Global & Local Economic Review

Volume 13 No. 2 2009

Tracce Edition

POLICY FOR THE SUBMISSION OF MANUSCRIPTS TO BE PUBLISHED IN THE GLOBAL & LOCAL ECONOMIC REVIEW

1. PRELIMINARY INSTRUCTIONS

For a manuscript to be accepted for publication in the Review, authors are required to abide by the following publication rules. In particular, authors must state that:

the manuscript, or relevant parts of the manuscripts, was not previously published;

the manuscript is not simultaneously under review by any other journal;

the manuscript will be not submitted elsewhere before the final outcome of peer-reviewing process by the Scientific Board.

2. MANUSCRIPTS SUBMISSION

Manuscripts must be sent to the Scientific Board by e-mail attachment to gler@fondazionepescarabruzzo.it. Papers, which must be submitted in printable formats, should include the manuscript text, footnotes, a bibliography, an abstract (both in Italian and in English) not longer than 300 words in length and the JEL code (available on http://www.aeaweb.org/journal/elclasjn.html).

Manuscripts written in English are also publishable, but it is the author's responsibility to ensure that papers are written in clear and comprehensible English, especially if they do not adhere to linguistic standards.

Institutional affiliation, telephone number and the e-mail address of the author (or the corresponding author in case of multiple authorship) should be placed on the first page of the manuscript.

All manuscripts should follow the format specified in the section "Instructions for authors", available on the Review website www.gler.it.

3. REVIEWING PROCESS

The manuscript is reviewed by the Scientific Board and, if judged suitable for publication in the Review, is then sent to three referees for a rigorous blind peer-reviewing process.

The initial decision by the Scientific Board takes 2 weeks.

Authors will be informed of the review decision and all relevant comments by the three referees within approximately a month.

Based on the referee's recommendations, the Scientific Board then decides whether the paper should be accepted as it is, revised or rejected. In any case, authors will be provided with a feedback.

In case of acceptance by the Scientific Board, the publication process will begin by assigning the manuscript to a number and a year of the Review.

Proof corrections and the printing process takes approximately a month.

4. MISCELLANEOUS

Authors submitting manuscripts to the Review are expected to be available for contributing as referees, in the event they would be so requested.

The articles published on Global & Local Economic Review will be listed in the ECONLIT and E-JEL bibliographies.

Global & Local Economic Review

Aut. Trib. PE n. 7 del 14.7.1999, No. 2/2009
Editor
Nicola Mattoscio

ISSN (*print*) 1722-4241 ISSN (*online*) 1974-5125

«Global & Local Economic Review» is included in JEL on CD, e-JEL and Econlit, the electronic indexing and abstracting service of the American Economic Association

Editorial Offices: Corso Umberto I, n. 83 - 65122 Pescara
Tel. +39 0854219109 – Fax +39 0854219380
Website: www.gler.it
E-mail: gler@fondazionepescarabruzzo.it

Global & Local Economic Review

Volume 13 No. 2 **2009**

SOMMARIO

TRACCE EDITION

Ennio Forte, Lucio Siviero, Lidia Carleo*

FROM FARE STANDARDIZATION TO FUNCTIONAL UNIFICATION OF URBAN TRANSPORT WITH REFERENCE TO THE LOGISTIC FUNCTIONALITY OF SERVICES

Abstract

This paper focuses on some problematic aspects of public transport fares, with particular reference to the Italian case. The transport economics approach is supported by a modelling application on theoretical basis concerning, at urban level, the potential reaction of demand to the supply of an improved functional integration between underground lines and surface road routes. In particular, through a logistic regression model, the study investigates the effects produced by the adoption of functional unification based on the reduction of waiting time for modal interchange. Hence structural trends toward fare and functional integration are investigated.

JEL classification: C13, C14

Keywords: Local public transport; integrated fares; functional unification; modal interchange

1. Introduction

In today's globalised world, local public transport (LPT) is no longer a historical 'ball and chain' soldered onto local government budgets. On the contrary, it is becoming more and more a business area with international supporting data and poles of excellence expressed in terms

* Department of Economics, University of Naples 'Federico II' - forte@unina.it. Even it is the result of study in common between the authors, this paper can be attributed to: paragraphs 1 and 7, Ennio Forte; paragraphs 2, 3, 4, 6, Lucio Siviero; paragraph 5, Lucio Siviero and Lidia Carleo.

of mergers, takeovers and incorporations of businesses so as to better exploit the advantages of the ‘near-monopoly’ positions which, by definition, characterise the LPT market. It is well known that in urban settings, with light and/or heavy railways and surface road routes, LPT can be summarised, regarding revenues, by the metaphor of ‘flesh’ and ‘bone’; road transport clearly representing the ‘bone’. The simple driver/passenger ratios of rail transport and road transport give an immediate idea of their different productivity levels with respect to travelling staff. For surface, road LPT there is also the difficulty of working in the setting of private traffic which often does not allow timeTab.s to be respected. Moreover, the network and the rolling stock are provided by the State, so LPT providers are asked to cover at least the operational costs of the business and, in the near future, to cover new network and rolling stock investment costs.

The most significant negative aspect is that the two systems (road and rail) are generally operated by different providers which do not foresee functional planning and unification of services (tariff standardization is only a first step). This results in a damaging autonomy of the single networks often without planning for connections and transfers and, therefore, without the strategic aim of the so-called ‘functional continuity’ and minimisation of waiting times for multimodal origin-destination (O-D) trips, a prerequisite of success according to users. Fares standardization favours the less well-off users to the detriment of marginal and inframarginal users (extramarginal users are the only ones who appear clearly satisfied, as is revealed during off-peak periods which, in many cases, have very low load factor which produce correspondingly high external diseconomies). In addition, adopting two different approaches with regard to the O-D’s home-work and work-home, different parameters of elasticity and, consequently, of willingness to pay, can be identified. Hence, operations planning should be done taking into consideration such economic parameters, better shaping the supply with respect to peak/non-peak hours of the work-home return period.

Therefore, it is necessary to take measures to avoid road LPT companies falling into bankruptcy, especially in a context of liberalisation. In the last ten years, in line with EU directives on the liberalization of the sector through a system of tenders aimed at accessing a more efficient market, highly contradictory legislative attempts have been made. More often than not, however, such solutions have all too often demonstrated the absence of clear and transparent regulations and of objective parameters for public contributions which, the higher they become, the more they represent a dissipation of resources. Some attempts aimed at efficiency have involved a number of bus lines in minor urban centres using systems that see greater user involvement (buses 'on request'), or services with control systems provided to specific groups of users.

Therefore the problem of the structural deficit of road transport remains. In large scale urban settings such transport is almost always provided by public operators rather than their private counterparts. Efficacy and efficiency are also measured by the extent of public contribution evaluated on a comparative basis (they should, however, also be calculated on the basis of demand that has been satisfied and not only of the supply that has been produced: i.e. according to the item passenger-kilometre, not number of seat-kilometre).

This paper deals with the topic of functional integration, as well as fares integration, of road-rail services with a view to adopting network models which foresee the rail systems as determining factors for the road services to be undertaken through a network architecture which has the road services playing a feeder role to the rail services, and therefore organised prevalently in relation to the latter.

2. Economic sustainability of public transport: a neglected target

In Italy one of the most widely used systems of tariff unification for urban public transport is generalised unification of multimodal systems, bound, generally, only by space limits (zones, areas, etc.) or time limits (on times of use, on duration of validity). Over the years, generic tariff unification has led to a quest for different systems of fare integration which, according to the analysis perspectives and to the objective assessment of utility and efficiency, have various positive and negative aspects for the two main players involved in the economic process of public transport, i.e. the providers and the users.

The third player involved in the economic process, the public institution, which should be *super partes* and play a key role not only as a distributor of public resources but as a regulator of a market which naturally tends towards a monopoly/oligopoly (in some rare cases towards monopolistic competition), occupies an anomalous position in relation to the supply/demand dynamics that, since they are public services, elude the rules of free market competition and therefore require intervention of a 'visible hand'.

The quest for greater efficiency on the part of businesses and the State, together with greater efficacy of services in favour of users are not, therefore, easily satisfied objectives, particularly in Italy where the economic performance is among the lowest in Europe. The sector is distinguished by the high incidence of labour costs as well as a significant dependence on resources coming from the public sector (the State, the Regions and local government structures).

Tab. 1 Balance of LPT Italian companies

Revenues	%	Costs	%
Traffic revenues	28	Raw materials	11
Public resources	62	Services	21
Other revenues	10	Labour	57
		Depreciation	8
		Other costs	1
Total revenues	100	Total costs	100

Source: ASSTRA elaboration from LPT company balance data 2006

Also unit costs comparisons show that costs have a higher trend than revenues. In 2002, net operating costs per km (before depreciation) were € 3.71, increasing to € 4.18 in 2006. Another interesting element is the ratio between traffic revenues and the number of passengers carried, an indicator which represents the average fare; while it is constant from 2005 to 2006, over the longer period from 2002 to 2006 it had a percentage increase slightly higher than half the percentage increase in operating costs per passenger.

Tab. 2 - Unit costs and revenues for urban public transport Italian companies (based on an ASSTRA sample) (Var. % 2002-2006; 2006-2005)

Indicators (€)	2002	2005	2006	Var. % 2006-2002	Var. % 2006-2005
Operating costs per vehicle km	3.71	4.05	4.18	12.7	3.2
Operating costs per passenger	0,96	1,05	1.07	11.5	1,9
Traffic revenues per vehicle	1.15	1.25	1.26	9.6	0.8
Traffic revenues per passenger	0,30	0,32	0.32	6.7	0,0
Total revenues per unit of labour	**64,928**	**70,000**	**74,005**	**14.0**	**5.7**

Source: ASSTRA

Fares integration is therefore believed to be a possible pathway for PT businesses to follow in order to increase traffic revenues and, at the same time, to allow greater generalised and non-specific use of the transport networks on the part of the users. It is clear that, in many cases, this has only been an apparent solution to the real problems of inefficiencies in public management, little concerned with covering operational costs with traffic revenues, and has provided users with integrated systems based on the potential use of number of kilometres offered, even when these are considerably more than necessary if correctly proportioned to the effective fluctuating demand during the different phases of the day (peak, off-peak). In Italy a single urban trip ticket has an average cost of € 1.0 (2008), which is greatly inferior to the € 1.7 average of other European countries (AASTRA 2007); this comparison suggests the need for more remunerative tariff levels and fare systems, better catering for the typologies offered by the operators. This could become the subject of more discriminatory compensation in favour of users, like, for example, the already proposed reduction of VAT on travel expenses, and forms of financial detraction of the expenses of season tickets for users of public transport, also if graduated according to income.

For most transport companies in Italy the price of the time ticket is precisely € 1; this confirms that the logic followed by administrations with the responsibility of setting fare levels has clearly been that of 'monetary simplicity' and not an economic assessment (ISFORT-ASSTRA, 2008). It is also to be noted that while the fare dynamic of the main sectors of public utilities are governed by a national Authority which regulates prices and tariffs (energy, gas, telecommunications), in Italy this instrument has not been adopted for LPT and therefore there is no independent national body able to guarantee the definition of the objective criteria for determining fares in such a way as to make them certain, transparent and, principally, economically in agreement with the consistency and typology of the services provided

3. Integrated fares system in urban public transport: some critical evidence

The spread of the integrated fares system in public transport, which occurred in Europe from the second part of the 1960's, was mainly dictated by the need to deal with factors regarding the development of mobility in urban centres, like: the difficulties for users caused by long waiting times generated in the absence of adequate forms of inter-modal, inter-vectorial and inter-company coordination; the objectives of environmental policies aimed at reducing the external costs of private transport and the opportunity of using the rail transport networks to the best in terms of economy of scale and of density.

By integrated fares system one may mean a contractual and payment method for public transport services which generally involves a number of public transport operators, with the scope of offering users the possibility of using a single travel document, the price of which does not depend on the means or the vectors used, nor on the number of transfers, but only on the quantity of transport purchased. Fares integration is only one of the essential prerequisites for making an integrated transport system based on the concept of 'functional continuity', that is:

a) *infrastructure integration*, which is carried out through the creation of interchange nodes, car parks, stations, bus terminals, systems of connections, etc.;

b) *mode integration*, that is, the possibility of using different transport modalities (road, rail) coordinated as much as possible, in such a way that the user has the perception of travelling on a single means of transport;

c) *tariff integration*, which consists in introducing a single fares system valid for all the modalities on offer. The criterion used for determination of fares must be unique, that is, independent of the specific means used by the user (Piacenza, Carpani, 2003).

In the literature there are various positions concerning the fact that, in order to be truly effective, fares integration should be 'exclusive' and therefore users should not have the possibility of buying travel documents for single trips or for specific couplings and combinations of vectors, businesses or means of transport.

Justification for this conjecture derives from the 'compensation effect' between users that use public transport very frequently and who enjoy a noTab. 'consumer surplus', given that the price paid is generally less than what they would have paid for using several non-integrated services, and users who consume minor quantities of transport (perhaps making a single journey on a single means), who find themselves paying higher relative prices under 'exclusive' integration. This effect is reflected in the form of traffic revenues for LPT businesses. The fare reduction guaranteed to high use users is 'compensated' by the lower use users and/or, this being specific to the transport network, the collectivity in general through the mechanism of state and regional financial contributes to LPT businesses, distributed according to supply. Indeed, 'pure' and 'exclusive' fares integration also foresees a generalised quantity of services produced, in many cases greater than the specific effective need, (for example for a single line, single zone of traffic or determined operating hours) which is distributed over the entire service network contributing in a not altogether rational way to increasing the vehicle-km produced and, as a consequence, to operating costs and costs of public contributions. As a matter of fact, routes with lower traffic and those with strong timeTab. related oscillations often have a very low utilisation coefficient (pass-km/seat-km), clearly demonstrating the production of 'empty' seat-kilometres.

From this there is a clear advantage for those who need to use public transport very frequently, as they pay the same fares while consuming more transport resources, while the effects on public balances and on the balances of businesses involved in LPT are not very clear given the far from simple identification of the economic-functional link between the production of services and the relative costs for businesses and the collectivity. The average operational costs in Italy are 3.6 €/km against a European average of 2.7 €/km, while, with reference to the revenues from public contributions, in Italy these run from € 13.8 per 1,000 seat-km supplied by A.T.M. in Milan, to € 45.0 per 1,000 seat-km supplied by the businesses operating in Naples (ANM, CPT and Metronapoli) (Boitani, 2008).

Supporting exclusivity in fares integration is the concern that otherwise there overall reductions of traffic revenues would be generated. However, some experiences in the application of 'non-exclusive' (with choice options) integrated tickets in Swiss cities have shown a non-decrease in the overall traffic revenues (Fitzroy e Smith 1994).

Besides, from a strictly economic viewpoint, 'second grade price discrimination' should be applied when dealing with services of public utility. This corresponds to a situation where a business sells different units of a service at different prices, but consumers buying equal quantities each pay the same price. In this case, though the business knows that different types of consumer exist, it is not able to or does not intend to stratify the demand through the mechanism of multiple pricing. The objective of non-stratification is trying to impede that 'strong' consumers may choose tariffs aimed at 'weak' consumers of public transport. Theoretically, the consumer should therefore be left the possibility of choosing to buy a single service or the integrated one, with fares structured in a way which creates an adequate system of incentives according to the real utility for each of the user types.

Finally, given the typical rigidity of the demand price of public transport, the offer of the integrated ticket in an exclusive and universal way for all user types could generate minor revenue for the businesses in cases where 'strong' consumers of public transport constitute the majority of users, who would be paying a fare which is inferior to the real utility or at least is inferior to the sum of the fares which intensive use of the public transport system would bring them to pay (Doxsey, 1984).

4. Fares and functional integration: approaching the road-rail modal interchange

With respect to fare integration alone, based on possible differing criteria of unification of the travel title, as for example the one prevalently applied in the main large Italian cities based on the maximum time of ticket validity, hence of use of the network, sometimes 'tailored' on a geographical basis with differentiation of 'zonal' or 'kilometric' fares, true integration could not be defined as such if it is not accompanied by redesigning of the multimodal transport network based on the principle characteristics which distinguish the functional integration of the services (Beria, Ponti, 2007):

- single mutually accessible services to favour modal interchange;
- working infrastructure in the interchange locations (private-public, rail-road, rail-rail, etc.);
- coordination of timeTab.s and frequencies;
- elimination of overlapping and redundancy;
- generation of a 'network effect' for the expansion of the demand and better use of the supply;
- reduction of 'empty' trips, especially in the off-peak hours and in zones with low demand;
- functional hierarchy, through strong lines (mainly rail) and feeder lines (hub and spoke at the urban level);
- correct dimensioning of the vehicles and the systems with respect to the real requirements of the demand.

A complete integration of multimodal public transport services, aimed at the possible higher level of operating efficiency and of quantity-quality efficacy in favour of the users, should not set aside either fares or functional integration with possibilities of choice for the users between different fare types according to specific needs and the expected and perceived total and marginal utility.

In urban settings with railways (underground, urban railways, trams), the architecture and the operational set up of networks is evolving more and more towards strongly hierarchical systems which see rail transport as representing the main backbone of the entire system of public transport, onto which a sub-network of road transport is 'hung'. Thus, in many cases underground stations and urban railway stations are to take on a role of hub with respect to the surface lines which are routed through them to destinations at a short-medium distance, generally at the level of the quarter and/or of the rather restricted area. Such a role is all the more reinforced for the more important or central railway stations which, in many cases, are concentrated the final destinations of road services which also serve more distant areas with a strong demand and/or not served by rail lines (Rome, Naples, etc.).

This function of interchange and distribution of flows from a node of the network to and from multiple origins/destinations can be found in hierarchically functionally integrated networks with strong lines and feeder lines operated in strict coordination and applying compensatory economic formulae between different means and companies of transport and no longer between 'strong' and 'weak' users of public transport, a most important aspect from the supply side and for public financing. All the more so, since from an economic point of view the rail systems register traffic revenues greatly superior to road networks (the Italian average of the ratio total revenues/total costs: undergrounds = 76,5%, road lines = 37,1%; CNIT, 2002).

The rail-road modal interchange thus becomes strongly strategic for several profiles due, on the supply side, to economic efficiency (lower costs and higher revenues through the increase of demand or compensation between vectors) while, on the demand side, to economic efficacy (qualitative improvements of the integrated system and lower waiting times). Defining and identifying an ideal functional integration fare for the metro-bus interchange for a single origin-destination trip (for example, one rail journey and one road journey) requires greater investigation regarding its potential effects on the demand side and on the supply side; these effects are analysed considering as strategic variables the average waiting time and the quality of the intermodal interchange which, within an integrated fares system for time and/or zone applied in an 'exclusive' way, does not tend to be specifically considered and assessed separately in utility functions and users' choice functions. Besides, with 'exclusive' fares integration 'by time' the concept of a 'temporal validity' of the ticket is to be adopted, not treated as total travel time but as time available before making the last interchange.

