Modelling the propensity in adhering to a carsharing system: a behavioral approach

Stefano de Lucaa, Roberta Di Pacea

*aUniversity of Salerno, Department of Civil Engineering, Via Giovanni Paolo 132, 84084 Fisciano (SA), Italy

Abstract

In this paper carsharing behavior was investigated with regard to a different and little investigated aspect of users’ behavior: modeling the propensity in adhering to a carsharing system. The propensity was modeled within the random utility framework through starting from a stated preferences survey. The main focus was on the investigation and estimation of a set of attributes able to interpret and measure the propensity. In particular, together with socio-economic and activity related attributes, the satisfaction variable (expected maximum utility) was tested in order to interpret the interest towards carsharing in light of the level of service supplied by the other transport modes and of the users’ socio-economic and activity-based characteristics. The satisfaction variable was specifically calibrated on actual mode choice behavior starting from revealed preferences data.

Keywords: carsharing; propensity; random utility theory; stated preferences; satisfaction/logsum variable.

1. Introduction

Carsharing has become popular in the early 90s and nowadays can rely on several consolidated projects in more than 600 cities in more than 18 countries. Most of the existing projects have highlighted that users’ socio-economic characteristics (age, education, income) and the geographical context of each city play a major role in a successful implementation and growth of any carsharing program (Burkhart and Millard-Ball, 2006; Shaheen and Cohen, 2007; Shaheen et al., 2004).

* Corresponding author. Tel.: +39-089-964-122; fax: +39-089-964-233.
E-mail address: sdeluca@unisa.it
In this context, most of the cited studies on carsharing are mainly concentrated in North America, and are focused primarily on the feasibility of carsharing programs and on the impact of carsharing on auto ownership and vehicle usage. In all the contributions carsharing has been interpreted as a service supplied in an urban context and, usually, as an alternative to the car transportation mode. Less investigated are carsharing programs on an extra-urban scale and as a transport mode, which are complimentary to public transport. Finally, although carsharing may rely on several case studies and is relatively consolidated in many cities, it should be noted that not many attempts to model choice behaviors exist in literature. Such an issue has most frequently been addressed through focus groups and/or analyses of real data or revealed preferences surveys (Walb and Loudon, 1986; Doherty et al., 1987; Steininger et al., 1996; Shaheen, 1999; Shaheen et al., 1999; Shaheen and Wright, 2001; Katzev, 2003; Huwer, 2004; Rodier and Shaheen, 2004; Vance et al., 2005; Lane, 2005; Shaheen and Rodier, 2005; Burkhardt and Millard-Ball, 2006; Celsor and Millard-Ball, 2007; Morency et al., 2007; Martin et al., 2010; Martin and Shaheen; 2011; Zhou, 2012), through econometric models based on RP data (Morency et al., 2009; Habib et al., 2012; Costain et al., 2012; Morency et al., 2012) and through descriptive and modelling approaches based on stated preferences data (Fukuda et al., 2005; Firnkorn and Muller, 2011; Huwer, 2004; Zheng et al., 2009; Nobis, 2006; Cervero et al., 2007; Stillwater et al., 2007; de Luca and Di Pace, in press).

Among modelling approaches, different solutions have been tested in order to predict changes in individual car ownership, mode choice and carsharing usage: logistic regressions (Shaheen, 1999), binary Logit models (Cervero, 2003 and 2007) to predict the use of carsharing; Multinomial Logit models to predict the likelihood of choosing carsharing as a travel mode among other travel modes and Ordered Probit models to examine factors influencing people’s acceptance of carsharing (Zhou et al., 2008).

The choice to adhere to a carsharing system may be interpreted as a choice process in which the users, before choosing carsharing as an alternative mode to travel, (i) acquire information on the carsharing system, (ii) develop interest in the system and finally (iii) choose on the basis of service characteristics. With regard to this conceptual framework, in this paper a different and little investigated aspect of users’ behavior was addressed: the propensity in adhering to a carsharing system.

The motivation for modeling the “interest” is twofold:

(i) to simulate the demand segment (the probability) interested in a carsharing system, in fact, models simulating carsharing choice behavior are only usually applied to the users interested in the alternative;

(ii) to investigate the determinants of the phenomenon; in fact, one of the main issues for a decision maker is understanding how to attract users and how to make the system more appealing.

