A multimodal perspective in the study of car sharing switching intentions

Original
A multimodal perspective in the study of car sharing switching intentions / Diana, Marco; Ceccato, Riccardo. - In: TRANSPORTATION LETTERS. - ISSN 1942-7867. - STAMPA. - (2019), pp. 1-7.

Availability:
This version is available at: 11583/2833492 since: 2020-06-08T09:48:45Z

Publisher:
TAYLOR & FRANCIS LTD

Published
DOI:10.1080/19427867.2019.1707351

Marco Diana & Riccardo Ceccato (2019): A multimodal perspective in the study of car sharing switching intentions, Transportation Letters, DOI: 10.1080/19427867.2019.1707351

A multimodal perspective in the study of car sharing switching intentions

Marco Diana[a], Riccardo Ceccato[a,*]

[a]Department of Environment, Land and Infrastructure Engineering, Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129, Torino, Italy

*Corresponding author e-mail: riccardo.ceccato@polito.it

Abstract

The introduction of innovative mobility services such as car sharing leads to changes in travel habits of users, inducing a shift of travel demand from existing travel modes. An analysis of such changes should be carried out in order to promote car sharing, while managing travel demand effectively. In particular, policies should be developed to induce the switch only from private modes, like private car, avoiding the shift from public transport and active modes. In order to reach this aim, data from a mobility survey carried out in the Turin Metropolitan area (Italy) were used to study modal choices. In particular, Decision Trees were adopted to complement the analyses following an econometric approach. One decision tree was estimated for each mode used by respondents in a specific trip under consideration, in order to identify trip attributes that might affect the intention to switch to car sharing. In this way, threshold values of each variable that entice a modal shift are mode-specific, thus better informing policies aimed at maximizing the social benefit of car sharing.

Keywords: stated preferences; travel surveys; travel demand; car sharing; SMOTE technique.

1 Introduction

After the introduction of car sharing, changes in travel habits of users are often reported [1], both considering private and public transport modes. Potentially car sharing has several advantages, including the reduction of car ownership [2] and vehicle miles traveled [3], which contribute to decrease carbon emissions and energy consumption [3,4], the diffusion of electric vehicles [5] and an increase of the user's attitude to combine different travel means [1]. However, in order to take advantage of these positive aspects in a suitable and effective way, authorities have to carry out policies to promote car sharing, avoiding competition with existing sustainable modes (i.e. active modes and public transport) [5], but complementing them. Hence, policies should be targeted in order to shift travel demand from private modes to car sharing [6]. In order to reach this aim, this paper analyses factors affecting the choice to switch to car sharing, developing separate models for each currently adopted mode. In this way, the use of car sharing can be promoted or avoided by varying mode-specific factors.

Several authors identified variables affecting the choice to use car sharing, even if their effects varied according to the specific area of analysis. Concerning socio-economic variables of travelers, car sharing adopters tent to be young [7], with higher employment rate [8], more educated [9], living in households with higher income [1] and fewer cars [10]. Other authors also considered household size [11] and composition [12], the number of owned bikes [13], and spatial information, such as working location [8] and the presence of private parking near home [6]. Moreover, characteristics of a specific trip performed by car sharing were used: cost [14], travel time and distance from the nearest vehicle [15], walking time to reach the car [12], parking cost [13], trip purpose [16], weather conditions [15] and potential alternative modes [6]. Finally, several authors adopted variables related to travel habits, such as usage frequencies of car [12], public transport and other non-motorized modes [1], and related to personal believes, like environmental awareness [11].

In this paper, we aim at studying which kinds of trips are more conducive to a shift from a given travel means to car sharing. Therefore, compared to the above-reviewed papers, only variables related to trip attributes were selected (such as travel time, distance and cost of car sharing trips), since policy makers can at least partly manage them, in order to promote or not the shift to car sharing. In particular, interventions to reduce travel time can be carried out, for example, by giving free access to public transit reserved lanes. Moreover, the walking time to reach the nearest vehicle can be decreased by supplying a large number of shared vehicles. Lastly, policy makers can also influence car sharing fares by providing financial support (subsidies or tax cuts) [17]. The related research question is to understand to which extent car sharing must outperform existing systems to induce a switch. On the other hand, socio-economic variables were not considered, since they are in general out of control. In order to understand and simulate travel behavior of users, starting from the work of McFadden [18], models based on random utility maximization theory have been extensively used [19,20], in particular multinomial logit [21–24] (MNL). These models are based on some statistical and mathematical assumptions on data used to calibrate them [20,25]. For instance, MNL requires independence of irrelevant alternatives (IIAs) [19,22,24–26], i.e. the effect of attributes are compensatory [20,22]. Several models were developed in order to overcome these limitations, such as probit models [27]. Furthermore, in order to introduce correlation effects among alternatives, nested logit, cross-nested logit, ordered generalized extreme values and mixed logit models were implemented [28].