For these reasons, we have built a theoretical model for the economic analysis of random choice (logistic regression) for estimating the utility of undertaking a multimodal urban movement (rail-road) from an origin to a destination, in a single direction and within a single zone of application of a system of time standardised fares. The aim of the analysis is to estimate the parameters of elasticity of demand with respect to variations in costs to the user (price of the service) and of the waiting times at the modal interchange, comparing a time integrated fares system type of a non-functionally integrated network, through parameters of average supply measured on the overall journey times from origin to destination carried out with a single travel ticket, and a system with average performance and improved functional integration (certain overall journey time and planned waiting time). The objective of the model application is the simulation of alternative choices between two public transport systems with differing integration characteristics (fares only and functional and fares) and therefore estimation of the potential process of substitution of a service with an alternative. The criteria and theorems of the general theory of demand, and in particular those relative to the elasticity of substitution, can be applied to this simulated process, deriving from the random generation of different levels of generalised transport costs for the users (monetary cost plus time cost) (Del Viscovo, 1990).

Furthermore, specific simulations regarding fares variations, linked to variations in waiting times and journey times, have yielded estimates on users' willingness to pay on the basis of functional changes and, therefore, service performance (efficacy). A probabilistic theoretical segmentation of the demand has been made on the basis of different levels of service offered, read in terms of 'time' fares integration alone and, as an alternative, functional and fares integration 'of interchange', between urban railway lines and feeder urban road lines for short-medium distance travel. Based on the measurements of elasticity of stratified demand for price and for time, the possible effect in terms of variation of total traffic revenues has been estimated for the situations modelled, with stratified demand and price discrimination (generalised cost for the user) within the application of the integrated functional fares of rail-road services.

5. Theoretical applications of a logistic regression model

The proposed model was developed for estimating economic parameters which representing users' willingness to pay with respect to fares variations and the elasticity of demand with respect to prices (fares). The analysis was made using a binary logistic model of joint probability (executed in the Microsoft® Office Excel), in which the alternatives are represented by a base hypothesis of simple 'time' fares integration and a second hypothesis of functional integration having the special characteristics of calibration of the waiting times and journey times (min-max), as a function of the hypothesised integrated journey, and of 'interchange' fares.

The adopted model allows calculation of the indirect utility function of the user, in which the three principle assumptions on which the model is based are respected: first, that the overall utility of goods depends on the utility of the single characteristics of the goods itself; second, that within a set of choices the consumer prefers the goods with higher overall utility; third, that it is impossible for the researcher to include all the elements which contribute to the utility function of a single individual in the analysis, for which reason it is reasonable to assume that utility is a stochastic function (McFdden, Train, 2000). In this specific case, the choice of the user represents the dependent variable and is a function of the different observed alternatives which express the choice of the individual.

The utility functions of the single user compared to the alternatives *i* e *j* are:

$$U_{in} = V_{in} + \varepsilon_{in}$$
$$U_{jn} = V_{jn} + \varepsilon_{jn}$$

This functions are given:

- by a systematic component (V), constituted from the contribution arising from the different variables taken into consideration, which in the alternative 'time' fares integration are represented by the time and by the cost of the journey, while in the functional integration and 'interchange' fare hypothesis are represented by the travel time, the waiting time and the cost of the journey;

- by a random component (ε) constituted from the non-observable components. It can be note that compared to the same alternatives, individuals choose differently. Choices of subjects may differ due to different individual preferences that are not directly observable and measurable. Probabilistic theory considers the existence of various sources of uncertainty. The main ones are:
- attributes not included in the alternative, not known or not observable;
- differences in preferences;
- measurement errors and imperfect information regarding the attributes of alternatives and the characteristics of individuals;
- use of instrumental variables (proxy) that imperfectly describe the real variables.

In the applied model the distribution of the error term ε is represented by a logistic distribution. It is assumed that ε_A and ε_B represent the maximum number of random variables that capture attributes not observable and errors of measurement and specification. *Gumbel's theorem* shows that the maximum of several random variables IID (independent and identically distributed) follows roughly the distribution of Gumbel. If ε_A and ε_B then follow a Gumbel distribution and are therefore independent and identically distributed, then $\varepsilon = \varepsilon_A - \varepsilon_B$ is distributed as a logistic (Marcucci, 2005).

According to the binary logistic model the probability that the alternative i is chosen by the individual n in front of the whole choice composed by *i* and *j*, is:

$$P_n(i) = \frac{1}{1+e^{-(V_{in}-V_{jn})}} = \frac{e^{V_{in}}}{e^{V_{in}}+e^{V_{jn}}}$$

where *V* is the deterministic component of the utility function.

If systematic utilities are linear in the parameters, the equation becomes:

$$P_n(i) = \frac{1}{1 + e^{-(\beta' X_{in} - \beta' X_{jn})}} = \frac{e^{\beta' X_{in}}}{e^{\beta' X_{in}} + e^{\beta' X_{jn}}}$$

with β′ represents the transposed vector of parameters and *X* is the vector of attributes.

Application hypothesis I

The two alternatives *A* e *B* considered are:
- PT = public transport with ‘time’ fares integration;
- Metrobus = public transport with functional integration and ‘interchange’ fares.

Utility functions associated with the two alternatives considered are:

$U_A = \beta_1\, Time\ PT + \beta_2\, Cost\ PT + \varepsilon_A$
$U_B = \beta_1\, Time\ Metrobus + \beta_3\ \ Waiting\ time\ Metrobus + \beta_2\, Cost\ \ Metrobus + \varepsilon_B$

The dataset is composed of a sample of 500 theoretical observations obtained through the generation of constrained random numbers (Horowitz, 1999, 2006). The hypotheses formulated on the single variables (constraints) are:
– 10′< Time PT< 90′
– Cost PT = 1,1 €
– 10′< Time Metrobus < 40′
– 1′< Waiting time Metrobus < 15′
– Cost Metrobus = 1,1 €

The model was calibrated by the method of *maximum likelihood*. This method consists in maximizing the logarithm of the likelihood function (*LogL*) according to the probability of observing a given sample distribution conditional to the values of the parameters (β) being estimated.

In formulae:

$$LogL = \max_{\beta} \sum_{n} \log[y_{A,n} P_n(A) + y_{B,n} P_n(B)]$$

where:

$y_{A,n}$ = 1 if *n* chooses alternative A, 0 if *n* chooses alternative B

$y_{B,n}$ = 0 if *n* chooses alternative A, 1 if *n* chooses alternative B

P_n = probability of choosing *n*

The coefficients obtained were:
β_1 = - 0.445
β_2 = - 1.763
β_3 = - 0.329

From the analysis of these coefficients, the value of time (VOT) is 11.32 €/h. This value may be considered as a test estimating the coherence of the model; the value obtained is fully compatible with the average values of the times taken for urban mobility at a national level (HEATACO, 2006).

Validation of the model was carried out with both informal testing and formal statistical testing. The informal testing, that is relative to the reasonableness of the results, can be said to be satisfactory in that the signs of the coefficients, as expected, are all negative and their reciprocal ratios, as for example the VOT, give credible values.

Then formal testing of the goodness of fit of the model was carried out, as well as tests of the single coefficients. In particular, in order to verify the goodness of the model, the ρ^2 (*coefficient of determination*) was calculated giving a result of 0.918. This coefficient takes values between 0 and 1. A value near to unity demonstrates good explicative capacity of the model.

From Fig. 1 and 2 it is possible to see the probability distribution of user choice as a result of Metrobus fares and waiting times variations, keeping the other parameters of the utility function constant

Fig. 1 – Probability distribution on fares variation

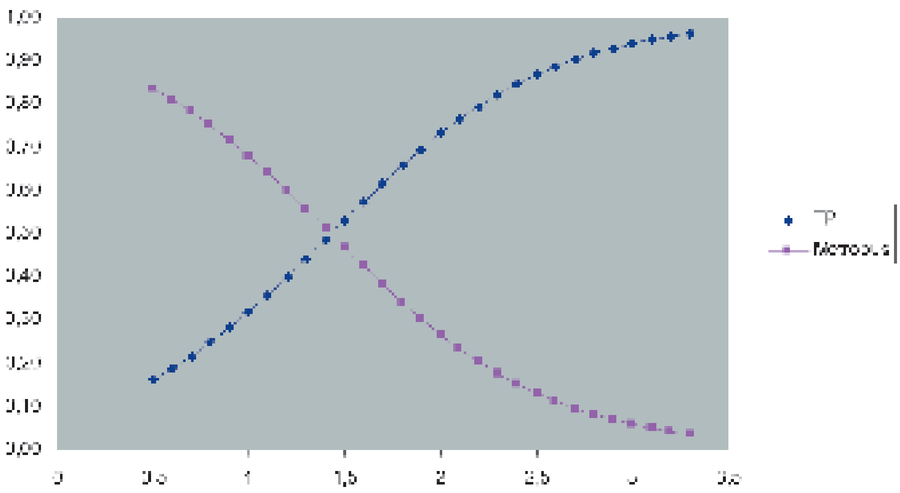
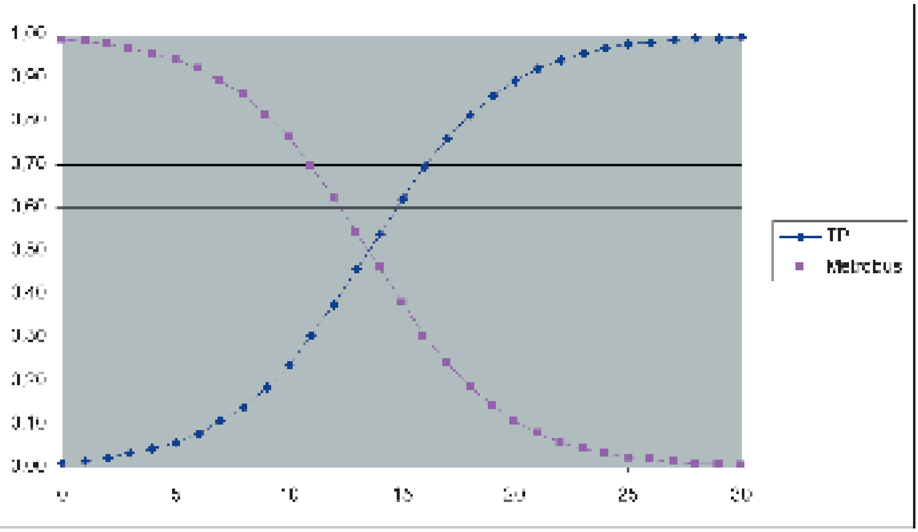


Fig. 2 – Probability distribution on waiting times variation



Application hypothesis II

A different operational hypothesis was also modelled and calibrated, this being a fares and functional integrated system which more closely reflects the mobility needs and choices of users who may be defined as 'inframarginal', i.e. those users who are not willing to support the marginal costs of the service beyond a given generalised cost threshold who, therefore, make alternative choices clearly closer to their own utility and willingness to pay. Such users, given the operational conditions of the system at a given time, prefer to abandon the network since they do not benefit from any surplus and therefore sustain a higher generalised cost which leads them to consider other forms or means of transport (car, motorbike, taxi, walking), evaluating the cost-time trade-off.

To this end the estimates of the coefficients generated by the model which determined a time value of 13.49 €/h were accepted as an indication of the stratification to be carried out with reference to a user group willing to pay more, especially considering possible savings of time. The system hypothesised by these considerations was defined 'Bus on call' (Bus C), which already exists in several Italian urban settings. It lies at an intermediate position between the hypotheses for public transport made previously. In particular, the bounding parameters for the generation of the random values on which the utility function depends, were modelled by matching typical performances of public transport on rail (underground, urban railways, etc.) with those of an feeder 'interquarter' distribution service gravitating around the stops of the urban rail lines (offered by mini-buses) which determine shorter waiting times (service 'on-call' comparable to taxi-sharing).

The dataset is composed of a sample of 500 theoretical observations obtained through the generation of constrained random numbers. The hypotheses formulated for the single variables (constraint) are:

– 10'< Time PT < 90'
– Cost PT =1.1 €
– 5'< Time Bus C < 30'
– 0'< Waiting time Bus C < 5'
– Cost Bus C = 2.5 €

The coefficients obtained by the maximum likelihood method are:
β_1 = - 0.987
β_2 = - 1.717
β_3 = - 0.386

From the analysis of these coefficients the value of time (VOT) is 13.49 €/h. Formal testing of the goodness of fit of the model and of the single variables were carried out. In order to verify the goodness of fit of the model the ρ^2 (*coefficient of determination*) was calculated, yielding a result of 0.942.

6. Economic considerations from the results of the model

The model has returned relevant statistical parameters for the economic interpretation of the variations occurring between the explanatory variables and the utility of the users in relation to the two different alternative systems of integrated fares. Besides, the elasticity of demand of the alternative called Metrobus was calculated with respect to the price (fare) and to the times both of the journey and of the wait for interchange. This has allowed evaluation of the willingness to pay for a service of the Metrobus type compared to the base situation identified in the hypothesis of 'time' fares integration.

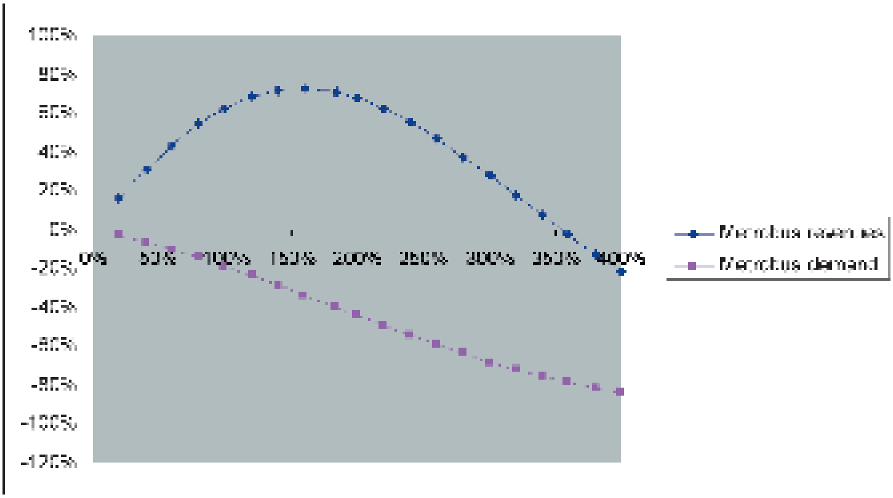
- *e* Metrobus (fare) = 0.735
- *e* Metrobus (waiting time) = 1.997

The elasticity with respect to the positive variation of price shows an inelastic trend in the interval of variation between the fare of the base hypothesis (€1.1) and the modelled fare in the hypothesis of 'indifference of choice' (equal utility) and probabilistic equidistribution of demand (€1.42). It may, therefore, be noted that users' willingness to pay for a better level of services is greater than the fare level of the base hypothesis ('time' fares integration).

The elasticity with respect to waiting times has an elastic trend in the interval of negative variation between the waiting time in the hypothesis of 'indifference of choice' of the demand (about 13,5 minutes) and the hypothesis of equal overall journey time (30 minutes), with the reduction of the waiting time to 10 minutes.

On the basis of the simulations carried out, with the scope of testing the elasticity of the variation of monetary costs of integrated Metrobus transport, a series of hypotheses were applied for the variation of the ticket price, and then the different probabilistic demand distribution was verified through the calibrated logit model. A revenues optimisation curve was calculated, hypothesising an increase in the Metrobus fare and leaving unaltered both the fare of the base hypothesis and the other variables taken into consideration. The results are reported in Fig. 3 from which one may see that the revenues would be optimised with an increase of 160% compared to a minimum fare of € 0.5 (at about € 1.3), with a decrease in traffic (segment of specific demand) of about 33%. The modest elasticity to the fare for the users, in particular for increases up to 20% compared to the price of the ticket in the base hypothesis (€1.3), indicates that the integrated Metrobus system would, in any case, be preferred by a proportion of demand such as to maximise revenues.

Fig. 3 – Revenues and demand as a function of fares variation



With reference to the second simulation (*application hypothesis II*), the model yielded interesting results in terms of the willingness to pay of inframarginal users who have greater values of time and therefore a different elasticity to the price and to the waiting time. For this reason, the potential proportion of demand interested in modifying its own choice on the basis of an improved offer was estimated, especially with reference to a minimisation of the waiting times for modal interchange, simulating 'on-call' service performances with availability in reduced times.

The results of the simulations were:

- *e* Bus C (fare) = 0.298
- *e* Bus C (waiting time) = 1.992

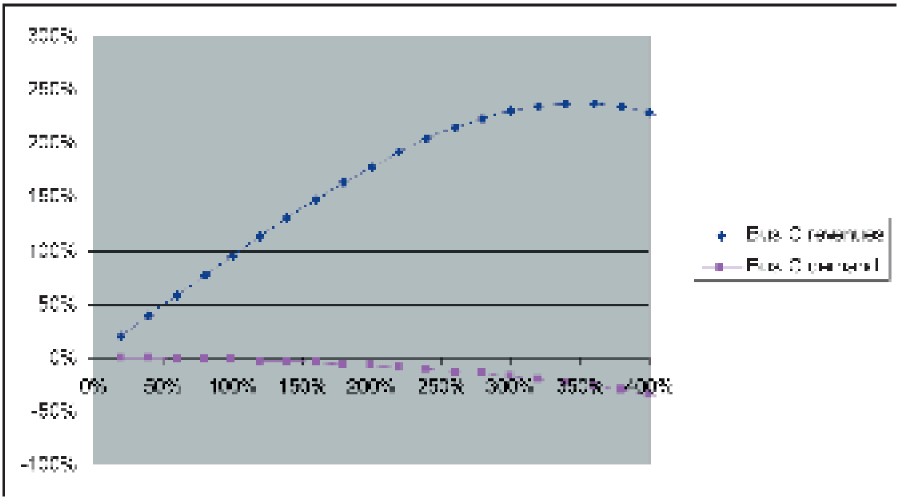
The elasticity with respect to the positive variation of price shows a rigid trend in the interval of variation between the fare of the base hypothesis (€ 1.1) and the modelled fare in the hypothesis of 'indifference of choice' (equal utility) and probabilistic equidistribution of demand (€ 2.85). One may note, therefore, that the willingness to pay for a higher level of service ('on-call' functional integration) is greatly superior at the fare level of the base hypothesis.

The elasticity with respect to the negative variation of waiting time has an elastic trend in the interval of variation between the waiting time in the hypothesis of 'indifference of choice' (19,4 minutes) and the hypothesis of equal overall journey time (30 minutes) with a reduction of waiting time to 10 minutes.

Such results are compatible with a stratification of the demand more interested in the time factor rather than the price factor and therefore willing to pay more for a service of greater quality, with reference to the waiting times, compared to the previous hypothesis.