Such issues were addressed through the specification and calibration of different modeling solutions founded on the behavioral paradigm of the utility theory (Domenich and McFadden, 1975). The propensity was modeled through a Binomial Logit model following a stated preferences survey and the main focus was on the investigation and estimation of a set of attributes able to interpret and measure the propensity. In particular, the satisfaction variable (expected maximum utility), specifically calibrated on actual mode choice behavior, was tested in order to interpret the propensity towards carsharing in light of the level of service supplied by the other transport modes and of the users’ socio-economic and activity-based characteristics. To this aim a set of mode choice models was specified and calibrated on revealed preferences data.

Elements of originality are twofold: (i) methodological: the propensity has not yet been explicitly investigated in literature, the analyses were based on stated preferences and taste variation among users was investigated; (ii) operational: some useful insights to support decision makers were drawn.

The paper is organized as follows: the methodological framework is proposed in section 2; the case study and the survey are briefly described in section 3; estimation results and conclusions are reported in sections 4 and 5.
2. Methodological framework

The propensity in adhering to a carsharing system was addressed through the specification and calibration of discrete choice models founded on the behavioral paradigm of the utility theory (Domencich and McFadden, 1975). In particular, the propensity was modeled as a binomial choice problem, where the alternatives were: “being interested” or “not being interested”. Homoscedastic and Heteroscedastic modelling solutions were investigated. In particular, the Binomial Logit model and the random coefficient formulation were specified.

The Binomial Logit (BNL) model is based on the assumption that the random residuals $\varepsilon_j$ are independently and identically distributed (iid) according to a Gumbel random variable of zero mean and parameter $T$. Under the assumptions made, the probability of choosing alternative one of the two alternatives can be expressed in closed form as:

$$p[j] = \frac{\exp(V_j / \theta)}{\sum_{m \in M} \exp(V_m / \theta)}$$

Where $V_j$ is the systematic utility of the generic alternative $j$, and $\theta$ is the parameter of the Gumbel probability distribution function assumed for perceived utilities (often known as scale parameter).

The chosen heteroscedastic modelling solution was the random parameter formulation. In this case, the utility of each decision-maker is specified as

$$U_j = V_j + \varepsilon_j = \left[ \sum_k \beta_k X_{kj} + \sum_h \gamma_h X_{hj} \right] + \varepsilon_j$$

where $X_{kj}$ are the explanatory attributes that relate to the alternative and decision-makers introduced before, $\beta_k$ is the generic coefficient of attribute $k$, $\gamma_h$ is the generic coefficient of attribute $h$ representing the decision-maker’s taste that is supposed to be distributed randomly with density $f(\gamma_h)$, $\varepsilon_j$ is the random term that is an iid Gumbel random variable of zero mean and parameter $\theta$. Assuming that $\gamma$ and $\beta$ are the vectors of parameters, Mixed Logit probabilities are the integrals of standard Logit probabilities over a density of parameters.

$$p[j] = \int \frac{\exp \left[ V_j \left( \tilde{\beta} / \theta, \tilde{\gamma} / \theta \right) \right]}{\sum_{m \in M} \exp \left[ V_m \left( \tilde{\beta} / \theta, \tilde{\gamma} / \theta \right) \right]} \cdot f(\tilde{\gamma} / \theta) \cdot d(\tilde{\gamma} / \theta)$$

If the choice function is well known and has been extensively applied for the last three decades, the main focus was on the specification of the systematic utility functions and, in particular, on the investigation and estimation of a set of attributes which are able to interpret and measure the propensity in adhering to a carsharing system.

Apart from the typical socio-economic and level-of-service attributes, specific attributes were tested in order to interpret the “propensity” in light of the accessibility currently supplied by the existing (and available) transport modes. In particular, it was assumed that the propensity in adhering to a carsharing system is mainly related to: the number of available transport modes, their real availability, the supplied level of service, the activity to be carried out at the destination and the users’ socio-economic attributes.