A different approach to analyze data to model mode choice of users is data mining. In transportation analysis, data mining techniques were mostly used to reproduce existing scenarios [29,30], modelling users' choice based on current conditions and options [20,31]. However, in this paper, exogenous variables in the model are related to a hypothetical future scenario, thus enhancing the predictive power of the adopted data mining approach. Like traditional mode choice models, that are based on random utility maximization theory, data mining techniques were used to predict future travel behavior of users, and in particular, mode choices of travelers [32]. Following this approach, mode choice can be defined as a pattern recognition task in which multiple behavioral attributes described by explanatory variables determine the prediction of the choice among different alternatives [22,32]. Therefore, data mining approach can be adopted for modal analysis and prediction.

Unlike other methods, data mining techniques do not require any statistical and mathematical assumption on data structure [19,31,33,34]. Furthermore, they have a flexible structure [19,20,22,35,36], extracting significant patterns from the dataset and leading to a deeper understanding of relationship among explanatory variables [19,20,22,24–26,30,34,35]. Moreover, they can be applied to large databases [28], even with high unbalanced data [36]. On the other hand, results that are quite useful for planning and forecasting purposes and that are commonly derived through an econometric approach, such as the Value Of Time and demand elasticities, cannot be obtained from such techniques, which are very sensitive to training data [28]. Furthermore they often lack of interpretability, indeed they tend to focus more on predictive accuracy rather than on counterfactual analysis [37]. Nevertheless, recently, some authors were able to extract interpretable economic information, such as elasticities, using a data mining approach [38].

Among different data mining techniques, Decision Trees were adopted to model mode choice of users [19,39,40]. Decision Tree is a classification method [41] that, trained on a base dataset, reproduces existing relationship between dependent and independent variables [24,30], without estimating any parameters [30]. Moreover, the general structure of the algorithm simulates the cognitive and decisional process of users [33,42,43]. In this sense, Decision Tree can be considered as a function that uses a vector of attributes as an input variable, and it returns a decision value, after a series of tests [26,44]. Therefore, unlike traditional econometric theory-based models, this method directly extracts information from data, with different assumptions and interpretability, thus contributing to the analysis and interpretation of travel behavior from a different perspective.

In details, Decision Tree considers both continuous and categorical variables [22,26], it evaluates the mixed effect of more than one variable [25], and it produces a visual representation of results to understand of the effect of explanatory variables on the travel behavior of users [19,20,26,34].

Several authors applied Decision Trees as a mode choice model [39,45], embedding them in a 4-step [35] or activity-based model [40]. Other authors used Decision Trees as a part of a more extensive mode choice model, to define groups of people with homogeneous travel behaviors [31] or to identify the most important variables for further analysis [32]. However, to the best of authors' knowledge, the use of a Decision Tree to define variables affecting car sharing adoption is very limited. In particular, only Wang et al. [46] adopted a Hierarchical Tree-based Regression to identify the most important socio-economic attributes and trip characteristics of users that might affect the choice to adopt a car sharing electric vehicle. However, they did not consider the mode that was previously adopted by users. Moreover, data used to calibrate the model were obtained from individuals that were not representative of the overall population of the study area.

Unlike previous works, in this paper, Decision Trees were estimated to select trips attributes that might affect the choice to perform a trip on car sharing, diversifying the analysis according to the mode currently adopted by users, in order to consider the effects of the original means. One of the most interesting results is that trip attributes are also ranked according to their importance in determining the modal shift. Decision Trees were calibrated using data from a mobility survey administered to a representative sample of the population living in the study area; therefore, results are reliable and can be effectively extended to the universe of users, in order to have a sound basis for policy makers to target mode-specific interventions to promote car sharing.

2 Method and data

A Decision Tree is a tree-like classification model, which predicts the value of a target attribute based on input attributes. The output is a graph containing different layers of nodes, where the first one is the root node, which is split into leaves nodes. Each internal node corresponds to a specific attribute, whereas terminal leaves define the predicted class. The value of the class is defined by the values of input attributes following the path from the root node to the leaf.

Among different algorithms to build a Decision Tree, in this paper, the C4.5 algorithm was adopted [47], since it can produce more than two branches for each node and it can manage both categorical and continuous variables [22]. In this case, a Decision Tree is generated using recursive partitioning, i.e., starting from the root node, each leaf node is repeatedly split. In particular, each node is split according to a defined criterion measuring the homogeneity of the node, i.e. only the most homogeneous attribute is considered for the split. In this way, the node contains most of the observations belonging to one predicted class. In order to evaluate the impurity of a node, several measures were proposed, such as accuracy, Gini index, information gain and gain ratio. Moreover, the tree is generated in such a way that each leaf node contains at least the *minimum leaf size* number of observations. The splitting procedure is repeated until one of the following criteria is reached:

1. the node contains a number of observations lower than a specified threshold (*minimal size for split*);
2. the spit does not produce a gain, in terms of impurity, greater than a fixed *minimal gain*;
3. the *maximal depth* of the tree is reached.