The results of the latter hypothesis in economic terms are reported in Fig. 4, which shows that the revenues would be optimised with an increase of fare of about 340% with respect to the minimum fare of € 0.5 (yielding € 2.2), with a reduction of traffic (segment of specific demand) of about 23%. The low elasticity to the fare for the users, in particular for increases up to 100% with respect to the price of a base hypothesis ticket (€ 2.2) indicates that the Bus C integrated system would in any case be preferred by a proportion of demand such as to maximise revenues

Fig. 4 – Revenues and demand as a function of fares variation



7. Conclusion

The theoretical model of economic analysis of random choice (logistical regression) has been applied to estimate the utility to the users of LPT in carrying out multimodal rail-road urban transfers. The main objective of the economic-statistical analysis was to estimate the parameters of elasticity of demand to variations in costs for the user and of waiting times for modal interchange with the scope of its possible stratification and segmentation. The initial assumption was that a 'time' fares integration system of a non-functionally integrated network does not allow such segmentation and generates compensatory effects between not well identified user groups which in many cases are not assessed economically, both on the demand side and on the supply side. The modelling application enabled simulation of the alternative choices between systems of public transport with different characteristics of integration (only fares, and fares and functional) and therefore to make estimates of the potential process of substitution of a service with an alternative. Besides, specific simulations regarding the fares variation, connected to the variations in waiting times and journey times, gave estimates of the users' willingness to pay because of functional and performance changes of the level of service. Services with functional integration were estimated for the rail-road modal interchange in the hypothesis called 'Metrobus' and in the hypothesis defined as 'on-call bus' and compared to a generic 'time valid ticket' integration system. The analyses and the simulations carried out suggest some measures of intervention aimed at reinforcing the functional integration between 'strong' services on rail and complementary services on road, such as:

- planning, sizing and organising the services of a unitary and integrated network conceived as a 'fish bone', in the view of hub and spoke, where the main axis is represented by the urban rail lines and the single stations which are presumed to be optimal and barycentric for the area and/or quarter. The main network, for the rail part, must be conceived as a succession of hubs from which operate services of short/medium routes, possibly realised with low environmental impact vehicle of small/medium dimensions;
- stratification of the users and 'personalisation' of the fares in terms of the O-D length and the proximity of the destination;
- liberalisation and privatisation of road services so as to create an integrated system of 'shared taxis' aimed at satisfying the needs of the users with possible margins of flexibility of itinerary in a modern way;
- the introduction, in line with current legislation, of progressive forms of transition with the scope of reducing the weight of public funding on the balances of business deficits, possibly adopting criteria for the repartition of the contributions as measures of the efficiency of the system.

In conclusion, no longer being able to hide behind the theorem of the social utility of the service and therefore of the indispensability of external public financial support, one may imagine four scenarios of which the first two are more extreme and the other two are intermediate and thus more feasible:

- underground lines and road lines running parallel over long distances with transverse connections between the two systems which lead to additional deficit and economic, with low user satisfaction;
- multiplication of taxis as in the USA and the progressive reduction

of the road lines, which would lead to the elimination of public financing;
- private (or semi-private) 'shared taxis', such as mini-buses and with flexible itineraries at small area dimension;
- buses of medium/small dimensions with itineraries in the quarter or to/from 'interquarter thresholds' connected with the rail network.

References

Asstra (2007), *Le tariffe nel trasporto pubblico locale*, Asstra, Rome.

Beria P., Ponti M. (2007), *Introduzione ai sistemi di Trasporto*, Pitagora, Bologna.

Bhat C.R. (2001), *Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model*, Transportation Research, 35B, 677-693.

Boitani A. (2008), *Il LPT tra crisi e riforme*, in Sistemi di Logistica 4

Del Viscovo M. (1990), Economia dei Trasporti, UTET, Turin.

Doxsey L.B. (1984), *Demand for Unlimited Use Transit Passes*, Journal of Transport Economics and Policy, 18, 7-22.

HEATCO (2006), J.D. Shires and G.C. de Jong, *An international meta-analysis of values of time*.

Horowitz A.J. et al. (1999), *Guidebook for Statewide Travel Forecasting*, prepared for the Federal Highway Administration.

Horowitz A.J.(2006), *Statewide Travel Forecasting Models*, National Cooperative Highway Research Program Synthesi.

Fitzroy F., Smith I. (1994), *The Demand for Public Transport: Some Estimates from Zurich*, International Journal of Transport Economics, 21(2), 197-207.

Forte E. (2009), *Trasporti, Logistica, Economia*, Cedam, Padua.

Isfort-Asstra (2008), *5° Rapporto sulla mobilità urbana in Italia*, Asstra, Rome.

Mac Fadden D., Train K (2000), *Mixed MNL models of discrete response*, Journal of Applied Econometrics, 15, 447-470.

Marcucci E. (2005), (a cura di), *I modelli a scelta discreta per l'analisi dei trasporti*, Carocci Editore, Rome.

Piacenza M., Carpani C. (2003), *Esperienze di integrazione tariffaria nel*

trasporto pubblico locale in Italia, HERMES.
RT-Isfort (2006), *Quaderno n. 9 'Mobilità urbana. Rassegna delle esperienze italiane ed europee'*, Isfort, Rome.
Train K. (2003), *Discrete choice methods with simulation, Cambridge University Press*, Cambridge.

G. & L. E. R., Vol. 13 No. 2, 2009

Pasquale Sarnacchiano[1], Antonello D'Ambra[2]

CUMULATIVE CORRESPONDENCE ANALYSIS AS A TOOL FOR OPTIMIZING FACTOR SETTING IN PUBLIC TRANSPORT

Abstract

In public transport an improvement of the customer satisfaction can attract further users and individual transport would be used less. Most studies of customer satisfaction deal with ordered categorical data. The aim of this paper is to propose Cumulative correspondence analysis as a tool to develop a procedure for choosing the optimal categorical response in a multifactor state system. We briefly review the cumulative correspondence analysis and the Taguchi's statistic and by a case study, recording the choice of the best scenario for a train service, we illustrate the proposed approach.

JEL classification: C13, C14

Keywords: Cumulative correspondence analysis; taguchi's statistic; design of experiments; customer satisfaction; orthogonal array

1. Introduction

Over the last few years, companies have gradually focused on service quality and customer satisfaction. This strategy is very profitable for both companies and customers, particularly for transit agencies and passengers.

[1] University of Naples Federico II sarnacch@unina.it
[2] Department of Accounting Management and Quantitative Methods - Second University of Naples antonello.dambra@unina2.it

A sure guide for business strategies is the knowledge of customers evaluation of the service. This knowledge represents a competitive advantage. In this way, the company may guide the renewal of its offer as well as the pricing itself. Moreover the enterprises of excellence tend necessarily to the maximum customers satisfaction. There have been several technical proposals for the quality assessment of the services, and others are being added, together with the theory improvement and brand new practical challenges.

An improvement of the supplied service quality can attract further users and individual transport would be used less. Most studies of customer satisfaction deal with ordered categorical data. Accounting for the ordering of such data plays an important role in determining effectively the optimal factor level of combination. The aim of this paper is to propose an optimal scenario based on a study of *a-priori* customer satisfaction regarding several scenarios in public transport. Cumulative correspondence analysis is used to develop a procedure for choosing the optimal categorical response in a multifactor state system based on Taguchi's statistic.

A main aspect of this paper is visualization. The proposed procedure facilitates graphical exploration of factors so that the really important factors and optimal levels are easily determined in a speedy manner.

The paper is organized as follows. In section 2 we expose the case study and introduce the Design of Experiments based on Taguchi approach in order to determinate nine scenarios for the train service Naples-Rome. In section 3 we present the Cumulative correspondence analysis using Taguchi's statistic. In the last section we show that the previous results could be used to analyze CS and to determinate the optimal settings for train service Naples-Rome.

2. Case study and design of experiment

The proposed research regarded the train service Naples-Rome, the aim was to choose the best scenario for this service. To the purpose we used the Taghuci Methods (TM) for Design Of Experiment (DOE).

DOE can be highly effective when you wish to: a) Optimize product and process designs, study the effects of multiple factors (i.e. variables, parameters, etc.) on the performance, and solve production problems by objectively laying out the investigative experiments; b) Study influence of individual factors on the performance and determine which factor has more influence, which ones have less. You can also find out which factor should have tighter tolerance and which tolerance should be relaxed. The information from the experiment will tell you how to allocate quality assurance resources based on the objective data. It will indicate how to combine different factors in their proper settings to get the best results.

The DOE using TM can economically satisfy the needs of problem solving and product/process design optimization projects. By learning and applying this technique, engineers, scientists and researchers can significantly reduce the time required for experimental investigations.

TM is a technique to lay out experimental (investigation, studies, survey, tests, etc) plan in most logical, economical, and statistical way. You can potentially benefit from it when you want to determine the most desirable design of your product/service, best parameters combination for your process, most robust recipe for your formulation, permanent solution for some of your production problems, most critical validation/durability test condition, most effective survey/data collection plan, etc.

The TM, also called orthogonal arrays, allows you to analyze many factors with few runs. TM are balanced, that is, no factor is weighted more or less in an experiment, thus allowing factors to be analyzed independently of each other. Engineering knowledge should guide the selection of factors, levels and responses.

In our case, we selected the main quality characteristics (control factors), then we decided the number of factor levels and we considered four control factors of three levels (Tab.1). The total possible runs engendered were 81.

Tab. 1. Control factors and their levels

	Levels 1	2	3
A. Journey times	€ 19,50	€ 27,60	€ 36,[illegible]0
B. Comfort	Each four hours	Each two hours	Each hour
C. Cost	70cm	75cm	80cm
D. Frequency	2:09	1:54	1:29

In order to analyse this experiment using TM, the most desirable and suitable. design, was an L-9 Orthogonal Array (Tab. 2).

Tab. 2. L_9 orthogonal array and factor assignment

	Column (factors)			
Exp. No	**A**	**B**	**C**	**D**
1	1	1	1	1
2	1	3	2	2
3	1	3	3	3
4	2	1	2	3
5	2	2	3	1
6	2	3	1	2
7	3	1	3	2
8	3	2	1	3
9	3	3	2	1

Then we have submitted the experimental trials to a sample of potential Customers (n=58), particularly we required an evaluation for each scenarios. The main results are summarized in Tab. 3.

Tab. 3. Train service Naples Rome - Customer Satisfaction

	Scenario I	II	III	IV	V	VI	VII	VIII	IX
Mean	2,51	3,81	2,67	3,31	1,86	4,22	3,66	4,62	3,83
Median	2.00	4,00	3.00	3,00	2.00	4.00	3.00	5.00	4,00
S.D.	3,56	5,66	5,71	7,24	4,66	15,78	0,91	31,83	9,16

3. Cumulative correspondence analysis and Taguchi's statistic

Consider a two-way contingencytable, N, that cross-classifies n individuals/units according to I row categories and K ordered (ascending) column categories. Denote n_{ik} as the (i; k)-th joint cell frequency with a relative frequency of $p_{ik} = n_{ik}/n$ which is the (i; k)-th element of the matrix P.

Let $n_{i\bullet}$, $n_{\bullet k}$ and D_1 be the i-th row and k-th column marginal frequencies and a diagonal matrix where the (i; i)-th element is $p_{i\bullet} = n_{i\bullet}/n$, respectively. Denote $Z_{ik} = \sum_{j=1}^{k} n_{ij}$ the cumulative frequency of the i-th row category up to the k-th column category and their consideration provides a way of ensuring that the ordinal structure of the column categories is preserved. Similarly, denote $d_k = \sum_{j=1}^{k} n_{\bullet j}/n = \sum_{j=1}^{k} p_{\bullet j}$ the cumulative relative frequency up to the k-th column category.

Taguchi (1966) proposed as a measure of the association between the row and ordered column variables the following statistic

$$T = \sum_{k=1}^{K-1} w_k \left[\sum_{i=1}^{I} n_{i\bullet} \left(\frac{Z_{ik}}{n_{i\bullet}} - d_k \right)^2 \right] \qquad 0 \le T \le \left[n(I-1) \right] \tag{1}$$

Nair (1987) considered the case where $w_k = \left[d_k (1-d_k) \right]^{-1}$ and showed that the distribution of T for $N \rightarrow \infty$ is proportional to a weighted sum of independent chi-squared random variables with one d.f. (Satterthwaite's, 1946). Nair (1987) also demonstrated the link between the Pearson chi-squared statistic and the Taguchi's statistic

$$T = \sum_{k=1}^{K-1} \chi_k^2 \tag{2}$$

Here χ_k^2 is the Pearson chi-squared for $I \times 2$ contingencytable obtained by aggregating column categories 1 to k and aggregating the column categories (k+1) to K. For this reason T is also referred to as "cumulative chi-squared statistic". The inertia for the Taguchi's statistic will be T/n .

Generally, correspondence analysis can be performed by partitioning the Pearson chi-squared statistic into the sum of squares of the singular values of the Pearson ratio's, or residuals (depending on the scaling used for the singular vectors).

If, we call:

$$Y = W^{1/2} A P' D_I^{-1/2} \tag{3}$$

Where

$$A = \begin{bmatrix} 1-d_1 & d_1 & \cdots & d_1 \\ 1-d_2 & 1-d_2 & \cdots & d_2 \\ \vdots & \vdots & \cdots & \vdots \\ 1-d_{k-1} & 1-d_{k-1} & \cdots & d_{k-1} \end{bmatrix} \tag{4}$$

then the total inertia can be expressed as:

$$\frac{T}{n} = tr\left[Y'Y\right] \tag{5}$$

Realizing the Singular Value Decomposition (SVD) of the matrix Y, we have

$$Y = \tilde{A} D_\lambda \tilde{B}' \tag{6}$$

Where

$\tilde{A}$ right eigen-vectors of dimension $\left[(K-1)\times K\right]$, $\tilde{B}$ left eigen-vectors of dimension $\left[I \times M\right]$. D_λ diagonal eigen-values matrix of dimension $\left[M \times M\right]$, $M = \min\left[I,(K-1)\right]$

Since $\tilde{A}'\tilde{A} = I$ and $\tilde{B}'\tilde{B} = I$ it's possible to verify that λ_m eigen-values can be expressed by

$$\lambda_m = \sum_{i=1}^{I}\sum_{k=1}^{K-1}\sqrt{\frac{w_k}{p_{i\bullet}}}\,\tilde{a}_{km}\,\tilde{Z}_k\,\tilde{b}_{im} \tag{7}$$

Where $\tilde{Z}_{ik} = {Z_{ik}}/{n}$ and $\tilde{a}_{kn}$ e $\tilde{b}_{im}$ are the elements of the eigen-vectors of the matrix $\tilde{A}$ and $\tilde{B}$ respectively.

In order to visualize the association between the ordered column categories and the nominal row categories the set columns of $\tilde{B}$ and $\tilde{A}$ may be considered. That is, the coordinate of the row category in M-dimensional space may be defined by $\left\{\tilde{b}_{im}{:}m{=}1,...,M\right\}$ while the variation between the *k*-th pair adjacent, and ordered, column categories can be visualized by considering the coordinate $\left\{\tilde{a}_{km}{:}m{=}1,...,M\right\}$ for $k = 1,...,K-1$. These coordinates are analogous to standard coordinates used in the classical correspondence analysis (Grenacre, 1984) and do not reflect the nature of the association between the two variables, Instead, one may consider the following row, column coordinates and supplementary column coordinates, respectively:

$$F = D_I^{-1/2}\tilde{B}D_\lambda \tag{8}$$

$$G = W^{-1/2}\tilde{A}D_\lambda \tag{9}$$

$$G^{-} = I'\tilde{B}D_\lambda \tag{10}$$

If the user wishes to identify how the ordered column categories impact upon the association between the two categorical variables it is possible to use Cumulative correspondence analysis of ordered categories (D'Ambra et al., 2008) in which the Taguchi's statistic was decomposed.

4. Results

Starting from the results of customers satisfaction obtained from the nine scenarios and presented in the paragraph 1, it is evident that the favourite scenario was the eighth followed by the sixth. Nevertheless, it is not possible to evaluate the contribution of each factor level to the customer satisfaction. Moreover it could be very interesting for the enterprise to have an overall customer satisfaction index for all the possible scenarios. In this way, considering also the economical and technical aspects, the enterprise could be setting the four control factors maximising customer satisfaction and considering the economical and technical constraints.

In order to determinate the optimal setting for train service the obtained results has been analysed with the cumulative correspondence analysis.

The data table for Cumulative Correspondence Analysis has to be ranged as follows: in the rows the different levels for each quality characteristics and in the columns the ordered modalities of the evaluation variable (Tab. 4).

Tab. 4. Data Matrix for Cumulative correspondence analysis of ordered categories

	Factor	I	II	III	IV	V
A1	Low cost	13	26	29	32	71
A2	Middle cost	8	52	65	53	16
A3	High cost	45	28	26	30	5
B1	Low frequency	42	67	41	13	11
B2	Middle frequency	1	42	47	47	27
B3	High frequency	13	27	42	38	54
C1	Low comfort	19	53	62	34	6
C2	Middle comfort	7	43	25	20	51
C3	High comfort	16	46	35	44	35
D1	Low Journey times	9	45	42	20	54
D2	Middle Journey times	26	31	42	44	31
D3	High Journey times	[illegible]	26	46	34	7

Starting from Tab. 4, it is possible to construct four contingency tables (dimensions 12 x 2) obtained by the unification of the ordered modalities of the evaluation variable (1 vs 2-3-4-5, 1-2 vs 3-4-5 and so on). All the chi-squared statistics computed have resulted significant (Tab. 5). The sum of the four chi-squared are equivalent to T-Taguchi's index (Taguchi, 1966; Hirotsu, 1986).

It is well known that classical Pearson's chi-squared test for contingency tables does not have good power against ordered alternatives (Nair, 1987).

Tab 5. Chi-squared and T-Taguchi's index for contingency Tabs

Index	Value	DoF	P-value	Percentage
$\chi^2_{1vs2345}$	93,66	11	0,000	15,06
$\chi^2_{12vs345}$	135,49	11	0,000	21,78
$\chi^2_{123vs45}$	171,51	11	0,000	27,58
$\chi^2_{1234vs5}$	221,31	11	0,000	35,58
T	621,97			

The Cumulative correspondence analysis of ordered categories computed on Tab. 4 showed that the first two eigenvalues explained more the 97% of the total inertia (Tab. 6). The sum of eigenvalues is equivalent to T-Taguchi's index divided by the total of the frequency of the Tab. 4 (D'Ambra et al., 2008).

