Relations Between Tour Type and Transport Mode Choice: Proposition of Models Estimated by Mixed SP and RP Data

Rosanna Guzzo* Gabriella Mazzulla*

*Dipartimento di Pianificazione Territoriale, University of Calabria
Via P.Bucci, cubo 46/B
87036 Rende, Italy
{r.guzzo, g.mazzulla}@unical.it

1 Introduction

Estimation of transport demand by mathematical models requires a phase of specification and calibration, and a successive phase of validation of these models. In other words, it is necessary to define the variables included and the functional form of the model (*specification*), to estimate the coefficients of the parameters (*calibration*) and to verify the capability to reproduce the data (*validation*).

Models estimation procedure must be supported by information about users' behavior through surveys. Surveys can be distinguished in two tipologies: those based on the users' choices in a real context, traditionally indicated as method of *Revealed Preferences* (RP), and those based on the statements made by interviewees about their preferences in different choice contexts (*Stated Preferences*, SP). Mixed use of RP and SP data allows to improve the estimation of parameters in the demand models obtained by SP data only. These estimation techniques, used since the early 1970s by some marketing researchers, are more and more frequently used for transport demand models calibration, particularly, of transport mode choice models (Cantarella *et al.*, 1995; Ortuzar, 1992; Postorino and Pirrello, 1993).
In the paper, these techniques are been applied in order to investigate about the existing relationships among tour type (*trip-tour* or *trip-chain*) and transport mode, and to verify if the need to make a trip-chain, a sequence of trips oriented to different activity, influence the transport mode choice. Multinomial Logit mode choice models are been calibrated to analyse these relationships; in their functional form has been inserted a variable equal to the number of trip in

the tour. The proposed models have been calibrated using mixed RP and SP data.

2 SP surveys methodologies and data analysis

SP tecniques are methodologies based on the statements made by interviewees about their preferences in different choice contexts, real, hypothetical or experimental. Therefore, an important innovation is introduced: the possibility to consider choice alternatives unavailable at the time of the surveys (Pearmain *et alii*, 1991). SP surveys methodologies involve the definition of choice alternatives, attributes (or factors) considered for each alternative, levels of variation of each attribute, choice contexts (scenarios) proposed to the decision-maker, type of preference asked, modality of interviews management (Cascetta, 2001). The number of possible scenarios depends on combinations of numerosity among the number of alternatives, the number of attributes and the number of levels of each attribute. The type of preference can be different: *choice*, *ranking* and *rating*. In the first case, the alternative chosen in that context must be indicated; in the second case, the available options must be ordered according to the degree of preference; in the third case, a preference measure must be assigned to each alternative, according to a predefined semantic scale.

The collected data by SP surveys can be used for demand models calibration relative to choice dimensions proposed to the decision-makers. Substantially, the estimation methods that can be adopted are those used for demand models calibration using RP surveys, but they differ for the type of preference used on SP experiment. The *choice* data can be analyzed as RP data; the only difference could consist in considering some hypothetical choice alternatives regarding those "revealed" in RP surveys. Also in *ranking* data the information can be analyzed in traditional way, considering alternatives two by two, and dealing the answers as if the "choice" was the alternative classified as better between two successive answers. Finally, in *rating* data, it is possible to assume as potentially chosen the alternative with higher value in a semantic scale, or to consider the "rate" associate to each alternative as indicative of utility.

In the case of *joint calibrations* using RP and SP data, usually it is applied the *scaling* estimation methodology; this methodology allows to consider the variability among different types of data used jointly in a statistic analysis (Bradley e Daly, 1992). In scientific literature, in MNL model calibration, various factor scale estimation methodologies are indicated. A sequential estimation method has been proposed by Ben-Akiva and Morikawa (1990); Bradley and Daly, in a 1991 paper, have proposed a simultaneous estimation method, based on construction of an artificial tree; an iterative estimation method has been proposed, two years later, by Postorino and Pirrello (1993).

3 Mode choice models proposed

In this paper, some MNL mode choice models are proposed, specified and calibrated using mixed RP and SP data. Data used for the calibration of models have been obtained by an experimental survey realized in the campus site of the University of Calabria, situated in the urban area of Cosenza (Italy). The campus is attended by 28,000 students and 2,000 members of staff approximately (March 2003). The surveys, realized in the spring of 2003, have interested a sample of 281 Engineering Faculty students, on a total of 6,600 students approximately, and have been planned to collect, on the same users' sample, RP and SP data (Guzzo and Mazzulla, *for press*). RP surveys have been done to collect some users' socio-economic characteristics and information regarding transport mode used to reach the university campus. In SP experiment, the users have expressed their degree of preference (in according to a semantic scale from 1 to 5) on 7 hypothetical choice scenarios; in each scenarios both *car* and *bus* alternatives are present. The *car* alternative is characterized by the *parking cost* attribute, that varies by actual level (free) to the intermediate and elevated level; the *bus* alternative is characterized by *frequency* (low and high) and *travel time* attributes (equal or reduced regarding to actual time). In total, the users have made 341 tours with at least one stop in the university area; 221 users have made one tour in a day only; 60 users have made two tours in a day. It has been analyzed the first tour make by users only. The access trip is realized primarily with an individual transport mode (74.2%), and only 25.8% of decision-makers uses the collective transport.