However, through this method, too complex Decision Trees are often generated, or they might overfit data, therefore pruning techniques are used [20,22,45]. Pruning is a strategy to simplify a tree and, in the adopted algorithm, it follows a pessimistic method, evaluating the range of expected error rates at each leaf node, based on a specified *confidence level* [22]. In this paper, Decision Trees were generated adopting the following parameters values (Table 1).

Table 1. Model parameters

Parameter	Values
Stopping criterion	Gini Index
Minimal size for split	20
Minimal leaf size	20
Minimal gain	0.5
Maximal depth	10
Confidence level	0.25

Decision Trees were calibrated and validated using data from a mobility survey carried out in the Turin metropolitan area, which includes the Municipality of Turin, with about 800,000 inhabitants and 23 traffic analysis zones, and the municipalities surrounding the city, with about 544,000 inhabitants and 31 traffic analysis zones. The Turin area has 6736 inhabitants per square kilometer. About 20% of the trips are performed during the morning peak period (from 8 to 10 o'clock) and about 23% during the evening peak period (from 16 to 19 o'clock). Moreover, most of the trips are unimodal trips (93%). About 48% of trips are performed by private car, 24% by public transport, 23% on walking and by bike, 3% by motorbike and 2% by taxi. The public transport ticket cost is 1.70€ for the urban area; the car sharing fare depends on the operator, and it is about 0.21 €/min, on average. The survey was administered to a representative sample of the population living in the study area. This sample was obtained through a stratified random sample technique, where strata were created according to gender, age, occupational status and traffic analysis zone where the individual lives.

The survey was structured through the following sections:

1. A brief introduction with preliminary screening questions (gender, age, occupation and zone) to understand which stratum the interviewee belongs to;
2. A travel diary and related activity patterns spanning over the 24 hours before the interview. Activity locations were entered and geocoded through Google Maps APIs, in order to better estimate duration of activities, travel times and covered distances, considering the specific transport modes used to reach each place;
3. Detailed questions (e.g. travel times with all means, walk and wait times, travel contingencies, info on vehicles, on-trip activities, and attitudinal questions) were posed about a randomly selected trip chain, among those reported by the

respondent in the travel diary. This chain was generated as a sequence of activities in order to help the respondents in focusing on a trip chain that makes sense to them [48,49];

4. Stated-preference experiments were administered to investigate mode switching attitudes for the selected trip chain, where the alternatives were "stay with the current means" and "switch to the proposed new means" if the same trip chain had to be performed in the future. In particular, following the approach presented by Diana [48,49], each respondent had to express her shifting likelihood on a 5 points scale ranging from "very unlikely" to "very likely", and the switching mode was one among car as driver, car sharing, bike, bike sharing, public transport, and a kind of shared taxi. Each respondent faced six binary choice tasks. The trip attributes of each mode were: costs (public transit ticket or subscription, tolls, fuel and parking fares), in-vehicle time, walking time to reach the public transit stop and waiting time at the stop (for public transport modes), or walking time to reach the parked vehicle (for car, car sharing and bike sharing). Attributes of the current mode were estimated by directly considering the characteristics of the trip reported by the interviewee in the previous sections ("RP attributes" in the following). Attributes of the alternative mode ("SP attributes" in the following) were evaluated by processing information on the trip chain (e.g. geographical locations, departure/arrival times) through Google Maps APIs, additionally integrating information related to public transit operations, car sharing and bike sharing services (namely, fares and subscription costs), along with average cost of fuels. Thus, the experiment is based on a real trip with realistic characteristics of the switching mode, thus increasing the realism and, therefore, the reliability of the answers. For each attribute, an upper, base and lower level were calculated for SP experiments. An orthogonal fractional factorial design with 18 questions in 3 blocks was generated;
5. Socio-economic questions both at individual and household level.

The same survey was administered through both CATI (Computer Assisted Web Interviewing) and CAWI (Computer Assisted Telephone Interviewing) protocols 7 days a week in three different 4-weeks periods, to control for seasonal effects. Data obtained from the three waves were aggregated obtaining 3280 valid interviews.

Since the algorithms predict the switch from the current mode to car sharing, only SP experiments in which one of the alternative is car sharing were retained. Furthermore, for each of them, neutral answers were discarded; the remaining observations were grouped in "shift" and "no-shift" answers.

Then, five Decision Trees were generated, one for each current mode (car as driver, car as passenger, public transport, bike and walking), which predict switching intention towards car sharing. Since the number of observation in the two classes of answers was highly imbalanced towards the negative switch, a SMOTE (Synthetic Minority Over-Sampling Technique) was adopted [19,50,51], which oversamples elements belonging to the minority class introducing synthetic examples [50]. In order to calibrate, validate and calculate model performances, a cross-validation procedure was applied [52]. In order to evaluate the performances of each Decision Tree, three indexes were considered:

1. Confusion matrix, i.e. a matrix containing counts of observations belonging to predicted and actual classes (class precision and class recall);
2. Model accuracy, i.e. the relative number of correctly classified observations;
3. Area Under the Curve (AUC) of the Receiver Operating Characteristics (ROC), plotting the True Positive Rate versus False Positive Rate in order to evaluate the distance between model classification output and random guess.