In this paper, it was assumed that the propensity towards a carsharing system is related to the determinants that affect the current transport mode choice process. To this aim, within the behavioral paradigm of the random utility theory (Cascetta, 2009), the satisfaction variable related to the current transport mode choices was specified and specifically estimated. Such variable represents the current users’ “satisfaction” towards the current transport modes and it is a sort of measure of the actual supplied accessibility (Ben-Akiva and Lerman, 1985).
As is well known, the random utility theory is based on the assumption that the generic decision-maker chooses from the available choice set, the alternative $j$ with maximum perceived utility $U_j$, where the perceived utilities are modeled as random variables. The variable $U_j$, therefore, denotes the perceived utility “obtained” by the decision-maker in the choice context ($I$). This variable is not observed by the analyst because it is the maximum value of unobserved perceived utilities. Therefore $U_j$ can be modeled as a random variable and it is usually expressed as the sum of two terms: a systematic utility and a random residual. The systematic utility $V_j$ represents the mean (expected value) utility perceived by all decision-makers having the same choice context (alternatives and attributes). The random residual $H_j$ is the (unknown) deviation of the utility perceived by user $i$ from this mean value; it captures the combined effects of the various factors that introduce uncertainty into choice modeling.

In the same manner, the generic user associates a utility to a given choice context that he/she faces. Such a utility, usually called expected maximum perceived utility ($s$) or satisfaction, is defined as the expected value of $U_j$ over the alternatives available in the choice set, vector $U$. Expected maximum perceived utility is a function of the systematic utilities of all the alternatives, vector $V$, and it depends on the joint probability density function of the random residuals, $f(H)$, as well as on the composition of the choice set $I$. Depending on the assumptions on $f(H)$ can be expressed in a closed-form or can be only estimated through simulation techniques.

With regard to this theoretical framework, the expected maximum perceived utility was estimated with regard to the choice set that users habitually face when travelling between the considered origin-destination pairs. To this aim, a specific set of random utility models was specifically calibrated. In particular, closed-form random utility choice models were preferred in order to obtain a closed-form expression for the expected maximum perceived utility. As matter of fact, random residuals were independently distributed Gumbel variables with the same scale parameter $\theta$. In this case the maximum of a set of such random variables is also distributed as a Gumbel variable with parameter $\theta$ and the mean may be expressed in a closed-form.

$$s_j = \theta \ln \sum_{j=1}^{I} \exp (V_j / \theta) = \theta Y_k$$

where $Y_k$ is the corresponding, well known, logsum variable (because of its analytical form).

3. Case study

The described methodological framework was specified and implemented starting from a stated preferences survey carried out between the municipalities of Salerno and Baronissi (Campania Region, Southern Italy – see figure 1). Salerno is the capital city of the Salerno Province, with about 140,000 residents and characterized by 10,000 daily commuters, whereas Baronissi belongs to the metropolitan area of Salerno, it is 10 km far from Salerno and has 20,000 inhabitants. The two municipalities significantly interact each other (bi-directional flows, distributed over the whole day), and commuting flows mainly travel by car, carpool or bus/train.

In particular, residents commuting between the two municipalities were intercepted at the main sites of each municipality (e.g. stations, squares and offices) and were asked to respond to a survey consisting into two stages. In the first stage, the carsharing system to be implemented was qualitatively introduced and described (e.g. one-way service, dedicated parking slots, free parking and free access to the existent restricted traffic areas) and the respondents were asked to declare the interest toward the system; in the second stage a realistic scenario (all the information introduced before plus operational details such as fares, location of the parking slots, type of car, etc.) was proposed and respondents were asked to declare if they would have chosen the new transport mode.

The survey data were collected from a sample of 200 individuals aged 18 and over, randomly selected to match census data (2011) proportions by gender (male, female), age (18-30; 31-60; >60 years old) and type of occupation.
4. Estimation results

4.1. Satisfaction attribute estimation

The Satisfaction attribute was estimated specifying and calibrating a mode choice model based on individual revealed preferences collected from a specific survey, different from the SP survey introduced before, but carried out on the same case study.