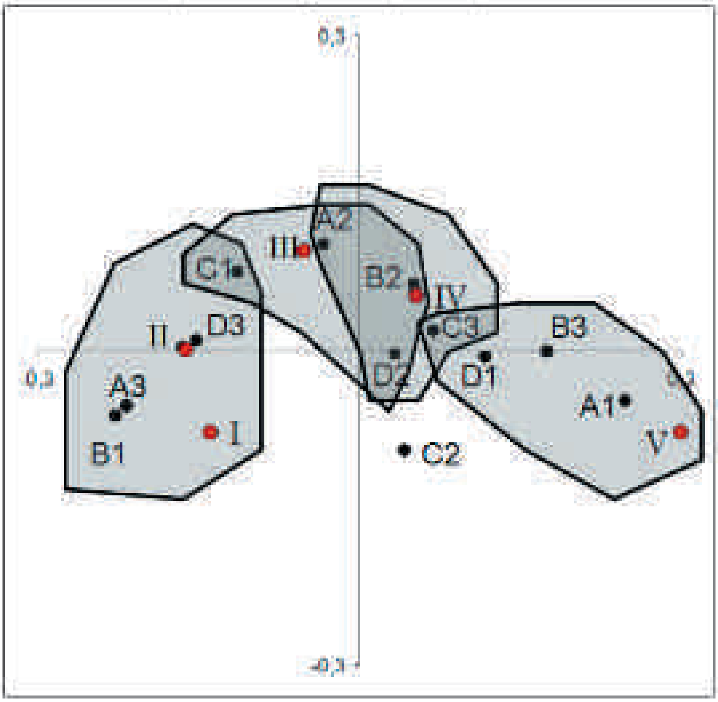
Tab 6. Cumulative correspondence analysis eigenvalue and decomposition of Taguchi's index

Eigenvalues	Percentage	Cumulative percentage
0,2505	84,11	84,11
0,0386	12,96	97,07
0,0080	2,70	99,77
0,0007	0,23	100

In order to visualize the association between the ordered column categories and the nominal row categories, we computed the row and column coordinates according to the formula (8) and (9).

Moreover we projected, as supplementary points, according to formula (10), the ordered modalities of the evaluation variable in the other space (Fig. 1). This graphical representation allowed to define a characteristic scenario for each level of customer satisfaction. This information matched with the enterprise economic planning allowed decision makers to have a very useful tool to set the transport service that maximise the customer satisfaction. In our case, as expected the best scenario was represented by the combination A1-B3-C3-D1. Nevertheless the informative information is that, if the enterprise for business opportunities can't offer this scenario, thanks to this approach the decision maker may define the level and the factor to modify that minimize the loss of Customer Satisfaction.

Fig. 1. Cumulative correspondence analysis: Graphical representation on the first two axis



For each of the five level of the overall assessment was constructed a scenario of train service provided (Fig. 2).

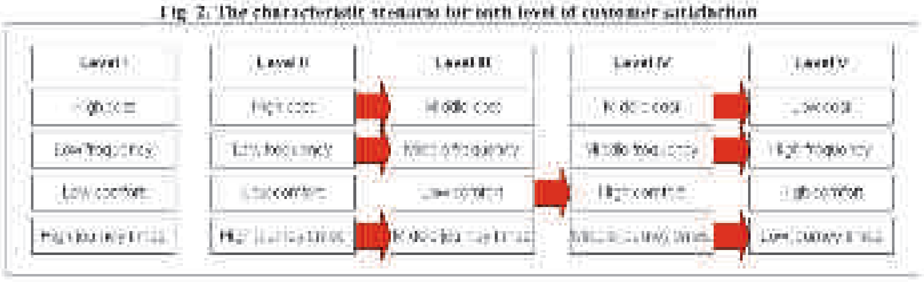
Analysing this Fig. we observed that the "characteristic scenario" corresponding to the lower level of customer satisfaction is not different from the scenario of the next level of Customer Satisfaction.

Differently, if the enterprise needs to pass from a level II to a level III of customer satisfaction, the enterprise has to improve simultaneously the factors "cost", "frequency" and "journey times".

In the same way to pass from level III to level IV, the client asks to improve significantly the factor "comfort".

The last scenario matching to the maximum level of Customer satisfaction (V) is not very interesting because it corresponds to the best level for each factor.

Fig. 2: The characteristic scenario for each level of customer satisfaction



References

Asiabar M.H., Ghomi S.M.T.F. (2006), *Analysis of ordered categorical data using expected loss minimization*, *Quality Engineering*, 18, 117 121.

Barlow R.E., Bartholomew, D. J., Bremner, J.M., Brunk H.D. (1972). *Statistical inference under order restrictions*, Wiley, New York.

Beh E.J. (1997), *Simple correspondence analysis of ordinal cross-classi¯cations using orthogonal polynomials*, Biometrical Journal, 39, 589-613.

Beh E.J., D'Ambra L., Simonetti B. (2007). *Cumulative correspondence analysis for ordered categorial data using Taguchi's Statistic*, in review.

Box G., Jones S. (1986), *Discussion of `Testing in industrial experiments with ordered categorical data, Technometrics*, 28(4), 295-301.

Chipman H.A., Hamada M.S. (1996), *Bayesian analysis of ordered categorical data from industrial experiments, Technometrics*, 35(1), 1-10.

D'Ambra L., Koksoy O., Simonetti B. (2008), *Cumulative correspondence analysis of ordered categorical data from industrial experiments*, Journal of applied statistics.

Greenacre M.J. (1984), *Theory and application of correspondence analysis*, Academic Press London.

Hamada M.,Wu C.F.J. (1990), *A critical look at accumulation analysis and related methods*, Technometrics, 32(2), 119-162.

Hirotsu C. (1986), *Cumulative chi-squared as a tool of goodness of fit*, Biometrika, 73, 165-173.

Jeng Y.C., Guo S.M. (1996), *Quality improvement for RC06 chip resistor*, Quality and Reliability Engineering International, 12(6), 439-445.

Koch G.G., Tangen C., Tudor G., Stokes M. (1990), *Strategies and issues for the analysis of ordered categorical data from multifactor studies in industry*, Technometrics, 32(2), 137-149.

Nair V.N. (1986), *Testing in industrial experiments with ordered categorical data*, Technometrics, 28(4), 283-291.

Nair V.N. (1987), *Chi-squared type tests for ordered alternatives in contingency Tables*, Journal of the American Statistical Association, 82, 283-291.

Parsa A.R., Smith B.S. (1993), *Scoring under ordered constraints in contingency tables*, Communications in Statistics - Theory and Methods, 22, 3537-3551.

Phadke S.M. (1989), *Quality engineering using robust design*, Prentice-Hall, New Jersey.

Ritov Y., Gilula Z. (1993), *Analysis of contingency Tables by correspondence models subject to ordered constraints*, Journal of the American Statistical Association, 88, 1380-1387.

Satterthwaite F.E. (1946), *An approximate distribution of estimates of variance components*, Biometrics Bulletin, 2, 110-114.

Schriever B.F. (1983), *Scaling of order dependent categorical variables with correspondence analysis*, International Statistical Review, 51, 225-238.

Taguchi G. (1966), *Statistical Analysis* (in Japanese), Maruzen, Tokyo.

Taguchi G. (1974), *A new statistical analysis for clinical data, the accumulating analysis, in contrast with the chi-square test*, Saishin Igaku, 29, 806-813.

Taguchi G. (1991a), *Taguchi methods: Case studies from the U.S. and Europe*, American Supplier Institute, Michigan.

Taguchi G. (1991b), *Taguchi methods: Research and development*, American Supplier Institute, Michigan.

Taguchi G. (1991c), *Taguchi methods: Signal-to-noise ratio for quality evaluation*, American Supplier Institute, Michigan.

Wang C.H., Tong L.I. (2003), *Quality improvement for dynamic ordered categorical response using grey relational analysis*, Advanced Manufacturing Technology, 21, 377-383.

Wu F., Yeh C. (2006), *A comparative study on optimization methods for experiments with ordered categorical data*, Computers & Industrial Engineering, 50, 220-232.

G. & L. E. R., Vol. 13 No. 2, 2009

Tonio Di Battista[1], Tonino Sclocco

SAMPLING DESIGN FOR LOCAL PUBLIC TRANSPORT ASSESSMENT

Abstract

The diffusion of the so-called Assessment Transportation has led local public transport companies to apply assessment procedures of Passengers Satisfaction. These procedures require statistical sampling surveys of the customers of the public transport from which to extract information useful for the management. A crucial step in defining the sampling design is the choice of how to select the units which, in the context of transport, are clustered and sparsely distributed over the study area. In these circumstances, standard sampling techniques cannot be applied and the authors propose the use of a specific sampling design called Intersect Units Sampling often used for monitoring biological populations. This technique selects the units along a transport service-line following a specific sampling design. The design proposed is easy to apply and leads to efficient estimators.

JEL classification: C13; C51

Keywords: local public transport, sampling design, Horvitz Thompson estimator, model-based sampling, design-based sampling

1. Introduction

In recent years, the importance of an appropriate local public transportation planning in the context of urban and territorial development policy has grown significantly. This has led to a more "realistic" approach, with the objective of using an integrated and coordinated spatial

[1] University of Chieti-Pescara - dibattis@unich.it

planning in order to increase the quality service in terms of passenger satisfaction. Local management has recognized the importance of any new decision (commercial, residential, etc..) and its effects on the existing transportation system.

A proper Transportation Assessment (*TA*) will help Local Public Administration (*LPA*) in assessing project's compatibility with the planning policy framework (usually the Local Development Framework) and with the relevant transport strategy (usually the Local Transport Plan). Furthermore, a proper *TA* will help *LTA* in evaluating all the transport implications and in identifying suitable measures for a more sustainable mobility. The *TA* can also address issues of interest to the local traffic authority related to their network management duties.

Current best practice recommends that a Public Transportation Assessment (*PTA*) be submitted as part of any planning application wherein the proposed development is likely to raise a significant amount of traffic.

In the PTA framework, the Local Transport Management (*LTM*) is interested in monitoring and assessing passenger satisfaction, the efficiency and the effectiveness of the service, travel times, waiting times etc.

For each of those aspects, that we call real phenomenon, we can identify a set of measures or variables of interest. Each variable can be considered to have a specific probability distribution in a based model point of view while in a fixed population framework it may have a set of fixed values.

For example, concerning the high-quality service supply, we can observe the obsolescence of means of transport or their CO_2 emissions. In order to evaluate the passenger satisfaction we can observe a suitable sequence of items measured in a Likert scale in order to recover the well known five latent dimensions such as empathy, tangible aspects, reassurance, reliability and responsiveness. With the purpose of planning the service, *LTM* is also interested in recording variables such as the network mobility, the morphology of the territory, etc. The assessment of the local public transport by means of statistical methods requires three essential aspects:

1. synthesis measures, such as the mean, the median etc. in order to evaluate the modality that better represents the entire real phenomenon.

2. variability measures, i.e. the dispersion of the values with respect to the synthesis measures. The purpose is to know the diversity of the measures of the real phenomenon.
3. Shape measures with the purpose of evaluating skewness and kurtosis of the data distribution.

In order to apply the statistical methods we can distinguish two approaches: the complete surveys and sampling surveys.

In the complete surveys setting, the knowledge of the entire population under study is required . But in practical applications, it is not possible to survey the entire target population. In the sampling framework, one has to decide which model to adopt: the model based approach or the design based approach. In the former, it is necessary to know the sampling distribution of the estimator while in the latter, it is necessary to know the list of units of the population. In this paper we focus on the design based approach. In particular, we propose an innovative sampling strategy suitable when the list of units of the population is not available. The proposal seems to be very useful for local transport assessment.

2. Sampling surveys

In order to obtain information on the *PTA* we need to choose a sampling survey strategy. As it is well known the sampling surveys can be conducted under a design-based approach or a model- based approach (Thompson, 1993). In the design–based approach, the values of the variable of interest are viewed as fixed quantities and the selection probabilities depend on the design adopted (Hedayat, Sinha, 1991). On the other hand, in the model-based approach, the values of the variables of interest in the population are viewed as random variables and the characteristics of the estimators depend on the joint distribution of these random variables which do not depend on the design choice.

The model-based approach has received more attention when it seems natural to the researchers to postulate a theoretical model which describes the real phenomenon under study. In this setting the variables $Y_1, Y_2, ..., Y_N$ associated to the population units can be assumed to be independent and identically distributed. In this case many standard statistical results apply without reference to how the sample is selected. However, it is difficult to encounter surveys in which a model of independent and identically distributed *y*-values can be assumed. In fact, the main problem of the model-based approach is that for many real populations it becomes mathematically complex while still not being suitably realistic. Moreover, in the *PTA* framework, results of many survey programs will be used by people with widely different viewpoints and often conflicting preferences regarding whether an estimate should be higher or lower. Thus, it would be hard to propose a model that would be acceptable to all the people involved in the PTA. In such a situation, the elimination of natural human selection biases through some sort of random selection procedure can be a powerful pragmatic argument in favour of an approach that is design-based.

Formally, let $\mathbf{y} = (y_1, y_2, ..., y_N)$ be the vector of *y*-values associated at each *N* units of the population. From the model viewpoint, these *y*-values are random variables with a joint distribution *F*. In the basic sampling theory for finite population, the population consists of a known, finite number *N* of units. The *y*-value of each unit in the population is a quantity, and even if it is unknown it can never be considered as a random variable. The units in the population are identifiable and labelled for example with numbers from 1 to *N*.

A sample **s** of $n<N$ unit is selected and observed. The procedure by which the sample is selected is called the sampling design. The design is determined by assigning to each possible sample the probability $P(\mathbf{s})$ of selecting that sample. The easiest design is the well known simple random sampling (*SRS*): a step-by-step procedure which consists of selecting a unit label at random from 1 to *N* and then selecting the next unit label at random from the remaining numbers between 1 to *N* and so on until *n* distinct units are selected.

In this setting, the inference problem is to estimate some summary characteristic, or parameter of the population, as a suitable function of $\mathbf{y} = (y_1, y_2,..., y_N)$. Formally, we indicate the parameters with $\theta = f(\mathbf{y})$ and with $\hat{\theta} = h(y_1, y_2,..., y_N)$ its estimate. For example, in the *PTA* viewpoint, $\mathbf{y} = (y_1, y_2,..., y_N)$ could be the passenger satisfaction, $\theta = f(\mathbf{y})$ the passengers satisfaction index in the population while $\hat{\theta} = h(y_1, y_2, \ldots, y_N)$ represents the passenger satisfaction estimate after observing the sample **s**. At the same time, it is possible to asses the accuracy of the estimate by means of the confidence interval. For this purpose, the variance of the estimator $\hat{\theta} = h(y_1, y_2, \ldots, y_n)$ must be introduced *i.e.* var $(\hat{\theta}) = E(\hat{\theta} - \theta)^2$.

Recalling $P(\mathbf{s})$, the probability of selecting the sample **s**, *i.e.* **s** is a sequence or a subset of the units in the population, an estimator $\hat{\theta} = h(y_1, y_2, \ldots, y_n)$ is said to be design-unbiased for $\theta = f(\mathbf{y})$ if its conditional expectation, given the realization of the N population y-values, is the realized value of $\theta = f(\mathbf{y})$:

$$E(\hat{\theta} \mid \mathrm{y}) = \theta$$

It is important to note that the design-unbiased estimator $\hat{\theta} = h(y_1, y_2, \ldots, y_n)$ derives its unbiasedness only from the design adopted (Sen, 1988). On the other hand, an estimator $\hat{\theta} = h(y_1, y_2, \ldots, y_N)$ is model-unbiased if, given any sample **s**, the conditional expectation $E(\hat{\theta} \mid \mathrm{s}) = E(\theta)$. In other words, the estimator is unbiased under the population distribution F for θ given the sample **s**.

3. Intersect units sampling

In Public Transportation Assessment (*PTA*), the passenger is frequently spatially clustered distributed. The population units are rare and then it becomes difficult to detect citizens that make use of public transport. In this setting, the drawback is to define a list of units. Thus, we have the need to implement a suitable sampling design.

Since we do not have a list of the population to be sampled, the sampling designs defined in the previous paragraph cannot be used. In these circumstances, in order to solve this drawback we propose the use of a particular sampling technique called Intersect units sampling (Di Battista, Valentini, 2007). This sampling design is frequently adopted in order to monitor biological populations such as animals or plants (Di Battista, 2003). In the particular case of *PTA*, the sampling units can be considered as a biological population that results to be unknown, rare and spatially clustered distributed. With the intersect units sampling, the units sampled are those encountered along a path over an area of study selected by a specific design.

Given the complexity of the design, the terms *path* and *intersect* will be specified when used.

As it is well know, the design-based sampling approach requires, as a first step, the computation of the first order inclusion probability denoted by $\pi_i > 0$ for $i = 1,2,\ldots,N$. In the intersect units sampling this probability corresponds, for each unit, with the probability to meet the units along the path. The second order inclusion probabilities, $\pi_{ij} > 0$ for $\forall\, i, j$, correspond for each pair of units with the probability to meet both units along the path.

Thus, if the sampling design ensures that every unit has got a first order inclusion probability different to zero, then it is possible to use the well known Horvitz-Thompson total estimator (Horvitz, Thompson, 1952). It results to be unbiased with variance depending on the second order inclusion probability $\pi_{ij} > 0$.

However, with the intersect units sampling, the unbiased estimator of the variance has numerous problems (Wolter, 1985). Indeed, it results to be difficult to develop designs in which the second inclusion probability is positive for each pair of units. Since the units are far from each other, it is difficult to meet both units on the same path. Furthermore, in real survey, the second order inclusion probabilities are not easy to compute, especially in local public transport.

To solve these drawbacks, we introduce a suitable technique named *replicated sampling designs* (Barabesi, Fattorini, 1998).

Indeed, if the sampling design chosen is replicable under the same condition *r* times, then the afore mentioned drawbacks, as we will see below, can be easily solved.

We point out that the replicated sampling seems to be appropriate with intersect units sampling designs. In fact, a path refers to a small part of the entire area under study, therefore it is not sufficient to ensure a representative sample of the entire area, so we prefer to choose at random and independently multiple paths.

4. Replicated sampling

Let θ denote a population parameter of interest that we want to estimate by means of intersect units sampling. Let $\hat{\theta} = t(S)$ be an estimator of θ obtained as a function of the sample data S where S is obtained by means of the design chosen. Suppose also that $\hat{\theta} = t(S)$ is an unbiased estimator of θ with variance $V(\hat{\theta})$. Therefore, if the design is r time replicable under the same conditions and the replications are mutually independent then we obtain r independent samples $S_1, S_2, \ldots, S_r$ that give r estimates of θ, *i.e.* $\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_r$. Obviously they represent the determination of r *i.i.d.* random variables with mean θ and variance $V(\hat{\theta})$. Quoting the well known properties of the mean of r *i.i.d* random variables, we have that the mean of the r estimates

$$\bar{\hat{\theta}}_r = \frac{1}{r}\sum_{i=1}^{r}\hat{\theta}_i$$

is still an unbiased estimator of θ with variance given by

$$V\left(\bar{\hat{\theta}}_r\right) = \frac{V\left(\hat{\theta}\right)}{r}.$$

Roughly speaking $\bar{\hat{\theta}}_r = \frac{1}{r}\sum_{i=1}^{r}\hat{\theta}_i$ results to be an estimator of θ r time more appropriate with respect to the estimator $\hat{\theta} = t(S)$ obtained using only a path. Moreover for $r \rightarrow \infty$, $\hat{\theta}_r$ results to be also a consistent estimator for θ.

