems of Qualitative Variables When some of the variables are unobservable. Part I-A modified latent structure approach. Am. J. Sociol. 79 (5), 1179–1259.
Gu, G., Feng, T., Yang, D., Timmermans, H., 2021. Modeling dynamics in household car ownership over life courses: a latent class competing risks model. Transportation 48 (2), 809–829.
Guo, J., Feng, T., Timmermans, H.J.P., 2019. Time-varying dependencies among mobility decisions and key life course events: an application of dynamic Bayesian decision networks. Transp. Res. A Policy Pract. 130, 82–92.
Guo, J., Feng, T., Timmermans, H.J.P., 2020a. Co-dependent workplace, residence and commuting mode choice: results of a multi-dimensional mixed logit model with panel effects. Cities 96, 102448.
Guo, J., Feng, T., Zhang, J., Timmermans, H.J.P., 2020b. Temporal interdependencies in mobility decisions over the life course: a household-level analysis using dynamic Bayesian networks. J. Transp. Geogr. 82, 102589.
Heckerman, D., Geiger, D., Chickering, D.M., 1995. Learning Bayesian networks: the combination of knowledge and statistical data. Mach. Learn. 20 (3), 197–243.
Hjorthol, R., 2016. Decreasing popularity of the car? Changes in driving licence and access to a car among young adults over a 25-year period in Norway. J. Transp. Geogr. 51, 140–146.
Hong, Y., 2006. Marital decision-making and the timing of first birth in rural China before the 1990s. Popul. Stud. 60 (3), 329–341.
Hoogendoorn-Lanser, S., Schaap, N.T.W., Oldekalter, M.J., 2015. The Netherlands mobility panel: an innovative design approach for web-based longitudinal travel data collection. Transp. Res. Proc. 11, 311–329.
Jain, A.K., 1981. The effect of female education on fertility: a simple explanation. Demography 18 (4), 577–595.
Jain, T., Rose, G., Johnson, M., 2021. Changes in private car ownership associated with car sharing: gauging differences by residential location and car share typology. Transportation 1–25.
Joubert, J.W., De Waal, A., 2020. Activity-based travel demand generation using Bayesian networks. Transp. Res. Part C: Emerg. Technol. 120, 102804.
Khan, N.A., Habib, M.A., 2021. When do households make vehicle transactions? What type of vehicle households choose after transaction events? Travel Behav. Soc. 22, 32–47.
Kim, S.H., Mokhtarian, P.L., 2018. Taste heterogeneity as an alternative form of endogeneity bias: investigating the attitude-moderated effects of built environment and socio-demographics on vehicle ownership using latent class modeling. Transp. Res. A Policy Pract. 116, 130–150.
Kim, J., Rasouli, S., Timmermans, H., 2016. A hybrid choice model with a nonlinear utility function and bounded distribution for latent variables: application to purchase intention decisions of electric cars. Transportmetrica A: Transp. Sci. 12 (10), 909–932.
Kim, J., Rasouli, S., Timmermans, H., 2017. Satisfaction and uncertainty in car-sharing decisions: an integration of hybrid choice and random regret-based models. Transp. Res. A Policy Pract. 95, 13–33.
Knez, M., Jereb, B., Obrecht, M., 2014. Factors influencing the purchasing decisions of low emission cars: a study of Slovenia. Transp. Res. Part D: Transp. Environ. 30, 53–61.
Knodel, J., Hermalin, A.I., 1984. Effects of birth rank, maternal age, birth interval, and sibship size on infant and child mortality: evidence from 18th and 19th century reproductive histories. Am. J. Public Health 74 (10), 1098–1106.
Kontkanen, P., Myllymäki, P., Silander, T., Tirri, H., Grünwald, P., 1997, January. Comparing predictive inference methods for discrete domains. In: Sixth International Workshop on Artificial Intelligence and Statistics. PMLR, pp. 311–318.
Kowalska-Pyzalska, A., Kott, M., Kott, J., 2021. How much polish consumers know about alternative fuel vehicles? Impact of knowledge on the willingness to buy. Energies 14 (5), 1438.