The choice set taken into account consisted of four transport modes: car as driver, car as passenger, car-pool and bus. Car as driver (A) is understood as the transport mode in which the user personally uses the car and bears all travel costs; car as passenger (P) understood as the way of transportation in which the user using a lift, manages to reach its destination by car without facing any travel costs; car-pool (CP) is understood as the transport mode in which two or more users move together in the car sharing travel expenses; bus (B) is the public transport mode.

The specified model (table 1) was based on the random utility theory, and the most effective modelling formulation was found to be the hierarchical Logit model. Indeed, estimation results showed a significant correlation between perceived utilities associated with individual transport but promiscuous modes (CP and P).

The values and signs of the coefficients of systematic utilities are consistent with behavioral assumptions and are all statistically significant. In particular, it was possible to disaggregate the travel time on public transport into a time on board, a waiting time and an access/egress time. The mutual ratios between the coefficients of representative attributes of travel time and the monetary cost attribute coefficient (willingness to pay in €/hour – value of time, V.O.T.) are equal to about 4 Euros for the time on board and 8 Euros for the waiting time and the access/egress time.

Among non-level of service attributes, it is interesting to note that the activity duration at the destination and the weekly trip frequency were statistically significant. As it was logical to expect, durations under 60 minutes penalize the bus and car-pool modes; very long stays (several hours) increase the utility of the car mode. In particular, the attributes Tstop (1,3), Tstop (3,5) and Tstop (> 5) define temporal thresholds that determine the variation of the utility of the car mode as the permanence time of users varies. Interestingly, this variation is not linear and, in general, longer stays increase the utility of the car mode and there is a maximum utility of using the car for permanence times ranging between 3 and 5 hours. The frequency makes it possible to understand how a high weekly trip frequency tends to encourage trips in organized forms of individual transport (car-pool) or rather the use of public transport.
Table 1. mode choice models for estimating satisfaction variable: estimation results

<table>
<thead>
<tr>
<th>Respondents</th>
<th>962</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of attributes</td>
<td>14</td>
</tr>
<tr>
<td>pseudo-$\rho^2$</td>
<td>0.53</td>
</tr>
<tr>
<td>pseudo- $\rho^2$ correct</td>
<td>0.52</td>
</tr>
</tbody>
</table>

**Systematic utilities**

$$V_A = \beta_1 T_{car} + \beta_2 CM_{car} + \beta_6 T_{stop(1,3)} + \beta_7 T_{stop(3,5)} + \beta_8 T_{stop(>5)}$$

$$V_P = \beta_1 T_{car} + \beta_3 PAX$$

$$V_{CP} = \beta_1 T_{car} + \beta_2 CM_{CAR\ POOL} + \beta_{10} Freq + \beta_9 T_{stop(<=1)} + \beta_{11} CarAV + \beta_{14} CAR\ POOL$$

$$V_B = \beta_1 T_{bus} + \beta_2 CM_{bus} + \beta_{12} T_{wait} + \beta_{13} T_{access} + \beta_{14} Gen + \beta_{10} Freq + \beta_9 T_{stop(<=1)}$$

- $T_{CAR}$: travel time by car (h)
- $T_{BUS}$: travel time by Bus (h)
- $T_{WAIT}$: waiting time at the stop (h)
- $T_{ACCESS}$: access time to bus stop (h)
- $CM$: monetary cost (€)
- $Gen$: binary attribute equal to 1 if the user is of female gender
- $Freq$: binary attribute equal to 1 if the trip frequency is smaller than three times
- $CarAV$: car mode availability (number of cars / number of family components)
- $T_{stop(<=1)}$: binary attribute equal to 1 if the stop time is less or equal to 1 hour
- $T_{stop(1,3)}$: binary attribute equal to 1 if the stop time is between 1 and 3 hours
- $T_{stop(3,5)}$: binary attribute equal to 1 if the stop time is between 3 and 5 hours
- $T_{stop(>5)}$: binary attribute equal to 1 if the stop time is greater or equal to 5 hours
- $PAX$: alternative specific attribute
- $\delta$: model parameter, index of the correlation level between the alternatives

"$\dagger$"All coefficient estimates are significantly different to zero with a probability of 95%

Systematic utility functions are completed with socio-economic attributes such as the availability of cars and different users’ characteristics. The gender attribute (being female) value may be linked to fears, to greater discomfort or less willingness to use individual modes. The CarAV attribute indicates, in general, a greater interest in choosing the car-pool mode if you own a car; "groups" are formed between users who have a car and in this way they do not only share costs but also the disadvantages of using their own car (risk of theft and/or accidents, driving stress, drop off).