Variables considered in the switch models were: starting time of the trip (START_HOUR), and travel time (DELTA_TIME in minutes), distance (DELTA_DIST in kilometers) and cost (DELTA_COST in Euros), which were calculated as the difference between the SP attribute of the alternative mode (i.e. car sharing) and the RP attribute of the current one.

3 Results

Results from all five Decision Trees are presented, but only two figures are reported for the sake of brevity. In each node of a Decision Tree, a value of 1 indicates a shift result, whereas 0 indicates a no-shift. Colored bars show the distribution of outputs in each node, where the blue and red parts represent the number of observations with a positive and negative shift, respectively.

Performance measures for each Decision Tree obtained from the calibration procedure are reported in Table 2.

Table 2. Model performances

		CD	CP	PT	B	W
Precision	1	72.1%	69.1%	76.4%	53.6%	74.9%
	0	67.3%	70.4%	74.2%	64.5%	80.9%
Recall	1	47.9%	68.7%	70.8%	57.7%	81.3%
	0	85.3%	70.9%	79.3%	60.6%	74.6%
Accuracy		68.7%	69.8%	75.2%	59.4%	77.8%
AUC		0.711	0.749	0.806	0.569	0.833

Notes: CD: car as driver; CP: car as passenger; PT: public transport; B: bike; W: walking.