The calibrated models have been distinguished in RP models, based on the choices made by users exclusively in the real context, and in SP models, based on the choices stated in the hypothetical contexts, characterizing as "choice" the alternative to which the users associate, in *rating* experiment, a greater degree of preference. In order to improve the estimation parameters obtained by previous elaborations, conjoint RP/SP models have been calibrated. In the proposed models, two modal alternatives have been only defined (*car* and *bus*); each alternatives is characterized by usual attributes of level of service (times and costs of travel, bus frequency). The time attribute, in minutes, represents the total time on the access trip for the car and bus alternatives. The *cost* attribute, in Euros, represents, for the *car* alternative, the kilometric cost valued for the access trip added to the possible cost for the car-parking, and the total ticket cost for the *bus* alternative. The bus frequency has been expressed as a dichotomous variable of value equal to one for elevated frequency and zero otherwise. Additionally, some socio-economic variables have been considered, as a variable indicative of the user's gender (variable dichotomous equal to one for female gender and zero otherwise), and a variable indicative of the student's "out-site" condition, which assumes value equal to one when the student is resident in a place distant from the university campus and zero otherwise. Also, a variable indicative of the tendency to combine different activities in a same tour has been considered, that is equal to the number of trip in the tour; the number of trip is equal to two for the trip-tour, and is more than two for the trip-chain. The level of income, initially indicate by interviews in according to a scale with values from 1 to 5, has been subsequently calculated through a linear regression model, calibrated in function of further factors which influence the family income, as professional condition and employment sector of head of household and consort, the number of cars owned by the family and number of components licensed. In table 1 the model coefficients and test results of estimation procedure are brought. In the model, the variables PC-HeadFam show the professional condition of head of household and consort (unemployed=0, artisan=1, worker=2, employed=3, manager=4, businessman=5), while Car/Licensed indicate the ratio between the cars owned and number of components licensed.

Table 1 – Regression model for income estimation

Variable	*Coefficient*	*t-Student*	*Confidence intervals (95%)*		*Regression statistics*	
			Lower limit	*Upper limit*	*F-Fisher*	*R^2*
Const	1.454	10.627	1.185	1.723	76.333	0.453
PC-HeadFam1	0.168	6.824	0.119	0.216		
PC-HeadFam2	0.206	8.311	0.157	0.255		
Car/Licensed	1.275	7.515	0.941	1.610		

The results of estimations of mode choice models are reported in table 2. By results is deduced that all parameters have a correct sign and assume a value statistically different from zero, to a 95% level of significance. The parameters TTime, NTrip and Sex, in some models, have a lower level of significance. As attended, the *income* variable, in the *car* alternative, assumes a positive sign, in order to indicate that an high level of income influences on the choice to use an individual mode, even if more expensive than a collective mode. Besides, also the *off-site* and *frequency* variables, in the *bus* alternative, assume a positive sign. In the first case, it emerges that the student with parents resident in a place distant from the campus is constrained to use a collective mode, because he cannot use the cars owned by other members of family, except if he is owner of a car. In the second case, an increase of bus frequency induces the user to mostly use the collective transport mode. Finally, the sign assumed by NTrip variable indicate that the choice to make a trip-chain discourages the use of collective transport mode because, in this case, the decision-maker prefers to use a transport mode less “constrained”.

The weight of NTrip parameter, equal to 0.22 in the RP model (in absolute value), shows that contribution to the systematic utility of the *bus* mode is equal to 1.10 for a trip-chain with 5 trips and 0.44 for a trip-tour. This value influences notably on the probabilities of choice calculated by model. For example, a student resident in the university area, of male gender and with index of income equal to 3.23, can reach to the campus by *car* (access trip duration of 25 minutes and cost of € 2.00) or by *bus* (access trip duration of 45 minutes and cost of € 0.78). If he realizes a trip-tour, the probabilities of choice calculated by the model are equal to 0.36 for the *car* alternative and 0.64 for the *bus* alternative; otherwise, if he realizes a trip-chain with 5 trips, the probabilities of choice for the two modal alternatives are equal to 0.53 and 0.47 respectively. Therefore, *ceteris paribus*, in the case of trip-tour the *bus* alternative is chosen, while in the case of trip-chain the *car* alternative is chosen. In joint RP/SP model, Ntrip parameter is equal to 0.16, and therefore the tour type has a lesser effect on the probabilities of choice of modal alternatives.