4.2. Modeling the propensity: estimation results

Specification and calibration of the Binomial Logit model (BNL) and the Mixed Binomial Logit model (MBNL) consisted, as usual, in several trials in which numerous and different attributes were tested. The proposed systematic utility functions are those that led to best goodness-of-fit and to statistically significant estimation results.
The following systematic utility functions turned out statistically significant:

\[ V_{\text{interest}} = \beta_1 \cdot s + \beta_2 \cdot \text{KNOW} + \beta_3 \cdot \text{BUS} + \beta_4 \cdot \text{DIST} + \beta_5 \cdot \text{FREQ} \]

\[ V_{\text{no-interest}} = \beta_{\text{ASC}} \cdot \text{ASC} \]

Where:

- \( s \), it is the satisfaction variable introduced in section 3 and related to the choice-set \( I \).
  
  It was estimated through the specification and calibration of specific mode choice model (introduced in the previous section).
- \( \text{KNOW} \), it is a binary attribute equal to 1 if user declared to be aware of what a carsharing system is and how it works;
- \( \text{BUS} \), it is a binary attribute equal to 1 if user usually travels by Bus;
- \( \text{DIST} \), it is the travelled distance (km);
- \( \text{FREQ} \), it is the number of weekly trips towards the declared destination;
- \( \text{ASC} \), it is the alternative specific constant.

It should be noted that typical socio-economic attributes - such as age, gender, income – were tested but did not turn out statistically significant. This result is, however, interesting since these attributes are significant in most models aiming to interpret and forecast carsharing behavior.

All the systematic utility coefficients were statistically significant with a probability greater than 80% and most of them had a probability greater than 95%. Signs are coherent with the expectations. Estimation results are reported in Table 2.

Comparing the homoscedastic and heteroscedastic formulations, it is worth to note that only one attribute (DIST) turned out normally distributed. Moreover, the BNL model did not outperform the BNL model, pointing out the negligibility of the explicit simulation of taste variation across respondents.

<table>
<thead>
<tr>
<th></th>
<th>BNL</th>
<th></th>
<th>MBNL</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>( s )</td>
<td>+0.765***</td>
<td>-</td>
<td>-0.981***</td>
<td>-</td>
</tr>
<tr>
<td>KNOW</td>
<td>+0.370**</td>
<td>-</td>
<td>+0.674**</td>
<td>-</td>
</tr>
<tr>
<td>BUS</td>
<td>+0.604***</td>
<td>-</td>
<td>+0.552***</td>
<td>-</td>
</tr>
<tr>
<td>DIST</td>
<td>-0.109***</td>
<td>-</td>
<td>-0.0830***</td>
<td>-</td>
</tr>
<tr>
<td>DIST (s.d.)</td>
<td>- 0.0550**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>FREQ</td>
<td>-0.503**</td>
<td>-</td>
<td>-0.585**</td>
<td>-</td>
</tr>
<tr>
<td>ASC</td>
<td>- +5.84***</td>
<td>-</td>
<td>-5.69***</td>
<td>-</td>
</tr>
<tr>
<td>Init log-likelihood</td>
<td>-157</td>
<td>-157</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Final log-likelihood</td>
<td>-47</td>
<td>-46</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Rho-square</td>
<td>0.702</td>
<td>0.704</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Adjusted rho-square</td>
<td>0.664</td>
<td>0.659</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*** results statistically significant with 95% confidence; ** with 85%

Through analyzing the coefficient values, the following conclusions may be drawn.

Satisfaction is the most statistically significant attribute and shows, as expected, a negative effect on being interested. In fact, as satisfaction increases, users perceive a smaller gain in a carsharing system. This result is noteworthy, since it allows us to understand that the perception of carsharing (and the consequent potential demand)
is affected by the actual mode choice behavior, thus by the transport modes available as well as the supplied level of service and users’ socio-economic and activity-based characteristics.