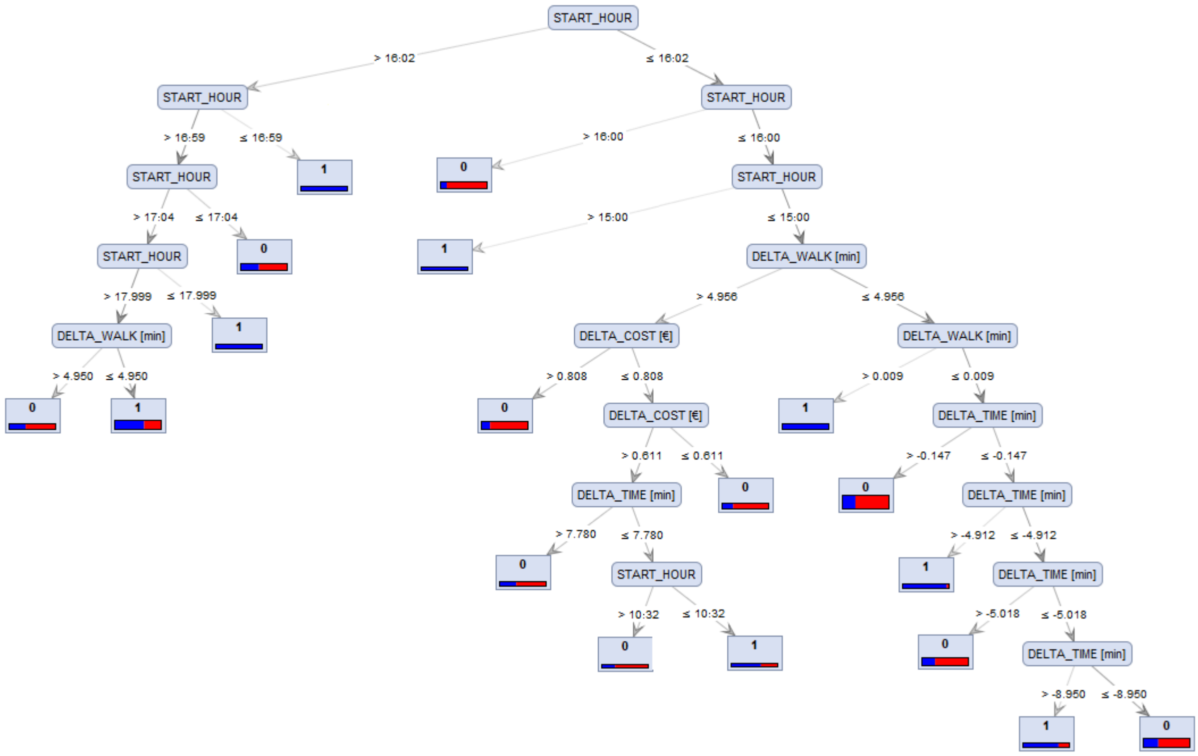

Figure 1. Decision Tree for car drivers

The first Decision Tree models the switch from car as driver (Figure 1). The most important variable is the starting time of the trip, with a threshold for the split of the first leaf estimated at 16 o'clock. In order to better understand the meaning of this threshold, the original dataset was analyzed pointing out that this time value divides the distribution of activities into two parts. In particular, the majority of trips starting after 16 o'clock (about 87%) are Home-Based trips, whereas trips starting before 16 have a distribution of activities more heterogeneous: 38% are Home-Based trips, 33% are trips for mandatory activities (work or school) and 29% for discretionary activities. This indicates that locations where discretionary activities take place are reached by car drivers in the late afternoon or in the evening (i.e. after 16 o'clock). In the whole tree there are some time windows in which car sharing is preferred, above all between 15 and 17 o'clock, in which trips to come back home are often carried out. The second most important variable is the walking time to reach the vehicle; in particular, about 65% of respondents are willing to adopt car sharing if vehicles are placed up to 5 minutes far from the users. On the other hand, travel time is not so important for the switching decision since this variable appears only on the bottom part of the tree and since negative switching intentions are reported even when travel time on car sharing is lower than the one of car. Moreover, concerning the cost of car sharing trips, negative switches are obtained for most of the cases (about 70%). In conclusion, car drivers might adopt car sharing if interventions to reduce the walking time to reach the vehicles are carried out. Other policies about reducing either the travel time or the cost of car sharing are likely to be not so effective.

The Decision Tree of switching intentions of people travelling by car as passengers is quite simple and similar to the one for car drivers. In particular, as for car drivers, this tree shows that users are willing to switch to car sharing if the walking time to reach the vehicle is less than 5 minutes. However, unlike car drivers, passengers might adopt car sharing if the cost of the trip is lower (about 63%).

The third Decision Tree considers public transport. Like for car drivers, the time to reach the vehicle is the most important variable that affects the switching decision. Moreover, the same threshold of 5 minutes is obtained, indicating that this is a common value for mode choices among travelers. On the other hand, if the travel time and cost of a car sharing trip are lower than the corresponding ones of public transport, users are willing to adopt the former mode, as expected. Some users (about 57%) might switch to car sharing even if the in-vehicle travel time is greater, but most of the positive switches are reported only if the cost of the trip is lower (about 75%). This suggests that travel cost is a key factor that affects the switch from public transport to car sharing, rather than travel time.

The fourth Decision Tree concerns cycling (Figure 2). Trips starting after 16 o'clock are more likely to be performed on car sharing, probably because of the higher comfort and safety provided by shared vehicles, especially during evening hours. For the other trips, it is possible to identify thresholds in terms of time and distance for which bikers are willing to adopt car sharing. In particular, trips longer than about 200 meters and with a travel time reduction ranging from 2 to 10 minutes are going to be replaced by car sharing. Furthermore, in this case, travelers are going to switch towards car sharing independently on the cost of the service, since the corresponding variable is not present in the DT.

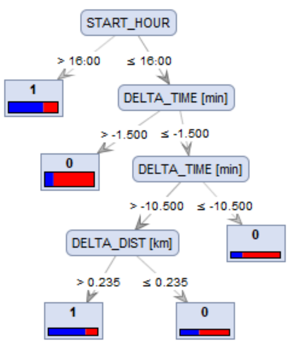

Figure 2. Decision Tree for cycling

The last Decision Tree evaluated the shift from walking to car sharing. The most important variable is DELTA_COST, since users are more likely to perceive this factor, because of the null price for walking. Results indicate that trips shorter than 100 meters cannot be substituted by car sharing, as expected, since these trips are likely to have a longer duration on shared car, due to urban congested streets. Moreover, walkers are willing to pay 0.3€ at maximum to reduce their travel time by at least 4 minutes. Independently on the travel time, users would pay up to 0.6€ for trips starting before 9 o'clock. Furthermore, travellers are willing to pay and to accept more travel time for trips starting after 17 o'clock, suggesting that during evening and night periods walkers prefer car sharing, probably because of comfort and safety.

4 Conclusions

In this paper, values of trip attributes that affect the choice to switch from a transport mode to car sharing are estimated through Decision Trees. Models were calibrated with data from a mobility survey carried out in the Turin Metropolitan area (Italy). In this survey, respondents had to face with Stated-preference experiments, where they had to declare their intention to perform a real trip with car sharing. In order to consider the effect of the currently adopted mode, one model was estimated for each of the current mode, i.e. car as driver, car as passenger, public transport, bike and walking.

Results are helpful to authorities who can target specific interventions to promote the switch from private modes to car sharing, avoiding the shift from public transport and active means (bike and walking). Considering private car drivers, the corresponding DT shows that the shift can occur between 15 and 17 o'clock for returning home trips. Therefore local authorities can foster the adoption of car sharing by providing reserved parking near activities location, where users can easily find available shared vehicles to come back home, for example. Moreover, results indicate that drivers are willing to walk up to 5 minutes to reach a shared car, with flexibility similar to the one of a private car. Consequently, policy makers should promote targets about the capillarity of shared vehicles, to ensure a walking time to reach the vehicle at least similar to that of private cars. On the other hand, travel time is found to be not important for the shift, therefore interventions to reduce the trip duration for car sharing (such as giving free access to public transport reserved lanes) are not essential. Furthermore the majority of car drivers reports negative switching intentions even if the cost of car sharing is lower, thus policies to reduce car sharing costs do not produce strong effects on them.