Regarding the variance estimation, we can derive an unbiased estimator for $V(\bar{\hat{\theta}})$ by means of

$$\hat{V}\left(\bar{\hat{\theta}}_r\right) = \frac{s_r^2}{r}$$

where

$$s_r^2 = \frac{1}{r-1}\sum_{i=1}^{r}\left(\hat{\theta}_i - \bar{\hat{\theta}}_r\right)^2$$

which is the sampling variance of the estimators $\hat{\theta}_i$ for i=1,2,…,r.

Finally, the central limit theorem ensures (Berger, 1998) that for $r \rightarrow \infty$

$$\Pr\left\{\frac{\bar{\hat{\theta}}_r - \theta_r}{V\left(\bar{\hat{\theta}}_r\right)}\right\} \le u \to \Phi(u) \tag{1}$$

where $\Phi(u)$ is the normal standardized cumulative distribution function.

A straightforward use of (1) allows us to asses tests of hypothesis and confidence intervals for the parameter θ as follows

$$\Pr\left\{\bar{\hat{\theta}}_r - u_{1-\alpha}\sqrt{V\left(\bar{\hat{\theta}}_r\right)} \le \theta \le \bar{\hat{\theta}}_r + u_{1-\alpha}\sqrt{V\left(\bar{\hat{\theta}}_r\right)}\right\} \to 1-\alpha$$

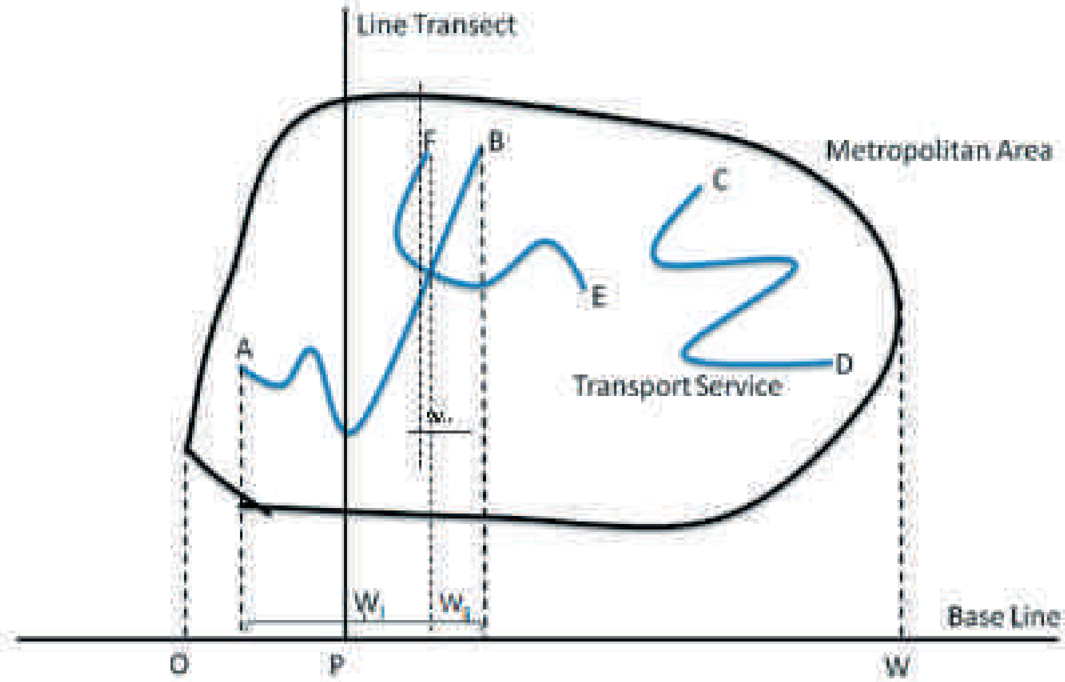
and for r sufficiently high $(r \to \infty)$ we have

$$\bar{\hat{\theta}}_r \pm u_{1-\alpha}\sqrt{\hat{V}\left(\bar{\hat{\theta}}_r\right)} = \bar{\hat{\theta}}_r \pm u_{1-\alpha}\frac{s_r}{\sqrt{r}}.$$

5. The use of line intercept sampling for Local Public transport

Suppose we want to assess passenger satisfaction of the network public transport. In this setting, the objective is to select the passengers in order to submit the questionnaire. The population is obviously elusive so that it is impossible to define the list and to get its total. Therefore, we suggest to adopt the line intercept sampling.

Fig.1 Line transaction design for local public transport



First of all, we need to define the study area such as the metropolitan area. Along the entire study area we plot a segment, said *base-line,* of length W. Line intercept sampling consists in selecting, at random, a point *P* on the *base-line* and then in sampling every transport service encountered along the path perpendicular to the base-line from the point *P*. This path is also named line-transect. Using this sampling design, a transport service will be selected when the random point *P* is in the projection of the service transport line on the base–line (Burnham, Anderson, 1976).

Let w_j be the length of the *j-th* projection on the *base-line*, then the first order inclusion probability for the *j-th* transport service is

$$\pi_j = \frac{w_i}{w}, \quad j = 1,2,\ldots,N$$

from which we obtain a straightforward Horvitz –Thompson estimator for the total and the mean.

The second order inclusion probability is given by

$$\pi_{ij} = \frac{w_j}{w}, \quad i \neq j = 1,2,\ldots,N$$

In real applications this kind of probability results to be difficult to compute because transport services are usually located at a distance for which there is not intersection on the base-line; so we suggest to use the approach based on replicated sampling described above.

In this setting the procedure consists in replicating *r* times the same design, that is to allocate randomly *r* points along the base-line.

As seen before, a straightforward design based inference can be easily (Kaiser, 1983) obtained by using the asymptotic distribution given by the expression (1).

References

Barabesi L., Fattorini L. (1998), *The use of replicated plot, line and point sampling for estimating species abundance and ecological diversity*, Environmental and Ecological Statistics, 5, 353-370.

Berger Y.G. (1998), *Rate of convergence to normal distribution for the Horvitz-Thompson estimator*, Journal of Statistical Planning and Inference, 74, 149-168.

Burnham K.P., Anderson D.R. (1976), *Mathematical models for non parametric inferences from line transect data*, Biometrics,32, 325-336.

Di Battista T. (2003), *Resampling Methods for Estimating Dispersion Index in Random and Adaptive designs*, Environmental and Ecological Statistics, 10, 83-93.

Di Battista T., Valentini P. (2007), *Complex sampling designs for the CSI estimation*, Statistica, LXVII, 3, 293-308.

Hedayat A.S., Sinha B.K. (1991), *Design and inference in finite population sampling*, Wiley, New York.

Horvitz D.G., Thompson D.J. (1952), *A generalization of sampling without replacement from a finite universe*, Journal of the American Statistical Association, 47, 663-685.

Kaiser L. (1983), *Unbiased estimation in line-intercept sampling*, Biometrics, 39,965-976

Sen P.K. (1988), Asymptotic in finite population sampling, in Handbook of Statistics, 6: Sampling, 291-332. Krishnaiah and C.R. Rao eds, North Holland, Amsterdam.

Thompson S.K. (1993), *Sampling*, Wiley, New York,.

Wolter K.M. (1985), *Introduction to variance estimation*, Springer-Verlag, New York.

G. & L. E. R., Vol. 13 No. 2, 2009

Michele Gallo[1], Enrico Ciavolino

MULTIVARIATE STATISTICAL APPROACHES FOR CUSTOMER SATISFACTION INTO THE TRANSPORTATION SECTOR

Abstract:

The aim of this paper is to study the effect of the dimensions of the transportation service on the Passenger Satisfaction (PS) taking into account the spatial effect due to the interaction across spatial units and spatial heterogeneity. The relationships between the service dimensions and PS are formalized by a Structural Equation Model (SEM) based on the Partial Least Squares (PLS) estimation method which includes the spatial effects in the measurement model. Moreover, in order to get a 'true' measure of satisfaction, the rating scale model is proposed.

JEL classification: C13; C51

Keywords: Latent trait; Rating Scale Model; Spatial Structural Equation Models; Partial Least Squares

1. Introduction

The customer satisfaction survey gives important sources of information for quality assurance in many economic sectors, where the customer satisfaction is a vital concern for companies and organizations in their efforts to improve service quality, and maintenance of customer loyalty. However, customer satisfaction cannot be measured by simple statistical tools. It is a result of a latent complex information process summarized in a multiple-items questionnaire, in which one set of alternative responses is used for estimating probabilities of responses. For this reason, in the analysis of multi-item data the multidimensional

[1] "L'Orientale" University of Naples - mgallo@unior.it

nature of customer satisfaction and the different nature of the data should be considered (Gallo, 2007). In the public transport sector, the measurement of customer satisfaction (hereafter referred as passenger satisfaction) might be influenced not only by the particular features of the transport sector, but also by some spatial effects attributable to the territorial dislocation of stations.

Consider the spatial effect into passenger satisfaction model present two reasons. First it is expected that the spatial effect of socio-demographic, economic or regional activity may be an important aspect of a modeling problem. Second one the observations associated with spatial units might reflect measurement error (Le Sage, 1999).

Following Papalia *et al.* (2008) and Ciavolino *et al.* (2009) a new strategy which accounts for the spatial effect into PS analysis is proposed, including, as in a unique process of analysis, the Rating Scale Model (Gallo, 2009).

The paper framework is based on three steps. First, a particular version of Rasch Analysis called rating scale model (Andrich, 1978) is used to get a 'true' measure of satisfaction. Second differential item functioning and the spatial dimension is used to build the structural equation model. Latter partial least squares algorithm is proposed to estimate which component has more influence on the passenger satisfaction.

2. Theory

The Rating Scale model and the PLS are presented in the next sections, in order to show the potentialities and the main characteristics of both of them in the framework of the Passenger Satisfaction.

2.1 Rating Scale Model (RASCH ANALYSIS)

More latent trait models could be used to measure passenger satisfaction, but Rasch models are distinguished from others by a fundamental statistical characteristic - subject sum score is a 'sufficient statistic' for the underlying unidimensional latent trait (Wright and Linacre, 1989). The model is based on the simple idea that passengers who have a high total score on an item are more satisfied overall than passengers with low scores. Likewise, items that receive lower ratings are more difficult to endorse than items that receive higher ratings. This way, on a single continuum of interest, it is possible to clearly identify which items are more difficult to generate satisfaction and which passengers are more satisfied than others.

When all items present the same set of alternatives, it seems reasonable to expect that the relative difficulties of the steps between categories should not vary from item to item. For these kinds of questionnaires the rating scale is the more appropriate version of Rasch models. Rating Scale Model – within a probabilistic framework – converts ordinal raw-score data, such as the scale strong dissatisfaction / dissatisfaction / satisfaction / strong satisfaction, into an interval-based measure, the log-odd metric or logit. Let $P_{ij(s)}$ be passenger *i*'s probability of scoring *s* on item *j*, the rating scale model can be written:

$$P_{ij(s)} = \frac{\exp\left(x_i - \phi_j - \gamma_s\right)}{\left[1 + \exp\left(x_i - \phi_j - \gamma_s\right)\right]} \tag{1}$$

where ϕ_j is the difficulty for item j to generate satisfaction, x_i is the attitude of the i-th passenger to be satisfied, and γ_s is the threshold parameter associated with the transition between response categories *s*-1 to *s*.

To estimate these parameters, the "Joint Maximum Likelihood Estimation" algorithm is used in this paper (Wright et al., 1969). This method is more flexible and it is independent from specific passenger and item distributional forms. Moreover the logits measure $\ln\left(P_{ij(s)} \big/ \left(1 - P_{ij(s)}\right)\right)$ of the items, passengers and rating scale categories, convert ordinal raw scores into linear interval measures.

One important part of the item analysis is to examine Differential Item Functioning (DIF) in the items. DIF refers to differences in item functioning after groups have been matched with respect to ability or attribute that the item purportedly measures. Where DIF shows a difference between the groups of passengers, it does not mean that there exists measurement bias since it might be a real difference in satisfaction level. In these cases, DIF measure can be used instead of the passenger one to study the different level of satisfaction between the groups.

2.2 Spatial Structural Equation Model (S-SEM)

The Partial Least Squares (PLS) estimation method was first formalized by Herman Wold (1966, 1973), for the use in multivariate analysis. The application in Structural Equation Modelling (SEM) was again developed by Wold (1975) and the main references on the PLS algorithm are Wold (1982, 1985). The main idea of PLS for the SEMs is an iterative combination of path analysis to give a measure of the relationships among the theoretical constructs (*Structural Model* or *Inner Model*, equation 2), then factorial analysis for measuring the latent construct (*Measurement Model* or *Outer Model*, equation 3):

$$\xi_{(m,1)} = B_{(m,m)} \bullet \xi_{(m,1)} + \hat{o}_{\ (m,1)} \tag{2}$$

$$x_{(q,1)} = \hat{\imath}_{(q,m)} \bullet \xi_{(m,1)} + \delta_{\ (q,1)} \tag{3}$$

In the *Structural Model*, equation (2), $\boldsymbol{\xi}$ is the vector of the m latent variables and $\mathbf{B}$ is the path coefficients matrix, with zeros on its diagonal representing the causal effect among the latent variables. The *Measurement Model*, equation (3), contains the $\mathbf{x}$ vector of the q manifest variables and the coefficient matrices $\boldsymbol{\Lambda}$ of the relationships between the latent constructs and the observed variables. The vectors $\boldsymbol{\tau}$ and $\boldsymbol{\delta}$ are the structural and the measurement error vectors and the $\boldsymbol{\Psi}$ and $\boldsymbol{\Theta}^{\boldsymbol{\delta}}$ are respectively the diagonal matrix variance of the structural error term $\boldsymbol{\tau}$ and the measurement error term $\boldsymbol{\delta}$.

Since we are observing H different units located in different positions, we have to take into account the effects which the geographic position can generate into the model. Spatial structures are generally associated to: a) *Absolute location* effects which are relevant to evaluate - for each observation - the impact of being located at a particular point in space, and to b) *Relative location* effects that consider relevant the position of an observation relative to other observations. The first effect called *spatial heterogeneity* assumes that each observation can have its own characteristic for the phenomenon under investigation. Moreover, in the latter case, it is assumed that the value observed in a sample in a specific location h can be affected by the value observed in another location k, with $h \neq k$. This effect, called *spatial dependence*, is due to the spatial interaction between contiguous observations.

The SEM formulation is therefore extended to take into account the spatial heterogeneity and the spatial dependence. The *spatial unobserved heterogeneity among* spatial observations is allowed by introducing fixed effects in the measurement model (Bernardini Papalia 2006, 2008a,b). We proceed by including an individual specific "dummy variable" to capture unobserved heterogeneity for each spatial observation h (h=1,…, H).

For the *spatial dependence*, we focus on one of the widely used approaches (called *spatial LAG model*) where the spatial correlation pertains to the dependent variable. In this context, it is assumed interdependence of latent variables across areas. This assumption may be formalized by including a spatial lag variable into the measurement model which represent the relationship between the manifest and latent variables. In doing this, a spatial weights matrix **W** of non-stochastic time constant weights has to be specified. This is a ($H \times H$) matrix in which the rows and columns correspond to the cross-sectional locations.

An element w_{hk} of the matrix expresses the prior strength of the interaction between location h (in the row of the matrix) and location k (columns). This can be interpreted as the presence and strength of a link between nodes in a network representation that matches the spatial weights structure. In most applications, the choice is driven by geographic criteria, such as contiguity (sharing a common border) or distance, including nearest neighbor distance (Anselin 1988; Lesage and Pace 2004).

More specifically, using the equation (3), the set of latent exogenous variables $\boldsymbol{\xi}$ is enlarged to include: *(i) Spatial Lag variable* (4), that is the first-order contiguity spatially lagged dependent variable; the *fixed effect*, that is the location dummy reported in equations (5); (iii) the set of q exogenous manifest variables $X_{H,q}$.

$$\mathbf{w}^{x} = Spatial - Lag = \mathbf{W}_{H,H} \cdot \mathbf{x}_{H,1} \tag{4}$$

$$\mathbf{d}^{s} = Dummy - Space = \mathbf{I}_{H,H} \tag{5}$$

To take into account the spatial lag variable and the fixed effect, the manifest variables of the equation (3) are rewritten as follow:

$$\mathbf{X}^{*}_{H,q+1+H} = \left[\mathbf{X}_{H,q} \mid \mathbf{W}_{H,H} \cdot \mathbf{x}_{H,1} \mid \mathbf{I}_{H,H} \right] \tag{6}$$

The equation (6) reports the specification of the locations in the measurement model, where, the vector $\mathbf{x}_{q,1}$ of the q manifest exogenous variables, is expressed in the form of matrix $\mathbf{X}_{H,q}$, with the locations reported in the rows and the variables in the columns. The matrix $\mathbf{I}_{H,H}$ is the identity matrix for the H locations.

The associated $\mathbf{\Lambda}$ matrix, which specifies the regression coefficients of the observed variables on the latent variables, is defined as $\mathbf{\Lambda}= [\boldsymbol{\tau} \mid \rho \mid \alpha]$, including: the set of the *manifest variables coefficients* ($\boldsymbol{\tau}$), the *spatial autoregressive parameter* (ρ) and the coefficient of the *spatial effects*.

The matrix formulation of the exogenous measurement model, equation (4), can be reformulated considering the spatial and the fixed effects, as below reported:

$$x_{q+1+H,1} = \hat{\imath}_{q+1+H,n} \bullet \xi_{n,1} + \delta_{q+1+H,1} \tag{7}$$

The measurement model is extended in this way adding to the q manifest exogenous variables, the spatial lag variable and the spatial effect, that means 1+H rows. In the estimation of a S-SEM, it is then essential to deal with the problem of endogeneity of the spatial lag term originated by the correlation between latent endogenous and exogenous variables and, as a consequence, the correlation between exogenous observed variables and errors. Our proposal is to use the PLS, which can be a powerful estimation method of analysis in case of small sample size, strong correlation among the items, missing data and no residual distribution assumption.

3. The measure of Passenger Satisfaction

To measure the passenger satisfaction a survey analysis was conducted on 2,473 passengers. The questionnaires were submitted by 10 different interviewers in the second week of October according to stratified random samples. Nine items ('station cleanness', 'train cleanness', 'passenger comfort', 'regularity of service', 'frequency of service', 'staff behavior', 'passenger information', 'safety', 'personal and financial security') are used where each item has a Likert scale with four ordinal levels (Likert scale), viz., 'strong dissatisfaction' / 'dissatisfaction' / 'satisfaction' / 'strong satisfaction'.