Kwisthout, J., 2011. Most probable explanations in Bayesian networks: complexity and tractability. Int. J. Approx. Reason. 52 (9), 1452–1469.
Lähdesmäki, H., Shmulevich, I., 2008. Learning the structure of dynamic Bayesian networks from time series and steady state measurements. Mach. Learn. 71 (2), 185–217.
Levenshtein, V.I., 1966. Binary codes capable of correcting deletions, insertions, and reversals. In: Soviet Physics Doklady, Vol. 10, No. 8, pp. 707–710.
Li, S., Zhao, P., 2017. Exploring car ownership and car use in neighborhoods near metro stations in Beijing: does the neighborhood built environment matter? Transp. Res. Part D: Transp. Environ. 56, 1–17.
Liao, F., 2019. Space–time prism bounds of activity programs: a goal-directed search in multi-state supernetworks. Int. J. Geogr. Inf. Sci. 33 (5), 900–921.
Maat, K., Kasraian, D., 2020. Dual car ownership: the role of intra-household interaction, the commute and the spatial context. In: Heart 2020: 9th Symposium of the European Association for Research in Transportation.
Maltha, Y., Kroesen, M., Van Wee, B., van Daalen, E., 2017. Changing influence of factors explaining household car ownership levels in the Netherlands. Transp. Res. Rec. 2666 (1), 103–111.
Mazur, C., Offer, G.J., Contestabile, M., Brandon, N.B., 2018. Comparing the effects of vehicle automation, policy-making and changed user preferences on the uptake of electric cars and emissions from transport. Sustainability 10 (3), 676.
McCutcheon, A.L., 1987. Latent Class Analysis (Issue 64). Sage.
Mihajlovic, V., Petkovic, M., 2001. Dynamic Bayesian Networks: A State of the Art. University of Twente Document Repository.

Mohammadian, A., Miller, E.J., 2002. Understanding and modeling dynamics of household automobile ownership decisions: a Canadian experience. In: Applications of Advanced Technologies in Transportation (2002), pp. 672–679.
Moro, A., Lonza, L., 2018. Electricity carbon intensity in European member states: impacts on GHG emissions of electric vehicles. Transp. Res. Part D: Transp. Environ. 64 (November 2016), 5–14.
Murphy, K.P., 2002. Dynamic Bayesian Networks: Representation, Inference and Learning. University of California, Berkeley.
Nolan, A., 2010. A dynamic analysis of household car ownership. Transp. Res. A Policy Pract. 44 (6), 446–455.
Oakil, A.T., 2013. Temporal Dependence in Life Trajectories and Mobility Decisions (Doctoral dissertation,. Utrecht University.
Oakil, A.T.M., Ettema, D., Arentze, T., Timmermans, H., 2014. Changing household car ownership level and life cycle events: an action in anticipation or an action on occurrence. Transportation 41 (4), 889–904.
Oakil, A.T.M., Manting, D., Nijland, H., 2016. Determinants of car ownership among young households in the Netherlands: the role of urbanisation and demographic and economic characteristics. J. Transp. Geogr. 51, 229–235.
Pearl, J., 1988. Embracing causality in default reasoning. Artif. Intell. 35 (2), 259–271.
Posel, D.R., 2001. Who are the heads of household, what do they do, and is the concept of headship useful? An analysis of headship in South Africa. Dev. South. Afr. 18 (5), 651–670.
Potoglou, D., Kanaroglou, P.S., 2008. Disaggregate demand analyses for conventional and alternative fueled automobiles: a review. Int. J. Sustain. Transp. 2 (4), 234–259.
Potoglou, D., Susilo, Y.O., 2008. Comparison of vehicle-ownership models. Transp. Res. Rec. 2076 (1), 97–105.
Qin, J., Liao, F., 2021. Space–time prism in multimodal supernetwork-part 1: methodology. Commun. Transp. Res. 1, 100016.
Rasouli, S., Timmermans, H., 2014. Accounting for heterogeneity in travel episode satisfaction using a random parameters panel effects regression model. Procedia Environ. Sci. 22, 35–42.
Roos, J., Bonnevay, S., Gavin, G., 2017. Dynamic Bayesian networks with Gaussian mixture models for short-term passenger flow forecasting. In: In 2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE). IEEE, pp. 1–8.