On the contrary, public transport users are found to be more likely to switch if the cost of car sharing is lower, independently on the travel time. From this perspective, policy makers should pay attention before providing subsidies or tax cuts to car sharing operators, especially if these benefits would reduce the fares of the service compared to those of public transport. Moreover, as for car drivers, even public transport users are willing to walk up to 5 minutes. Therefore interventions to promote the diffusion of shared vehicles could induce shifts also from public transport. In order to compensate for this negative effect, local authorities should pay specific attention to the accessibility of public transport to increase the attractiveness of public transit respect to car sharing.

The Decision Tree for bike trips suggests that users can switch if car sharing travel time is lower (from 2 to 10 minutes), or for trips starting after 16 o'clock. Moreover, their switching intentions are independent of travel costs. Results for walking trips indicate that no shift could occur for short trips (up to 100 meters). In addition, travelers are willing to pay to reduce their travel time, and positive switching intentions are also reported for limited car sharing fares (up to 0.6€).

By analyzing these results, the previously explained interventions can positively manage the shift from bike and walking. In particular, taking no actions to reduce travel time, which is ineffective for both drivers and public transport users, also allows preventing switches from bike and walking to car sharing. Moreover, like for transit, policies to induce operators to decrease car sharing costs could increase the shift from walking and, therefore, they should be avoided. Furthermore, thresholds provided by the Decision Tree could be useful to set quantitative targets to policies on car sharing. Lastly, the public transport system should be strengthened during the evening and in night hours, in order to compete with car sharing in these periods, thus reducing the adoption of car sharing for both bikers and walkers.

To sum up, though this approach, sound basis for policies to manage the shift towards car sharing are derived. In this way, policy makers can adopt coherent interventions to promote the switch from private car, simultaneously preventing the switch from public transport and active means.

The present research should be extended on some points to give a better overall picture. In particular, socio-economic characteristics of users are an obvious determinant of modal choices but were not considered so far. However such analysis was not yet developed, since at the present stage the variables under investigation were selected according to the perspective of the decision makers, who can control only some of the modal attributes of a trip to properly address travel demand.

Acknowledgements

This study was partly sponsored by the European project "Shared mobility opporTunities And challenges foR European citieS" (STARS), which has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement no. 769513. Financial support also came through a "Ricerca dei Talenti" grant from Fondazione CRT.

References

[1] Regina R. Clewlow, "Carsharing and sustainable travel behavior: Results from the San Francisco Bay Area," *Transport Policy* 51 (2016): 158–164, doi:10.1016/j.tranpol.2016.01.013.

[2] Henrik Becker, Francesco Ciari, and Kay W. Axhausen, "Measuring the car ownership impact of free-floating car-sharing – A case study in Basel, Switzerland," *Transportation Research Part D: Transport and Environment* 65.August (2018): 51–62, doi:10.1016/j.trd.2018.08.003.

[3] Elliot Martin and Susan Shaheen, *Impacts of car2go on Vehicle Ownership, Modal Shift, Vehicle Miles Travelled, and Greenhouse Gas Emissions: An Analysis of Five North American Cities*, 2016.

[4] Elliot Martin and Susan A. Shaheen, "Greenhouse gas emissions impacts of carsharing in North America," *Transactions on Intelligent Transportation Systems* 12.4 (2011): 1–114, doi:10.1109/TITS.2011.2158539.

[5] Frances Sprei et al., "Free-floating car-sharing electrification and mode displacement: Travel time and usage patterns from 12 cities in Europe and the United States," *Transportation Research Part D: Transport and Environment* (2018), doi:10.1016/j.trd.2018.12.018.

[6] Riccardo Ceccato and Marco Diana, "Substitution and complementarity patterns between traditional transport means and car sharing: a person and trip level analysis," *Transportation* (2018), doi:10.1007/s11116-018-9901-8.

[7] Henrik Becker, Francesco Ciari, and Kay W. Axhausen, "Comparing car-sharing schemes in Switzerland: User groups and usage patterns," *Transportation Research Part A: Policy and Practice* 97 (2017): 17–29, doi:10.1016/j.tra.2017.01.004.

[8] Páraic Carroll, Brian Caulfield, and Aoife Ahern, "Examining the potential for car-shedding in the Greater Dublin Area," *Transportation Research Part A: Policy and Practice* 106.November (2017): 440–452, doi:10.1016/j.tra.2017.10.019.

[9] Uta Burghard and Elisabeth Dütschke, "Who wants shared mobility? Lessons from early adopters and mainstream drivers on electric carsharing in Germany," *Transportation Research Part D: Transport and Environment* 71.June 2018 (2019): 96–109, doi:10.1016/j.trd.2018.11.011.

[10] Elliot Martin, Susan Shaheen, and Jeffrey Lidicker, "Impact of Carsharing on Household Vehicle Holdings," *Transportation Research Record: Journal of the Transportation Research Board* 2143 (2010): 150–158, doi:10.3141/2143-19.