The analysis of the Passenger Satisfaction consists of two parts: in the *first part*, the Rasch analysis is used to focus on the psychometric properties of the items, passengers, and rating scale categories. When the Rash diagnostic results guarantee the passenger satisfaction measure in terms of validity and reliability, the DIF measures were used to obtain the items measure for each station. The WINSTEPS program (Linacre and Wright, 2000) was used to obtain the results from these data.

To measure the relationships among the several aspects of passenger satisfaction, in the *second part*, a SEM is defined by considering three latent variables, '*Transportation*', '*Information & Security*' and '*Comfort & Cleaness*', which explain the *Passenger Satisfaction*. Tab. 1 reports the latent variables and the manifest variables used in the measurement model.

Tab. 1. Manifest and Latent Variables

LATENT VARIABLES	MANIFEST VARIABLES
Passenger Satisfaction	*Passenger Satisfaction*
Transportation	*Frequency of Service*
	Regularity of Service
	Security
Information & Security	*Passenger Information* *Staff Behavior* *Personal and Financial Security*
Comfort & Cleaness	*Passenger Comfort*
	Station Cleanness
	Train Cleanness

The spatial dependence between the station is formalized by using the spatial lag variable as defined in formula 4. The definition of the spatial weights matrix **W** is based on distances between the 34 Stations, calculated by using walking Google Map.

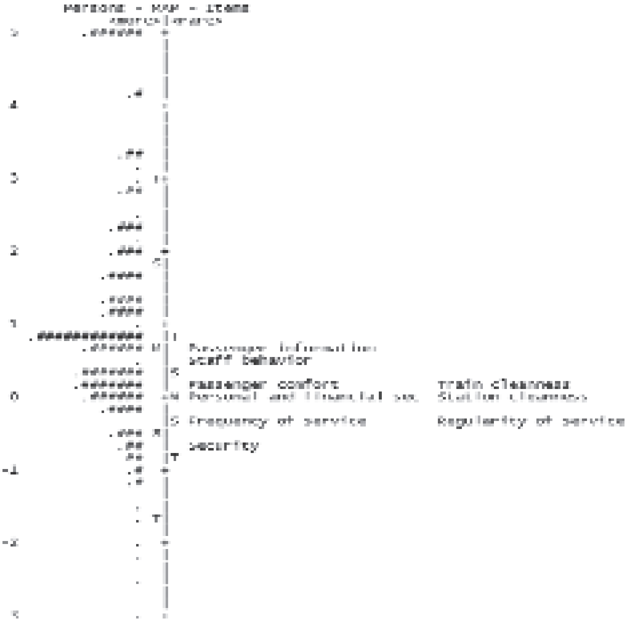
We provide a Passenger satisfaction Model which defines the PS rate in a specific location (Station) h, as a function of the PS rate in a location k (with $h \neq k$). As defined above (*supra* §2.2), the estimation method used is the PLS, performed by using a specific routine developed in Matlab.

3.1 Rating scale model results

The compatibility of the raw data with the Rasch measurement model is verified by several fit statistics. In this case the reliability index observed for items and passengers is 0.99, where the values range between 0 and 1. The estimates for items show how well the replicability of items placement across other passengers measure the same construct index. This result is confirmed by the separation index, whose observed value is equal to 1.

The results for the rating scale analysis of passenger satisfaction are shown in Fig. 1. The vertical line represents the variable passenger satisfaction into log-odds scale. Passengers are aligned to the left and represented by the symbol "#". The more satisfied are on top. Items are aligned to the right. The more the items are on top, the more difficult is to generate satisfaction. It is verified that the distribution of passenger is normal and displayed into higher position than the item distribution. Therefore, passengers have more probability to get satisfaction from metro service.

Fig. 1 - Person-item map for passenger satisfaction



Notes: Each "#" is 27 passengers and aligned to the left of the corresponding log-odds measure of satisfaction. Items are aligned to the right of the corresponding log-odds measure of difficulty to generate satisfaction

More details for item measure are given into Tab. 2. This table lists items in measure order. 'Passenger information' is the attribute of service that has more difficulty to generate satisfaction followed by 'Staff behavior' and 'Train cleanness'. The attributes that have less difficulty to generate satisfaction are 'Security' and 'Regularity of service'. Two types of fit statistics are given for each item. Ideally, the infit and outfit mean-square should be 1.0 for rating scale model, but values included between 0.6 and 1.4 indicate that the deviation from expectation is acceptable (Bond e Fox, 2001). In particular, infit mean-square statistic 1.16 for the item 'Passenger information' is the highest variation between observed data and the Rasch model predicted (16% more variation). 'Train cleanness' and 'Station cleanness' have 18% less variation in the observed response than modeled. Similarly, outfit mean-square for the item 'Passenger information' has the highest variation (20%) and 'Station cleanness' has 17% less variation in the observed data than the model.

Finally, the point-measure correlation is, for each item, a positive value included between 0.58 and 0.68. These values show absence of mis-scoring and normal polarity.

Tab. 2 - Items statistics

Item	Model		Infit	Outfit	Ptmea	Exact	Match
	Measure	S.E.	MnSq	MnSq	Corr.	Obs%	Exp%
Passenger information	.63	.03	1.16	1.20	.63	49.0	50.9
Staff behavior	.45	.03	1.08	1.10	.65	52.4	52.5
Train cleanness	[illegible]	.03	.83	.84	.68	60.2	54.5
Passenger conduct	.10	.03	.89	.93	.64	61.2	53.3
Station cleanness	.08	.03	.82	.83	.67	61.6	55.3
Personal and financial security	0.2	.03	1.11	1.10	.62	54.9	55.6
Frequency of service	-.34	.03	1.05	1.09	.59	56.8	55.7
Regularity of service	-.39	.03	1.07	1.07	.58	56.6	57.8
Security	-.67	.03	.99	.96	.59	61.8	58.8

Notes: Measure is the estimate for the item difficulty to generate satisfaction. S.E. is the standard error of the estimate. Infit MnSq and Outfit MnSq are the infit and outfit mean-square statistic, respectively. Ptmea Corr is the point measure correlation.

The logit measure of each item is obtained for each station by the Differential item function measure (Tab. 3).

Tab. 3 - DIF measure for each item and station

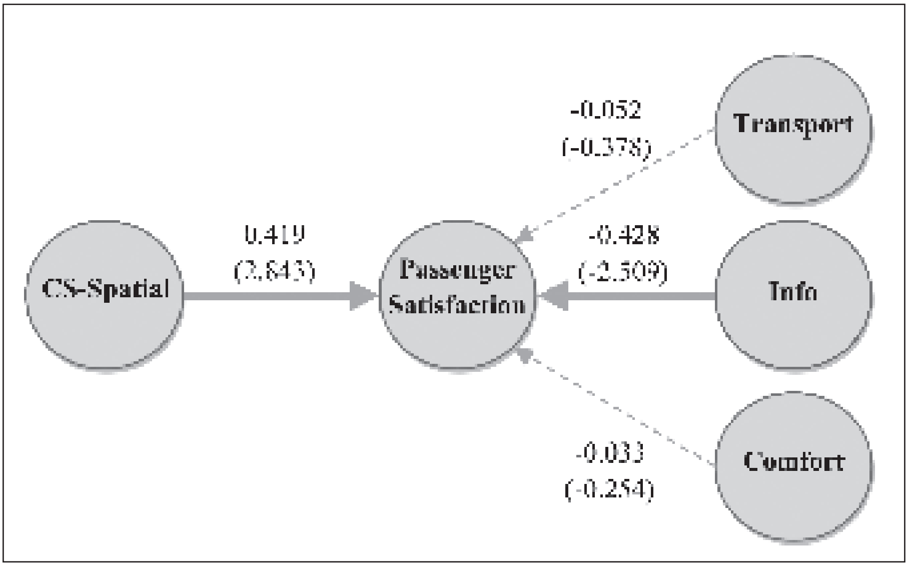
Station	Frequency of service	Regularity of service	Passenger information	Staff behavior	Personal security	Security	Station cleanness	Train cleanness	Passenger comfort
Station1	-0.35	-0.69	0.63	0.57	0.05	-0.62	0.23	0.19	-0.07
Station2	-0.14	-0.02	0.55	0.29	-0.2	-0.32	-0.03	0.21	0.19
Station3	-0.54	-0.34	0.35	0.27	0.01	-0.34	-0.01	0.09	0.53
Station4	-0.68	-0.5	0.45	0.62	0.35	-0.67	-0.15	0.78	0.72
Station5	-0.26	-0.51	0.72	0.74	0.02	-0.59	-0.02	-0.02	-0.13
Station6	-0.31	-0.11	0.5	0.12	0.25	-0.8	-0.13	0.02	0.08
Station7	0.27	0.24	0.54	0.01	0.39	1.21	0.01	0.16	-0.3
Station8	0.21	0.14	0.55	0.67	0.21	1.41	0.31	0.31	0.01
Station9	-0.15	[illegible]	0.38	0.04	0	-1.06	0.3	0.34	0
Station10	0.06	0.06	0.69	0.12	-0.3	-0.92	0.05	-0.02	0.22
Station11	0.23	0.2	0.41	0.23	-0.43	-0.77	-0.17	0.06	0.25
Station12	-0.16	0.08	0.82	0.24	-0.22	-1.17	-0.08	0	0.46
Station13	0.29	0.29	0.27	0.35	0.19	0.28	0	0.12	0.04
Station14	-0.29	0.36	0.42	0.29	[illegible]	-0.86	-0.06	-0.16	0.26
Station15	-0.15	0.28	0.19	0.36	-0.24	-0.93	-0.15	0.19	0.44
Station16	0.22	[illegible]	1.56	[illegible]	0.12	-0.58	0.22	-0.16	-0.16
Station17	-0.47	-0.47	0.97	0.14	0.14	-0.47	0.14	-0.15	0.14
Station18	0.1	0.1	0.63	0.1	-0.34	-0.51	-0.19	-0.31	0.41
Station19	-1.13	-1.25	1.15	0.24	0.39	-0.51	0.47	0.54	-0.42
Station20	-0.24	0.82	1.31	0.31	-0.24	-1.63	-0.24	0.82	-1.55
Station21	0.25	-0.37	0.77	-0.39	0.35	-1.73	0.52	0.72	-0.74
Station22	-0.35	-0.35	0.39	0.04	0.39	-0.55	0.22	0.3	-0.66
Station23	-0.41	-0.41	0.65	0.65	-1.41	-1.75	0.65	0.65	0.33
Station24	-0.29	1	1.43	0.18	-0.07	-0.74	-0.01	0.01	1.21
Station25	0.08	0.09	0.19	0.11	0.01	0.25	0.03	0.06	0.01
Station26	-0.29	-0.39	0.28	0.18	-0.07	-0.74	0.05	1.23	0.01
Station27	-1.16	-1.03	1.93	0.85	-0.05	-0.91	0.05	0.32	0.06
Station28	-0.6	-0.6	1.72	1.73	-0.57	-0.66	-0.56	-0.55	-0.56
Station29	-1.29	-1.29	1.16	1.16	-0.15	-0.15	-0.15	-0.15	0.33
Station30	-0.43	-0.97	1.99	1.59	-0.58	-0.97	-0.73	-0.36	-0.43
Station31	1.08	1.3	0.63	0.66	0.11	0.52	0.56	0.56	0.01
Station32	1.07	1.33	0.85	0.85	0.25	0.53	0.01	0.24	0.06
Station33	[illegible]	-1.57	0.45	0.6	-0.71	-0.51	0.6	1.04	0.25

3.2 S-SEM results

The estimations of the Passenger Satisfaction Model are represented by the path diagram reported in Fig. 2, which is a graphical representation of a priori specified structures and assumptions. The latent variables are drawn by circles and defined using Greek letters. The unidirectional straight arrows in the path diagram represent the causal influence of one variable on another.

Fig. 2 shows also the estimated path coefficients, where the significant relationships have been highlighted by a bold line and non-significant relationships are shown by broken lines. The number in the brackets is the T-test value. The significance of the variables is calculated via bootstrap re-sampling, considering the 100 samples of dimension 120.

Fig. 2 - The Passenger Satisfaction path model



The S-SEM explains the spatial interaction that characterizes the change in passenger satisfaction among the stations. The spatial autoregressive parameter (ρ) and the (β) structural coefficient estimated are significant, while parameter estimates of the other latent variables are not significant.

The spatial lag result suggests that the PS rate in a location *h* has a significant positive dependence from the PS level in another location *k* (0.419). This means the satisfaction in a specific station steps up if the satisfaction in the neighboring areas increases.

The latent variables are measured as the complexity to generate satisfaction. Therefore a negative value of the path coefficient increases the satisfaction, while a positive value decreases the satisfaction. The only latent variable that is significant, with a T-test value greater than 2, is the *Info*, with a path coefficient equal to -0.428, thus implying a good impact in increasing the satisfaction.

For improving the interpretation of the results and for giving a valid support to decision makers, Tab. 4 reports the estimated values of the Latent Variables, obtained as the weighted average value based on the tau (τ) coefficients of the manifest variables.

Tab. 4 - Estimated Value of the Latent Variables

Latent Variables	Estimated Value
Transportation	-0.134
Info	0.276
Comfort	-0.010

The estimated values of the LVs *Transportation* (-0.134) and *Comfort* (-0.010) show that these variables have a low value of difficulty to generate satisfaction, that is the passengers give a good evaluation of these aspects. *Info* (0.276) is evaluated with a high level of difficulty, so the passengers are not satisfied with this aspect.

By combining the results of the path coefficients and the level of PS it is possible to define an intervention matrix by the categorization into two groups of difficulty level as reported in Tab. 5. The LV of the first group has a relative low level of difficulty (good evaluation) and the variables in the second group have a high level of difficulty (bad evaluation).

Tab. 5 - Interventions Matrix

		Difficulty	
		Low	High
Importance	Low (or non significant)	*Transportation* *Comfort*	
	High		*Info*

Tab. 5 can be interpreted as the aspects that can be improved: *Info* is the most important variable in this case study.

Tab. 6 can help in the analysis of which aspects of the *Info* Latent Variable have to be improved. It is possible to read that *PassInfo* (0.8722) and *Staff-B* (0.1173) are the items with major difficulty in creating satisfaction, where instead *PersFin* is the only with a low level of difficulty (-0.2593).

Tab. 6 - Lambda coefficients of the Manifest Variables

Latent Variables	Manifest Variables	Weights	(s.e.)	T-Statistic
Transportation	***FreqServ***	0.3978	0.1771	2.2464
	RegServ	0.6366	0.1507	4.2233
	Security	-0.0429	0.2534	-0.1693
Info	***PassInfo***	0.8722	0.1231	7.0874
	Staff-B	0.1173	0.1593	0.7363
	PersFin	-0.2593	0.157	-1.6518
Comfort	***P-Comf***	-0.3177	0.3168	-1.0029
	St-Clean	0.8226	0.2914	2.8234
	T-Clean	0.2233	0.2709	0.8244

References

Andrich D. (1978), *A rating formulation for ordered response categories*, Psychometrika, 43, 561-574.

Anselin L. (2001), *Spatial Econometrics. In A companion to theoretical econometrics*, ed. B.H.

Bernardini Papalia R. (2006), *Modeling mixed spatial processes and spatio-temporal dynamics in Information theoretic frameworks*, COMPSTAT, in Rizzi A., Vichi M. (Eds), Physica Verlag Heidelberg, New York, 1483-1491.

Bernardini Papalia R., Ciavolino E. (2008), *GME estimation of spatial structural equations models*, in: MTISD (2008), *Methods, Models and Information Technologies for Decision Support Systems*, Lecce, 18 - 20 Settembre 2008, 228-231.

Bond T.G., Fox C.M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*, Erlbaum, London.

Ciavolino E., Al-Nasser A.D. (2009), *Comparing Generalized Maximum Entropy and Partial Least Squares methods for Structural Equation Models*, Journal of Nonparametric Statistics, 21(8), 1017-1036.

Gallo M. (2007), *The scaling problems in service quality evaluation*, Metodološki Zvezki, 2, 165-176.

Gallo M. (2009), *Measuring the passenger satisfaction: A strategy based on Rasch Analysis and Analysis of Means (ANOM)*. Submitted.

Linacre J.M., Wright B.D. (2000), *A user's guide to WINSTEPS: Rasch model computer program,* MESA Press, Chicago.

Wright B.D., Masters G.N. (1982), *Rating scale analysis*, MESA Press, Chicago.

Wold H. (1966), *Estimation of principal components and related models by iterative least squares*, in Multivariate Analysis, ed. P.R. Krishnajah, Academic Press, New York, 391–420.

Wold H. (1973), *Nonlinear Iterative Partial Least Squares (NIPALS) Modelling - Some Current Developments*, in Multivariate Analysis, 3, ed. P.R. Krishnajah, Academic Press, New York, 383–407.

Wold H. (1975), *Path Models with Latent Variables. The NIPALS Approach,* in Quantitative Sociology, ed. H.M. Blalock, New York: Seminar Press, 307–357.

Wold H. (1982), *Soft Modelling: The basic design and some extensions*, in

Systems Under Indirect Observations, part 2, eds. K.G. Jöreskog and H. Wold, North-Holland, Amsterdam, 1–54.

Wold H. (1985), *Partial Least Squares*, in Encyclopaedia of Statistical Sciences, 6, eds. S. Kotz and N.L. Johnson, John Wiley & Sons, New York, 581–591.

G. & L. E. R., Vol. 13 No. 2, 2009

Michele Gallo[1], Antonello D'Ambra, Ida Camminatiello

THE EVALUATION OF PASSENGER SATISFACTION IN THE LOCAL PUBLIC TRANSPORT: A STRATEGY FOR DATA ANALYSIS

Abstract

The diffusion of ISO 9001:2000 certification and adoption of mobility charter led the companies of Local Public Transport (LPT) to carry out surveys of Passenger Satisfaction (PS). However, often, the data analysis is limited to the application of descriptive and explorative statistical techniques. In this way, the collected data are used in an inefficient way and the information that is transferred to the company management is not sufficient to make decisions.

A good analysis strategy requires the use of a combination of parametric and nonparametric techniques. In this paper we propose the combined application of Rasch Analysis and Simple Components analysis based on the RV coefficient (SCA-RV).

JEL classification: C13; C14

Keywords: Passenger Satisfaction; Rasch Analysis; Rating scale model; Simple Components Analysis; SCA-RV; LPT

1. Introduction

The diffusion of ISO 9001:2000 certification and adoption of the principles defined in the Charter of the services of Transport sector led an increasing number of LPT companies to activate procedures for assessing the quality. The concept of quality, however, over the years has been considerably modified by binding more closely to that of satisfaction. In particular, the American Society for Quality defines the quality: "*…the characteristics of a product or service that bear on its ability to satisfy*

[1] "L'Orientale" University of Naples - mgallo@unior.it

stated or implied needs". From that statement one can see how even in the context of LPT, objective indicators measure the purely virtual or potential quality, only the perceived service quality by the customers (passengers, citizens, stakeholders ...) is the real quality.

It is necessary to move the attention to what is perceived, rather than what is supplied. However, this leads to considerable difficulties, in particular as regards the instruments to be used for its assessment.