From an operational point of view, as is reasonable, the level of service attributes partially affect market penetration of carsharing that, on the other hand, has to respond to users’ needs and attitudes.

Travelling by bus greatly affects the interest in a carsharing system. In fact, the bus alternative offers less comfort than all the other existing alternatives (including carsharing itself). Moreover, the transit system is less reliable, often crowded and characterized by low frequencies (for some line based < 1 hour). By contrast, carsharing offers the same flexibility and reliability as the car transport mode. As expected, the market segment to address is the bus users segment.

The weekly travel frequency increases the probability of not being interested. The interpretation is threefold: (i) travel costs become incomparable with bus and car transport modes; (ii) the need for booking several days in advance represents a mental obstacle; (iii) users that travel systematically are less interested in an alternative transport mode solution.

The distance represents a sort of spatial impedance. In fact, as distance increases, the probability of being interested decreases, independently from the level of service supplied by the available transport modes.

Users that are further from the final destination are less interested in a carsharing service, meaning that the perception of the advantages of this travel alternative decreases with the distance. The interpretations might be manifold: for longer distances users prefer to always use the same transport mode; for car users that travel for longer distances users prefer to drive the same car, furthermore their own car represents a sort of livable space with all the comforts aimed to make the journey less boring. Bus users travelling for longer distances prefer to travel with the same line, the same bus, the same driver and the same timetable.

The dummy attribute KNOWN showed a positive value. This attribute segments potential users into two classes and highlights that interest, and thus carsharing demand, is affected by the level of knowledge of the service itself and could be greatly increased through specific marketing policies.

Finally, it should be pointed out that the alternative specific constant showed no negligible values.

5. Conclusions

Although carsharing has become a consolidated transport alternative in many urban contexts, carsharing behavior have been mainly investigated through ex-post analysis and in terms of vehicle usage and/or ownership rate. In this paper the propensity in adhering to a carsharing system was investigated in terms of an extra-urban context and through an ex-ante approach based on a stated preferences survey. Behavior was modeled within the random utility paradigm. The aim was: (i) to investigate if modeling the interest was possible and/or reasonable, (ii) to investigate the main determinants of the phenomenon, (iii) to propose a model able to estimate the potential market segment interested in choosing carsharing.

The main insights and conclusions are resumed below.

As regards the modeling approach, a statistically significant mathematical formulation was obtained. The goodness-of-fit was good and the attributes composing the systematic utilities were coherent with the expectations.

One of the most interesting results was the statistical significance of the satisfaction variable (expected maximum utility). This result is noteworthy since it allows for a behavioral interpretation of the phenomenon based on behavioral assumptions. Indeed, the propensity towards carsharing does not only depend on the level of service of the available transport modes, but it also depends on the utility of the available transport modes. As is well known, the utility is not only comprised of the level of service attributes, but also of the socio-economic and activity-based attributes. From an operational viewpoint, such a result suggests that determinants of the actual mode choice behavior should be investigated in order to understand the potential demand for a carsharing system.

Unlike existing insights in literature on carsharing choice behavior, the interest is not directly affected by age, gender or income attributes. Such attributes most probably play a role only when the user has to choose whether to choose to the service, however, on the other hand it should be noted that such attributes concur with the satisfaction variable.

Together with the satisfaction variable, the weekly trip frequency and the travel distance reduce the propensity towards a carsharing system. Therefore, systematic users who travel for longer distances are a challenging segment
to attract. Finally, the interest increases for those users who are already familiar with the service and for those who usually travel by bus. Explicit simulation of taste variation among users, where significant, did not lead to significant gain in the models’ goodness-of-fit.

In conclusion, some research perspectives seem worthy of interest: (i) expand the case study to the whole metropolitan area of Salerno, in order to confirm and validate the obtained results; (ii) simultaneously and explicitly model the interest and the choice of adhere to carsharing service; (iii) simultaneously but implicitly model the interest within the choice of adhere to carsharing service.

References


Firnkorn, J. and M. Müller. What will be the environmental effects of new free-floating car-sharing systems? The case of car2go in Ulm. Ecological Economics 70, 2011, pp. 1519-1528.