[11] Dimitrios Efthymiou, Constantinos Antoniou, and Paul Waddell, "Factors affecting the adoption of vehicle sharing systems by young drivers," *Transport Policy* 29 (2013): 64–73, doi:10.1016/j.tranpol.2013.04.009.

[12] Jinhee Kim, Soora Rasouli, and Harry Timmermans, "Satisfaction and uncertainty in car-sharing decisions: An integration of hybrid choice and random regret-based models," *Transportation Research Part A: Policy and Practice* 95 (2017): 13–33, doi:10.1016/j.tra.2016.11.005.

[13] Maria Juschten et al., "Carsharing in Switzerland: identifying new markets by predicting membership based on data on supply and demand," *Transportation* (2017): 1–24, doi:10.1007/s11116-017-9818-7.

[14] Scott Le Vine, Orestes Adamou, and John Polak, "Predicting new forms of activity/mobility patterns enabled by shared-mobility services through a needs-based stated-response method: Case study of grocery shopping," *Transport Policy* 32 (2014): 60–68, doi:10.1016/j.tranpol.2013.12.008.

[15] Henrik Becker, Francesco Ciari, and Kay W. Axhausen, "Modeling free-floating car-sharing use in Switzerland: A spatial regression and conditional logit approach," *Transportation Research Part C: Emerging Technologies* 81 (2017): 286–299, doi:10.1016/j.trc.2017.06.008.

[16] Scott Le Vine et al., "A new approach to predict the market and impacts of round-trip and point-to-point carsharing systems: Case study of London," *Transportation Research Part D: Transport and Environment* 32 (2014): 218–229, doi:10.1016/j.trd.2014.07.005.

[17] Marion Lagadic, Alia Verloes, and Nicolas Louvet, "Can carsharing services be profitable? A critical review of established and developing business models," *Transport Policy* 77.October 2018 (2019): 68–78, doi:10.1016/j.tranpol.2019.02.006.

[18] Daniel McFadden, "Conditional Logit Analysis of Qualitative Choice Behavior," in *P. Zarembka (Ed.), Frontiers in econometrics, Academic Press, New York*, (1974), 105–142.

[19] Liang Tang, Chenfeng Xiong, and Lei Zhang, "Decision tree method for modeling travel mode switching in a dynamic behavioral process," *Transportation Planning and Technology* 38.8 (2015): 833–850, doi:10.1080/03081060.2015.1079385.

[20] Toshiyuki Yamamoto, Ryuichi Kitamura, and Junichiro Fujii, "Drivers' Route Choice Behavior: Analysis by Data Mining Algorithms," *Transportation Research Record: Journal of the Transportation Research Board* 1807.1 (2007): 59–66, doi:10.3141/1807-08.

[21] Ch Ravi Sekhar, Minal, and E. Madhu, "Mode Choice Analysis Using Random Forrest Decision Trees," *Transportation Research Procedia* 17.December 2014 (2016): 644–652, doi:10.1016/j.trpro.2016.11.119.

[22] Chi Xie, Jinyang Lu, and Emily Parkany, "Work Travel Mode Choice Modeling with Data Mining: Decision Trees and Neural Networks," *Transportation Research Record: Journal of the Transportation Research Board* 1854.1 (2007): 50–61, doi:10.3141/1854-06.

[23] Elke Moons, Geert Wets, and Marc Aerts, "Nonlinear Models for Determining Mode Choice," *Progress in Artificial Intelligence* (2007): 183–194, doi:10.1007/978-3-540-77002-2_16.

[24] Julian Hagenauer and Marco Helbich, "A comparative study of machine learning classifiers for modeling travel mode choice," *Expert Systems with Applications* 78 (2017): 273–282, doi:10.1016/j.eswa.2017.01.057.

[25] Xuewu Chen et al., "Applying a random forest method approach to model travel mode choice behavior," *Travel Behaviour and Society* 14.May 2018 (2018): 1–10, doi:10.1016/j.tbs.2018.09.002.

[26] Anabele Lindner, Cira Souza Pitombo, and André Luiz Cunha, "Estimating motorized travel mode choice using classifiers: An application for high-dimensional multicollinear data," *Travel Behaviour and Society* 6 (2017): 100–109, doi:10.1016/j.tbs.2016.08.003.

[27] Kenneth E. Train, "Discrete choice methods with simulation," *Discrete Choice Methods with Simulation* 9780521816 (2003): 1–334, doi:10.1017/CBO9780511753930.

[28] Zheng Zhu et al., "A mixed Bayesian network for two-dimensional decision modeling of departure time and mode choice,"

Transportation 45.5 (2018): 1499–1522, doi:10.1007/s11116-017-9770-6.

[29] Xinyi Wang and Sung Hoo Kim, "Prediction and Factor Identification for Crash Severity: Comparison of Discrete Choice and Tree-Based Models," *Transportation Research Record* (2019), doi:10.1177/0361198119844456.

[30] C. S. Pitombo, E. Kawamoto, and A. J. Sousa, "An exploratory analysis of relationships between socioeconomic, land use, activity participation variables and travel patterns," *Transport Policy* 18.2 (2011): 347–357, doi:10.1016/j.tranpol.2010.10.010.