In fact, the objective parameters are always uniquely determined, and their measurement does not lend itself to different interpretations, while the measurement of satisfaction is more complex because of the its subjective nature and the presence of other factors that are not always easily separable from the aspect that one wants to evaluate.

Since the measurement of passenger satisfaction depends on cognitive emotional and psychological aspects (Oliver, 1993), it is a latent variable and its measurement can only be done through a set of items which change according to adopted conceptual model. Consequently, the analysis of satisfaction can only be achieved through appropriate statistical techniques to estimate these latent aspects and keeping, as far as possible, the effects attributable to such factors separate.

A further problem concerns the multivariate nature of quality (Parasuraman et al., 1988) that requires an analysis of the phenomenon using statistical techniques that consider the different dimensions and allow us to define whether a service is really able to meet the expressed and implied needs of customers.

Given those problems, we have to pay close attention to the followed procedures for the collection, analysis and interpretation of data related to satisfaction. In particular, several conceptual models and procedures have been proposed in the literature. The UNI 11098:2003 represents a first attempt of standardization. However it has the disadvantage that wants to adapt to types of very different services, failing to capture the specific needs of particular services such as LPT those.

To develop a standard that is shared and applied by all companies of LPT is essential, especially if the commissioning body wants to use PS data to implement a system of rewarding or penalizing. One last point concerns the customer on which to focus research, of course, it is immediate to see that the research terms change if we are interested in analyzing only passengers or all potential users of the service.

For all the considerations relating to procedures to be adopted for the PS survey and the choice of the conceptual model to be applied, we refer to D'Ambra *et al.* (2006), Gallo *et al.* (2009) and Gallo (2009). We refer at the same contribution to the considerations relating to the set of parametric and nonparametric methods usable for the study of PS, below, instead, we address the issues concerning the application of Rasch Analysis and Simple Components analysis based on the RV coefficient (SCA-RV).

2. Data analysis

The PS surveys are carried out by administration of a questionnaire composed of a set of items structured to consider the different aspects that go to make a greater or lesser satisfaction of passengers.

These assessments are made using a Likert scale, where the collected data are on ordinal scale, so, before proceeding to their analysis it is necessary to quantify the data or to use a technique that allows their transformation into pseudo-interval values. Moreover the transformation in linear and quantitative measures has to guarantee a right calibration throughout the whole set of the real numbers avoiding an "compression effect" (Wright e Linacre, 1989).

A good analysis strategy requires the use of a combination of nonparametric techniques for exploring the structure of data and, if the modelling assumptions are verified, the use of the parametric methods for generalizing the results to all passengers. Furthermore, the results have to be analyzed to verify the presence or not of a type of passengers that differs significantly in terms of satisfaction, from others.

The combined application of the Rasch Analysis and SCA-RV can be an answer to such needs.

Based on the separability theorem, the Rasch Analysis is a parametric technique that allows to identify what aspects of the LPT service provide greater satisfaction. Among the main characteristics of this technique is the possibility of using ordinal data without their quantification, to obtain interval measures, to get optimal estimates of the parameters and the absence of distributional assumptions.

Unlike the Rasch Analysis, the SCA-RV is a nonparametric statistical technique that can only be used after transforming the ordinal data into pseudo-interval data. The SCA-RV, however, allows to investigate the structure of relationship among the items and to identify latent variables that otherwise would not be observable.

2.1. Rasch analysis

The need to evaluate a not directly observable aspect requires the administration of a questionnaire consisting of several items formulated to investigate the various dimensions of satisfaction. All items have the same set of response categories and it is assumed that the passengers choose a higher category in the scale of assessment only if they are more satisfied.

From these remarks, the assessment can be made on the basis of the ratio between the frequency with which a particular category has been chosen and the frequency with which it has not been chosen, this ratio is a continuous and easily linearizable measure with a logarithmic transformation.

Based on the assumption that the measure of satisfaction is achieved by the logarithm of the ratio between the probability that the passenger *i* scoring in category *m* of the item *j* and the probability that the passenger *i, having chose* the category *m-1* for the item *j,* do not reach the category $m\ (1-P_{ij\,(m)})$:

$$\log it = \ln\left(\frac{P_{ij\,(m)}}{(1-P_{ij\,(m)})}\right) \tag{1}$$

the Rasch model not only allows to get a continuous and linear measure of PS latent variable, but Andrich (1978) and Masters (1982) showed that these estimators also have the optimal properties, such as sufficiency, consistency, efficiency and unbiasedness.

If the survey is consisted of items that are on the same ordinal scale, the best Rasch model for measuring latent variable is the rating scale. Such Rasch model decompose log *it* in the following three quantity:

$$\log it = \xi_i - \delta_j - \gamma_m \tag{2}$$

where ξ_i is the propensity of passengers *i*-th to be satisfied, δ_j is the propensity of item *j*-th to generate satisfaction, γ_m is the difficulty of category threshold *m*.

Sometimes, in addition to quantities already described, it can be useful to identify the existence or not of a different level of satisfaction generated by different companies of LPT or by different lines of the same company. In the latter cases, the probability that the passenger *i* scoring in category *m* of the item *j* has to be calculated on the basis that the passenger chose or not the company (line) k-th, Therefore, the equation (2) can be rewritten:

$$\log\left(\frac{P_{ijk(m)}}{1-P_{ijk(m)}}\right) = \xi_i - \delta_j - \upsilon_k - \gamma_m \tag{3}$$

where υ_k is the difficulty of the company (line) *k*-th to generate high level of satisfaction.

2.2 Simple Component Analysis based on RV coefficient

Among the explorative methods, Principal Component Analysis (PCA) has optimal properties. The data for Customer Satisfaction (CS), however, present the peculiarity of generating all the positive correlations. Therefore, the first component is typically the mean or the sum of observed scores on different items. It is merely a measure of satisfaction.

Moreover all the variables are represented on the same quadrant obtaining little interesting and useful results. To improve the interpretability of the results, several methodological developments have been proposed: Lasso (Jolliffe *et al.*, 2003), Constrained Multivariate Analysis (Choulakian *at al.*, 2006), Simple Component Analysis (Rousson *et al.*, 2003), Simple Principal Components (Vines, 2000), SCA-RV (Gallo *et al.*, 2005). These methods lead to sub-optimal but more useful results for make decisions.

In this paper we apply the SCA-RV. It maximizes the quantity:

$$\left[p_1' C p_1 + \sum_{j=2}^{q} p_j' C_{(j-1)} p_j \right] / tr(C) \tag{4}$$

where C is the correlation matrix among the original variables, q is the rank of C, p_1 is the first principal component, p_j is the *j*-th principal component, *tr* is the trace operator.

This maximization criterion assures equivalent results to PCA only in case of uncorrelated components, while it is a penalized version of PCA criterion for correlated components, because less variability is extracted from the original variables.

SCA-RV is based on two stages SCA algorithm proposed by Rousson and Gassen (2003). Fixed the number of components q and the number of block b, the first stage of SCA classifies the p original variables into b disjoint blocks. The approximate block-structure in the correlation matrix leads to a maximal within block correlations and in the meantime to a minimal between blocks correlations. The authors solved this problem with an agglomerative hierarchical procedure based on a dissimilarity measure between clusters called "median" linkage alternative to the possible "single" or "complete" linkages. Coming from the loadings matrix corresponding to the b simple block-components of the first stage, the second step of the algorithm is based on a suitable difference-component shrinkage procedure of the sequential first components of the residual variables obtained by regressing step by step the original variables on the first ($J-1$) simple components.

The SCA-RV modifies the first stage criterion. Instead to use an agglomerative hierarchical procedure based on simple correlation coefficient, which can lead to very different solution with a choice of a possible different link criterion, it uses the RV vectorial correlation coefficient (Robert and Escoufier, 1976). This coefficient gives a measure of similarity of the two configurations, taking into account the possibly distinct metrics to be used on them to measure the distances between points.

3. A case study

The capability of the techniques is illustrated by a PS survey carried out on a sample of 1933 passengers. The survey is composed of seven items for evaluating the PS.

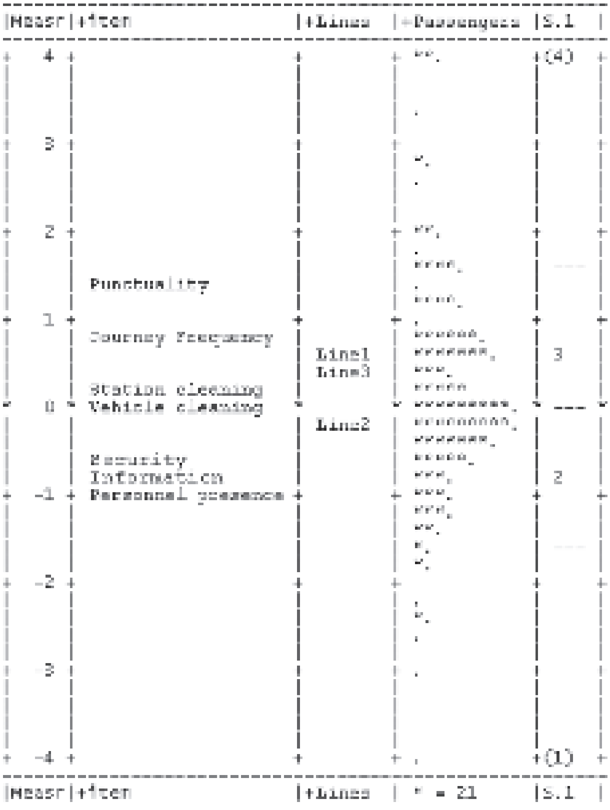
All the items have four response categories (1 - negative, 4 - positive) and concern "punctuality" "frequency of the journeys", "cleaning of stations", "cleaning of vehicles", "information", "presence of personnel" and "security".

The 1933 questionnaires has been collected on three separate lines that characterize the LPT company under consideration. In particular, 1124 interviews have been carried out on Line 1, 354 on line 2 and 455 on line 3. The statistical software Facets has been utilised for Rasch Analysis and the meta-language S-plus for SCA-RV.

3.1 The results of Rasch analysis

The Rasch Analysis shows that the items that generate higher customer satisfaction are the "punctuality" and the "frequency of journey", while the aspects of service that are less satisfactory are the "presence of personnel", the "information" and the "security". From the point of view of the three lines, it can be showed that the line 1 and 3 generate a higher satisfaction level than the line 2.

Fig. 1 - Item map



To verify the goodness of fit of the model we focus on Infit and Outfit statistics and point-biserial correlation coefficients. The Tab. 1 show the measure for the items and their Standard Errors (S.E.), Infit and Outfit statistics and the point-biserial correlation coefficients. The Infit and Outfit statistics are in the range [0.8; 1.3] so the model fits each item well. There are no outliers.

Tab. 1 - Item statistics: measure, S.E., Infit, Outfit, point-biserial correlation coefficients

N	item	Measure	Model S.E.	Infit MnSq	Outfit MnSq	PtBis
1	Punctuality	1.31	.04	1.1	1.0	.44
2	Journey frequency	.81	.03	1.1	1.2	.36
3	Station cleaning	.23	.03	.8	.8	.50
4	Veheicle cleaning	.01	.03	.8	.8	.52
7	Security	-.63	.03	1.2	1.3	.34
5	Information	-.78	.03	1.0	1.1	.36
6	Personnel presence	-.95	.03	1.0	1.0	.39

The goodness of fit is confirmed by reliability index and by point-biserial correlation coefficients and by chi-squared test. The reliability index, equal to 1.0, ensures the reproducibility of obtained results. All the point-biserial correlation coefficients are positive values, so there no problem in the dataset. The chi-squared test is significant ($\chi_6^2 = 3889.6$), the items do not generate the same level of satisfaction.

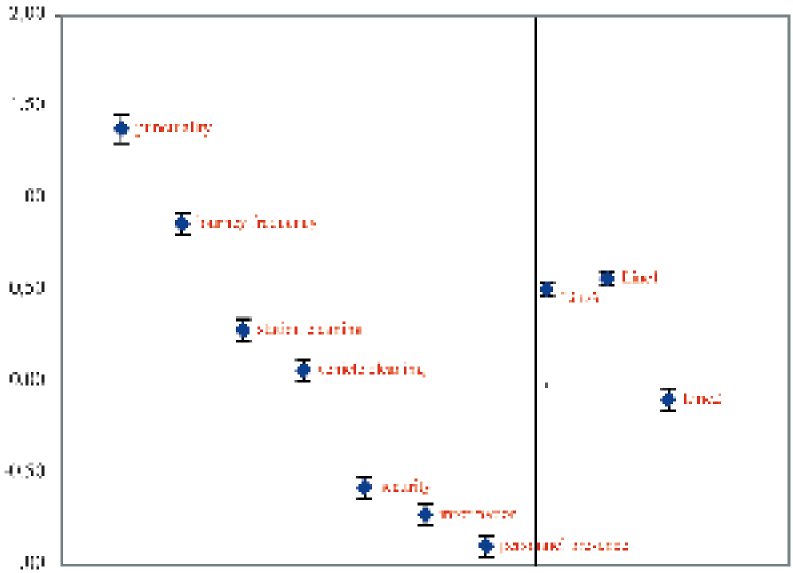
Also for the lines the goodness of estimates is confirmed by reliability index, equal to 0.99, by point-biserial correlation coefficients (all positive) and significant chi-squared statistics.

Tab. 2 - Line statistics: measure, S.E., Infit, Outfit, point-biserial correlation coefficients

N	Lines	Measure	Model S.E.	Infit MnSq	Outfit MnSq	PtBis
1	Line1	.53	.02	1.0	1.0	.35
3	Line3	.47	.02	1.0	1.0	.36
2	Line2	-.15	.03	1.0	1.1	.29

The Fig. 2 presents an hierarchy for the items and the lines. Moreover it shows that all the items are well separated among them and there are no significant differences between the lines 1 and 2.

Fig. 2 - Confidence interval for items and lines



The Tab. 3 gives an overview of results for the categories. We can observe that only the 16% of the passengers chose the category "negative", the 23% "slightly negative", while the 32% "fairly positive" and the 28% "positive." The diagnostics confirm that the estimates are good, so the results are reproducible

Tab. 3 - Category statistics.

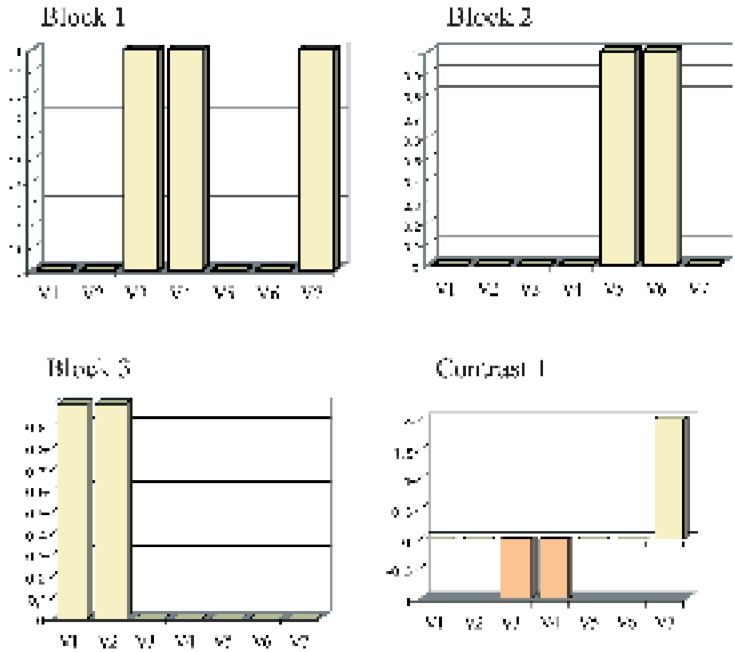
DATA Category Score	DATA Counts Used	DATA %	Cum. %	QUALITY CONTROL Avge Meas	QUALITY CONTROL Exp. Meas	QUALITY CONTROL OUTFIT MnSq	STEP CALIBRATIONS Measure	STEP CALIBRATIONS S.E.	EXPECTATION Measure at Category	EXPECTATION Measure at -0.5
1	2129	16%	16%	-1.17	-1.15	1.0			(-2.34)	
2	2950	23%	39%	-.25	-.26	1.0	-1.02	.03	-.73	-1.60
3	4197	32%	72%	.60	.59	1.0	-.19	.02	.67	-.05
4	3846	28%	100%	1.59	1.60	1.1	1.22	.02	(2.45)	1.84

3.2 The results of SCA-RV

After transforming the ordinal data into pseudo-interval data trough the Rasch analysis, we applied the SCA-RV on the seven items for evaluating the PS.

The SCA-RV shows the presence of three blocks of variables and a contrast. In particular, observing the Fig. 3 we can see the first block (that explains the 28.2% of variability) is consisted of the items "Security of Station" (V7), "cleaning of vehicle" (V3) and "cleaning of station" (V4), the second block (that explains the 23.5% of variability) is characterized by the items "information"(V5) and "presence of personnel" (V6), the block 3 (that explains the 21.6% of variability) is characterized by the items "punctuality" (V1) and "frequency of journeys", finally the first contrast (that explains the the 11.1% of variability) sets against the items which constitute the first group emphasizing that the greatest differences occur between the cleanliness and safety.

Fig. 3 - SCA-RV results: the first component explains the 28.2% of variability (Block 1), the second one the 23.5% (Block 2), the third the 21.6% (Block 3), the fourth the 11,1% (Contrast 1).

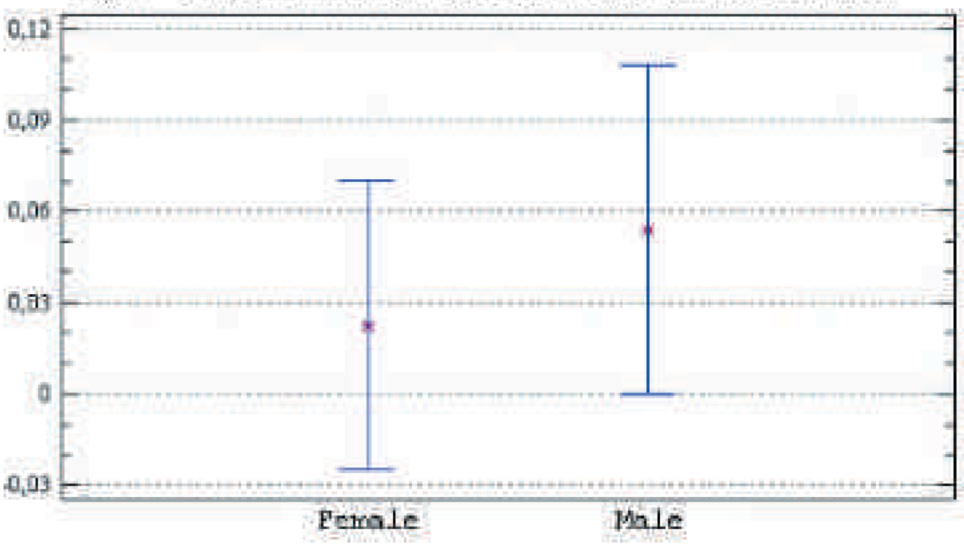


The variability explained by the first four principal components is 84%, The variability explained by SCA-RV solutions is 76%. Thus, the SCA solution is 90% optimal.