Table 2 – Mode choice model calibration results

			RP Model		*SP Model*		*RP/SP Model*	
Alternative	*Variable*	*Parameters*	*Value estimated*	*t-Student*	*Value estimated*	*t-Student*	*Value estimated*	*t-Student*
CAR	TTime	β_1	-0.0220	-1.571	-0.0163	-3.522	-0.0230	-4.073
	TCost	β_2	-0.6955	-2.204	-1.237	-12.000	-1.5710	-12.730
	Income	β_3	0.4457	2.986	0.186	3.728	0.2746	4.587
BUS	TTime	β_1	-0.0220	-1.571	-0.0163	-3.522	-0.0230	-4.073
	TCost	β_2	-0.6955	-2.204	-1.237	-12.000	-1.5710	-12.730
	Sex	β_4	0.8385	2.111	-0.2880	-2.068	-0.1843	-1.072
	OffSite	β_5	2.0670	5.903	0.7060	6.232	1.1690	8.383
	NTrip	β_6	-0.2243	-1.503	-0.1050	-2.063	-0.1615	-2.603
	FreqBus	γ_1	-	-	1.2320	11.000	1.7280	12.690
V.O.T. (€/h)			1.90		0.79		0.88	
θ			-		-		0.73	
LogL(β)			-126.594		-975.265		-1115.76	
LogL($\mathbf{0}$)			-189.229		-1165.87		-1355.10	
LR			125.269 (χ^2 = 12.592)		381.217 (χ^2 = 14.067)		478.68 (χ^2 = 14.067)	
$\bar{\rho}^2$			0.299		0.157		0.172	
%RIGHT			78.55% (216/275)		69.27% (1188/1715)		69.84% (1385/1983)	

All models verify the statistical tests on the goodness-of-fit; the $\bar{\rho}^2$ statistics assumes the maximum value in the RP model (equal to 0.299). In the three proposed models, the estimation of the monetary value of the time (VOT) results equal to 1.90, 0.79 and 0.88 €/h respectively. The values obtained are apparently low; however, the interviewed sample is composed of students with the family which has a middle or middle-low income in the 85% of the cases. Additionally, the large disagreement between the VOT in RP model and in SP and RP/SP models can be explained considering that, in the last two models, the *cost* parameter, for the *car* alternative, includes both the kilometric cost valued for the access trip and cost for the car-parking; consequently, the additional cost for the car-parking entails than the user isn't disposed to pay a high fee in order to save an hour of time.

The scale factor, estimated by the procedure proposed by Postorino and Pirrello (1993), assumes a value equal to 0.73; the obtained value indicates an under-estimation of the calibrated parameters by the SP model and the necessity “to readjust” their value amplifying it of a factor equal to 1.37 (1/0.73).

4 Conclusions

In this paper, RP and SP joint analysis techniques have been adopted for the demand models estimation to predict the choices made by users and their preferences variations while the choice context changes. The study has confirmed the utility of SP techniques for analysis of users’ travel behavior in hypothetical choice contexts; it has confirmed, moreover, than RP and SP conjoint analysis improves the estimation of parameters in discrete choice models. The results show the choice to make a trip-chain, to execute more activities in the same tour, discourages the use of transport collective mode.

In the research development the proposed models transferability is inquired, so that models can be used for forecasting in different contexts.

References

M. Ben-Akiva and T. Morikawa, “Estimation of Travel Demand Models from Multiple Data Sources”, presented at 11th International Symposium on Transportation and Traffic Theory, Yokohama, Japan, 1990.

M. Bradley and A. Daly, “Uses of the Logit Scaling Approach in Stated Preference Analysis”, presented at 6th World Conference on Transportation Research, Lyon, 1992.

G.E. Cantarella, E. Cascetta and M.N. Postorino, “Le Catene di Spostamento e la Scelta Modale: Analisi tramite l’uso congiunto di dati RP ed SP”, in *Il Trasporto Pubblico nei Sistemi Urbani e Metropolitani*, S. Amoroso e A. Crotti (eds), FrancoAngeli, 175-198, Milano, 1995.

E. Cascetta, *Transportation Systems Engineering: Theory and Methods*, Kluwer Academic Publishers, Dordrecht, The Netherlands (2001).

R. Guzzo and G. Mazzulla, “Modal Choice Models Estimation Using Mixed Revealed and Stated Preferences Data”, in *Urban Transport X - Urban Transport and Environment in the 21st Century*, L.J. Sucharov and C.A. Brebbia (eds), WITPress, Southampton, UK, (*for press*).

J. de D. Ortuzar, “Stated Preference in Travel Demand Modelling”, presented at 6th World Conference on Transportation Research, Lyon, 1992.

D. Pearmain, J. Swanson, M. Bradley and E. Kroes, “Stated Preference Techniques: A Guide to Practice”, Second Edition, Steer Davies Gleave and Hague Consulting Group, 1991.

M.N. Postorino and P. Pirrello, “Calibration of a Modal Split Model Though combined RP and SP Surveys”, presented at Second Meeting of the Euro Working Group on Urban Traffic and Transportation, Paris, France, 1993.