[31] Rui Zhang, Enjian Yao, and Zhili Liu, "School travel mode choice in Beijing, China," *Journal of Transport Geography* 62.June (2017): 98–110, doi:10.1016/j.jtrangeo.2017.06.001.

[32] Cira Souza Pitombo et al., "A two-step method for mode choice estimation with socioeconomic and spatial information," *Spatial Statistics* 11 (2015): 45–64, doi:10.1016/j.spasta.2014.12.002.

[33] Jean-Claude Thill and Aaron Wheeler, "Tree Induction of Spatial Choice Behavior," *Transportation Research Record: Journal of the Transportation Research Board* 1719.1 (2007): 250–258, doi:10.3141/1719-33.

[34] Li Yen Chang and Wen Chieh Chen, "Data mining of tree-based models to analyze freeway accident frequency," *Journal of Safety Research* 36.4 (2005): 365–375, doi:10.1016/j.jsr.2005.06.013.

[35] Nadezda Zenina, Yuri Merkuryev, and Andrejs Romanovs, "Transport travel demand simulation model development for mixed-use building," *Proceedings of the 5th IEEE Workshop on Advances in Information, Electronic and Electrical Engineering, AIEEE 2017* 2018-Janua (2018): 1–6, doi:10.1109/AIEEE.2017.8270557.

[36] Fangru Wang and Catherine L. Ross, "Machine Learning Travel Mode Choices: Comparing the Performance of an Extreme Gradient Boosting Model with a Multinomial Logit Model," *Transportation Research Record* (2018), doi:10.1177/0361198118773556.

[37] Paul Waddell and Arezoo Besharati-Zadeh, "A comparison of statistical and machine learning algorithms for predicting rents in the San Francisco Bay Area," *Transportation Research Board 98th Annual Meeting* (2019).

[38] Shenhao Wang and Jinhua Zhao, "An Empirical Study of Using Deep Neural Network to Analyze Travel Mode Choice with Interpretable Economic Information," *Transportation Research Board 98th Annual Meeting. 13-17 Janaury. Washington DC, United States* (2018).

[39] L. O. Oral and V. Tecim, "Using Decision Trees for Estimating Mode Choice of Trips in Buca-Izmir," *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XL-4/W1.1982 (2013): 139–145, doi:10.5194/isprsarchives-xl-4-w1-139-2013.

[40] Theo Arentze and Harry Timmermans, "Parametric action decision trees: Incorporating continuous attribute variables into rule-based models of discrete choice," *Transportation Research Part B: Methodological* 41.7 (2007): 772–783, doi:10.1016/j.trb.2007.01.001.

[41] L.; Breiman et al., *Classification and Regression Trees* (New York: Routledge, 1984).

[42] Theo Arentze and Harry J. P. Timmermans, "Modeling learning and adaptation processes in activity-travel choice," *Transportation* 2003 (2003): 37–62.

[43] Theo A. Arentze and Harry J. P. Timmermans, "A learning-based transportation oriented simulation system," *Transportation Research Part B: Methodological* 38.7 (2004): 613–633, doi:10.1016/j.trb.2002.10.001.

[44] Stuart J. Russell and Peter Norvig, *Artificaial Intelligence. A modern Approach* (Pearson, 2002).

[45] Geert Wets et al., "Identifying Decision Structures Underlying Activity Patterns. An Exploration of Data Mining Algorithms," *Transportation Research Record: Journal of the Transportation Research Board* 1718.1 (2000): 1–9, doi:https://doi.org/10.3141/1718-01.

[46] Yun Wang et al., "Individuals' acceptance to free-floating electric carsharing mode: A web-based survey in China," *International Journal of Environmental Research and Public Health* 14.5 (2017), doi:10.3390/ijerph14050476.

[47] J. Ross Quinlan, *C4.5 Programs for Machine Learning* (San Mateo, California: Morgan Kaufmann Publishers, 1993).

[48] Marco Diana, "Making the 'primary utility of travel' concept operational: A measurement model for the assessment of the intrinsic utility of reported trips," *Transportation Research Part A: Policy and Practice* 42.3 (2008): 455–474, doi:10.1016/j.tra.2007.12.005.

[49] Marco Diana, "From mode choice to modal diversion: A new behavioural paradigm and an application to the study of the demand for innovative transport services," *Technological Forecasting and Social Change* 77.3 (2010): 429–441, doi:10.1016/j.techfore.2009.10.005.

[50] Alberto Fernandez et al., “SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary,” *Journal of Artificial Intelligence Research* 61 (2018): 863–905, doi:10.1613/jair.1.11192.

[51] Nitesh V. Chawla et al., “SMOTE: Synthetic Minority Over-sampling Technique,” *Journal of Artificial Intelligence Research* 16 (2002): 321–357, doi:10.1613/jair.953.

[52] R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2* (1995): 1137–1143, doi:10.1067/mod.2000.109031.