Saadeh, F.B., Clark, M.A., Rogers, M.L., Linkletter, C.D., Phipps, M.G., Padbury, J.F., Vivier, P.M., 2013. Pregnant and moving: understanding residential mobility during pregnancy and in the first year of life using a prospective birth cohort. Matern. Child Health J. 17 (2), 330–343.
Serrenho, A.C., Norman, J.B., Allwood, J.M., 2017. The impact of reducing car weight on global emissions: the future fleet in Great Britain. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 375 (2095).
Stehman, S.V., 1997. Selecting and interpreting measures of thematic classification accuracy. Remote Sens. Environ. 62 (1), 77–89.
Sun, L., Erath, A., 2015. A Bayesian network approach for population synthesis. Transp. Res. Part C: Emerg. Technol. 61, 49–62.
Suzuki, J., 2017. An efficient Bayesian network structure learning strategy. N. Gener. Comput. 35 (1), 105–124.
Testa, M.R., 2014. On the positive correlation between education and fertility intentions in Europe: individual-and country-level evidence. Adv. Life Course Res. 21, 28–42.
Ton, D., Zomer, L.B., Schneider, F., Hoogendoorn-Lanser, S., Duives, D., Cats, O., Hoogendoorn, S., 2020. Latent classes of daily mobility patterns: the relationship with attitudes towards modes. Transportation 47 (4), 1843–1866.
Train, K., Lohrer, M., Systematics, C., Commission, C. E, 1982. Vehicle ownership and usage: An integrated system of disaggregate demand models. In: The 62nd Annual Meeting of the Transportation Research Board, CD-ROM. Washington, DC.
Tsamardinos, I., Brown, L.E., Aliferis, C.F., 2006. The max-min hill-climbing Bayesian network structure learning algorithm. Mach. Learn. 65 (1), 31–78.
Uusitalo, L., 2007. Advantages and challenges of Bayesian networks in environmental modelling. Ecol. Model. 203 (3–4), 312–318.
Valeri, E., Cherchi, E., 2016. Does habitual behavior affect the choice of alternative fuel vehicles? Int. J. Sustain. Transp. 10 (9), 825–835.
Van Acker, V., Mokhtarian, P.L., Witlox, F., 2014. Car availability explained by the structural relationships between lifestyles, residential location, and underlying residential and travel attitudes. Transp. Policy 35, 88–99.
van Ham, M., Mulder, C.H., Hooimeijer, P., 2001. Spatial flexibility in job mobility: macrolevel opportunities and microlevel restrictions. Environ. Plan. A 33 (5), 921–940.
Velders, G.J.M., Wesseling, J., Geilenkirchen, G.P., Ligterink, N.E., 2013. The Euro Emission Standards for Cars and Trucks in Relation to NO2 Limit Value Exceedances in the Netherlands.
Vermunt, J., 2010. Latent class modeling with covariates: two improved three-step approaches. Polit. Anal. 18 (4), 450–469.
Vilchez, J.J.G., Julea, A., Peduzzi, E., Pisoni, E., Krause, J., Siskos, P., Thiel, C., 2019. Modelling the impacts of EU countries' electric car deployment plans on atmospheric emissions and concentrations. Eur. Transp. Res. Rev. 11 (1), 1–17.
Wang, J., Matsumoto, S., 2021. Can subsidy programs lead consumers to select "greener" products? Evidence from the eco-car program in Japan. Res. Transp. Econ. 101066.
Weinberger, M.B., 1987. The relationship between women's education and fertility: selected findings from the world fertility surveys. Int. Fam. Plan. Perspect. 35–46.
Xiong, C., Yang, D., Zhang, L., 2018. A high-order hidden Markov model and its applications for dynamic car ownership analysis. Transp. Sci. 52 (6), 1365–1375.
Zhao, J., Knoop, V.L., Wang, M., 2022. Microscopic Traffic Modeling inside Intersections: Interactions between Drivers. Transportation Science.
Zou, C., Feng, J., 2009. Granger causality vs. dynamic Bayesian network inference: a comparative study. BMC Bioinform. 10 (1), 1–17.