4. Conclusions

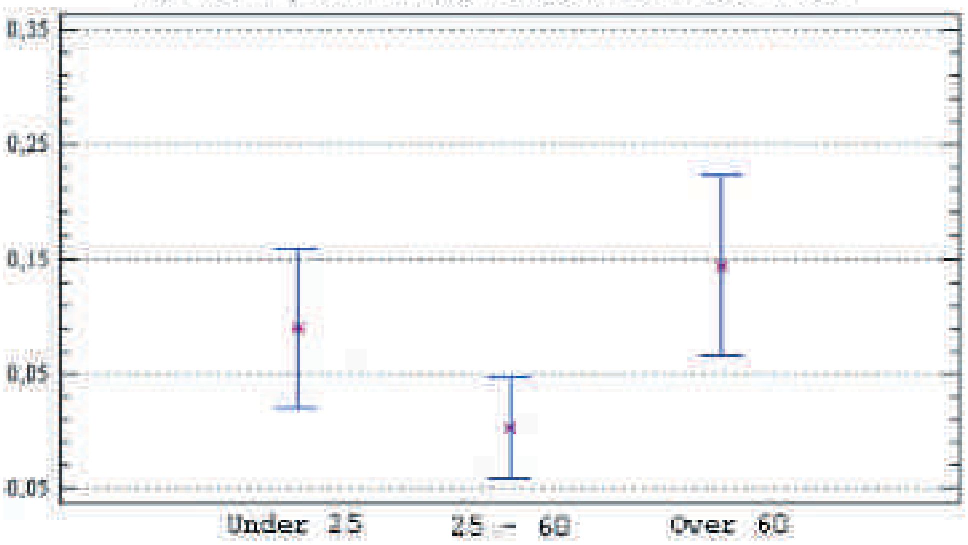
In this contribution we tried to illustrate in a simplified way that a combined approach of different statistical techniques can lead to useful and not repetitive information. We can further observe that quantitative measures allow to company managers to identify which types of passengers are less satisfied. In particular, considering the gender of passengers it is possible to point out that the women are less satisfied than the men. However, these differences are not statistically significant at 5% level.

Fig. 4 - Comparison between the level of PS for females and males.



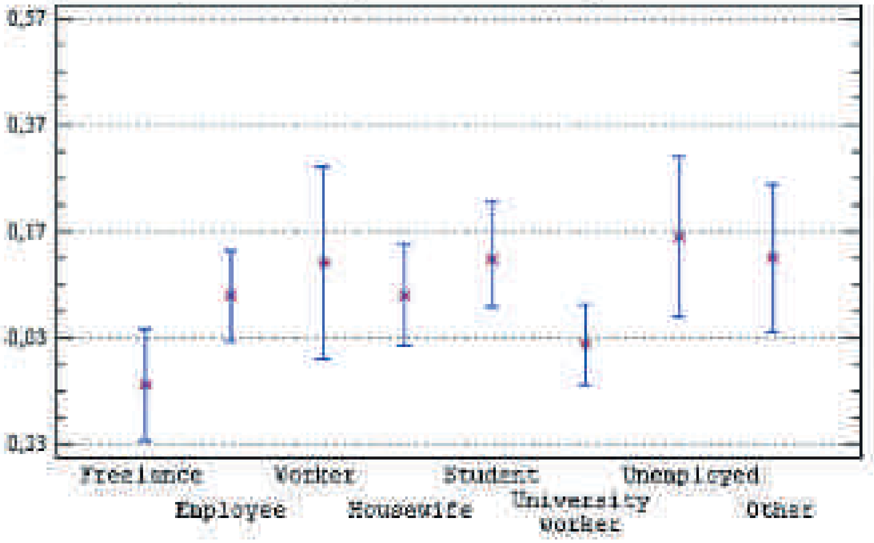
Concerning the age of the passengers, we can highlight a significant difference in the level of satisfaction between the passengers over 60 years and those aged 25-60, instead, there are no significant differences between the younger (under 25 years) and the other two categories, although the average satisfaction level of younger is higher than the passengers aged 25-60.

Fig. 5 - Comparison among the level of PS for age classes.



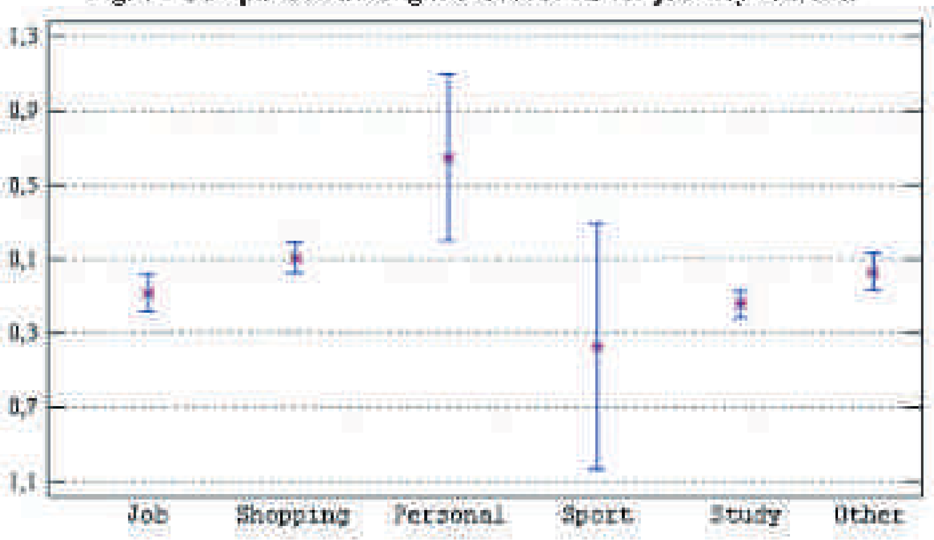
Considering the profession of the passengers, there are significant differences between the level of satisfaction of freelancers and students. The freelancers are less satisfied than the students. There are no significant differences among the other professions.

Fig. 6 - Comparison among the level of PS for professional activities.



Concerning the trip reason, the passengers who travel for personal reasons are statistically more satisfied than those using the LPT services for work, study or other reasons.

Fig. 7 - Comparison among the level of PS for journey reasons.



In conclusion, we can observe that, in many cases, the companies carry out PS surveys (expressly provided for the mobility charter). However, often, the data analysis is limited to the application of descriptive and explorative statistical techniques. In this way, the collected data are used in inefficient way and information that are transferred to the company management are not sufficient to make decisions. Furthermore, such information could include the only interviewed passengers.

References

Andrich D. (1978), *A rating scale formulation for ordered response categories*, Psychometrika, 43, 561-573.

Choulakian V., D'Ambra L., Simonetti B. (2006), *Hausman principal component analysis*, in From Data and Information Analysis to knowledge enginnering, Springer.

Cronin J.J., Taylor S.A. (1992), *Measuring Service Quality: A Reexamination and Extension*, Journal of Marketing, 56, 3.

D'Ambra L., Ciavolino E. (2006), *La Customer Satisfaction nell'ambito del Trasporto Pubblico Locale*, Proceedings AMTS.

Franceschini F. (2001), *Dai prodotti ai servizi. Le nuove frontiere per la misura della qualità,* UTET Libreria.

Gallo M. (2005), *Un Approccio Statistico per la Misurazione della Propensione agli Spostamenti*, Amenta P., D'Ambra L., Squillante M., Ventre A., *Metodi, Modelli e Tecnologie dell'informazione a supporto delle decisioni*, Franco Angeli, Milan.

Gallo M., Amenta P., D'Ambra L. (2005), *Simple component analysis based on RV coefficient,* Proceedings Cladag'05, 357 – 360.

Gallo M., Fiorenzo D. (2009), *A strategy to evaluate the satisfaction of different passenger groups,* In: Classification and Data Analysis. Catania, September 9-11, cleup, 1, 501-504, ISBN/ISSN: 9788861294066.

Gallo M. (2009), *Measuring the passenger satisfaction: A strategy based on Rasch Analysis and Analysis of Means (ANOM),* Submit**.**

Jolliffe I.T., Uddin M. (2003), *A modified principal component technique based on the lasso*, Journal of Computational and Graphical Statistics, 12, 531 - 547.

Linacre J.M. (1989), *Many-Facet Rasch Measurement*, MESA Press, Chicago.

Linacre J.M. (2004), *Facets Rasch measurement computer program*, Winsteps.com, Chicago.

Masters G.N. (1982), *A Rasch model for partial credit scoring*, Psychometrika, 47.

Oliver R.L. (1993), *A conceptual model of service quality and service satisfaction: compatible goals, different concepts*, in Advances in Service Marketing and Management: Research and Practice, 2. JAI Press.

Parasuraman A., Zeithaml V. Berry L.L. (1988), *Servqual: a multiple item scale for measuring customer perceptions of service quality*, Journal of Retailing, 64(1), 12-37.

Rousson V., Gassen T. (2003), *Some case studies of simple component analysis*. Manuscript on http://www.unizh.ch/biostat/Manuscripts.

Tesio L. (2003), *Measuring Behaviours and Perceptions: Rasch Analysis as a Tool for Rehabilitation Research*, J. Rehabil Med., 35, 105-115.

UNI 11098 (2003), *Sistemi di gestione per la qualità - Linee guida per la rilevazione della soddisfazione del cliente e per la misurazione degli indicatori del relativo processo,* Ente Italiano di Unificazione.

Vines S.K. (2000), *Simple Principal Components*, Applied Statistics, 49, 441–451.

Wright B.D., Masters G. N. (1982), *Rating Scale Analysis*, MESA Press, Chicago.

Wright B.D., Linacre J.M. (1989), *Observations are Always Ordinal: Measures, However, Must be Interval,* Archives of Physical Medicine and Rehabilitation, 70.

Zanella A. (1999), *A Statistical Model for the Analysis of Customer Satisfaction: some theoretical aspects,* Statistica, 59.

G. & L. E. R., Vol. 13 No. 2, 2009

Cristina Bernini[1], Alessandro Lubisco[2]

MODELLING DYNAMIC CUSTOMER SATISFACTION IN URBAN PUBLIC TRANSPORTATION

Abstract

The dynamic analysis of Customer Satisfaction is particularly useful for monitoring either customer satisfaction over time or customer behaviour reaction to company strategies, and therefore for measuring their effectiveness and efficiency. In the paper we address the following questions: does customer satisfaction change over time? What are the main factors affecting overall satisfaction? Does customer characteristics influence overall satisfaction? In investigating these hypotheses, we propose an Extended dynamic version of LISREL model, re-written in the Observed Form Solution. The model is estimated by 2SLS estimators on data drawn from the Tram Customer Satisfaction Surveys (Rimini) over the period 2000- 2006.

JEL classification: C13; C14

Keywords: public transport; Customer Satisfaction; dynamic LISREL models; instrumental estimator; pseudo-panel

1. Introduction

In the last decades, the aim to reduce car traffic and the need of saving time and money to travel has yielded a rapid growth of urban public transportation. Travel companies and public administrations have shown an increasing interest in understanding what determines individual travel mode choices, what are the individual preferences

[1] Department of Statistical Science "Paolo Fortunati" cristina.bernini@unibo.it
[2] Department of Statistical Science "Paolo Fortunati" alessandro.lubisco@unibo.it

about the urban transport and whether they change over time. In this framework, the Customer Satisfaction analysis reveals to be useful in investigating relationships between overall customer satisfaction and perceived quality of services supplied by travel companies (Golob 2003; Thögersen 2001; Cagnone *et al.* 2003). Moreover, the dynamic analysis of Customer Satisfaction is helpful in monitoring as customer behaviour and satisfaction change in respect to company strategies, contributing to evaluate their effectiveness and efficiency.

The paper attempts to answer the following questions: What are the main factors affecting the overall satisfaction towards the public transport? Does overall customer satisfaction change over time? Do determinants of satisfaction change over time? Do customer characteristics affect public transport evaluations?

We propose investigating these hypotheses within the LISREL modelling framework (Jöreskog and Sörbom, 1977). A generalisation of the Simplex model (Jöreskog 1970, 2001) to an original dynamic version, the so called Extended Simplex Model – ESM, (Bernini and Lubisco 2005a, 2005b, 2006, 2007) is used in the empirical analysis. As regards model estimation, two different approaches are considered. The traditional WLS (Weight Least Square estimator) for structural equation latent variable models is compared to the 2SLS estimator for latent variable models re-written in the Observed Form Specification (OFS). As suggested by Bollen (1996), instrumental estimators may be superior to full information estimators when normality and misspecification assumptions are not longer satisfied, as in our case.

Dynamic OFS-LISREL models are estimated on an empirical data set obtained from the Customer Satisfaction Survey, carried out by the Faculty of Statistics for the TRAM Agency of public transport in Rimini in the period 2000-2006. Dynamic analysis requires panel data but unfortunately they are not usual, as in our case. More frequently, data are collected from Independent Repeated Surveys (IRS) in which a different sample of customers from the same population is observed in several time occasions. In the analysis we propose using the average cohort technique to overcame the problem of the lack of panel data. This technique transforms IRS data to the so called pseudo-panel, (Deaton 1985; Browing *et al.* 1985), allowing satisfaction longitudinal analysis by dynamic LISREL models.

2. The conceptual model and research hypotheses

The aim of public transport agency is to analyse customer satisfaction in order to increase the public transport choice. In this context, the identification of the determinants influencing public transport mode choice becomes relevant. To this goal we propose a conceptual model which guides the following research. The model extends and integrates several research streams on public transport field.

As regard the determinants influencing travel mode choices, the simplest distinction is between *external conditions* and *individual traveller characteristics*. Travel mode depends on the availability of a public transport alternative for destination and time, on the locations of shops, jobs, and homes and on the transport infrastructure. Environment conditions, such as road construction and weather, have also been shown to influence the choice of travel mode. Among individual characteristics influencing travel mode choices are age, gender, attitudes, income and habits (Jakobsson *et al.* 2002; Fujii *et al.* 2001, Gärling *et al.* 2001). There is empirical evidence that the travel mode choice depends on individual age and there are also arguments that using public transport tends to become habitual. In particular, travel mode choice is dynamic: determinants are influenced by (past) behaviour. Thus habits are the result of performing the same behaviour frequently and in a stable context. Behaviour may also lead to change in attitudes and in perceived control due to experience-based learning, self-inferences and cognitive dissonance.

Psychologists and marketing researchers also distinguish between *volitional* and *non-volitional determinants* of behaviour (Bagozzi and Warshaw 1990; Bagozzi 1994; Peter *et al.* 1999). External conditions are considered non-volitional, such as some of the individual characteristics (habits and individual characteristics). Another non-volitional individual characteristic that has been found to influence travel mode choices is the knowledge of public transport services such as routes, timetable, ticket price. Moreover, traveller evaluations, preferences and motives (volitional determinants) directly affect travel mode choice.

The focus in the paper is limited to the determinants that have been found to have a fairly direct influence on public transport choice. Consistent with the outlined conceptual framework, the study is based on the assumption that public transport choices are partly volitional (influenced by the traveller satisfaction and expectation) and partly determined by non-volitional individual characteristics (transport services, habits, age). Therefore travel mode choices are based on some co-determinants: the traveller evaluation (about travel characteristics), personal abilities (expectations, habits) and individual characteristics (age).

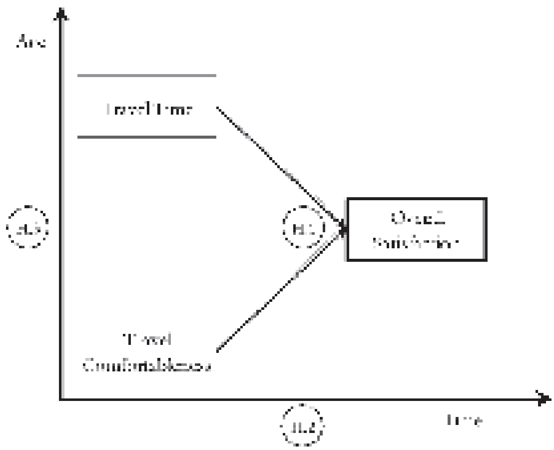
In this framework, our conceptual model considers time travel and travel comfortableness as antecedents of overall customer satisfaction. We suppose they are separate constructs allowing to change over time and over age (Fig. 1). Then, we posit the following hypotheses:

H.1. the greater the extent of satisfaction with time travel and comfortableness, the higher the level of overall satisfaction

H.2 the influence of time travel and comfortableness on the overall satisfaction changes over time

H.3. the overall satisfaction changes with respect to customer age.

Fig. 1 – Public transport conceptual model of Customer Satisfaction



3. The case study: the public transpot company - Tram

In the following, we present the main features of the survey that TRAM, the Public Transport Agency in Rimini (Italy), and the Faculty of Statistical Sciences - University of Bologna carried out from year 2000 to 2006. The aim is to evaluate travel mode choice, perceived quality of public services and level of customer satisfaction.

3.1 The data collection procedure

The population of interest consists of urban transport customers who choose the bus to move in the province of Rimini. The sampling design is based on a stratification with respect to several characteristics of transport services (bus-routes, time bands, days) and service typologies (urban, suburban).

For every year of observation, one thousand questionnaires were administrated to customers, randomly chosen, by mean of a "face to face" technique on the bus during the trip. The surveys was conducted in the 2nd week of December of every year.

3.2 The questionnaire

The questionnaire is composed of four main sections: "Evaluations", "Judgments", "Habits" and "Personal data".

The "Evaluations" section starts with a question to pick out the most important reason of the current trip. The main part of this section includes 15 items representing various aspects of the service, such as Bus frequency and speed, Travel comfort and safety, Information and Tickets availability, Ticket price, Services for disabled people, Care of environment, Driver kindness. For each item, customer was asked to give a score, expressed in terms of satisfaction and importance, using a Likert scale (1 – 10).

In the "Judgment" section, all respondents were asked to give an evaluation on the trip expectation, declaring whether the transport services were better, worse or equal to what they expected to have. Customers were also required to give a score to the Global satisfaction of public transport services.

"Habits" section observes the characteristics of TRAM customers for what concerns how and why they use the urban transport service to move. Respondents were also asked about the principal reason for which they "usually" use the bus, allowing to distinguish between what we called "forced" and "unforced" customers.

In the "Personal data" section information on customer age, gender, nationality, residence, educational degree and occupation were collected.

3.3 Data description

The descriptive analysis of satisfaction scores over time evidences some interesting findings.

Tab. 1 summarizes the variables used in the analysis, their operationalizations and mean values for the samples in 2000, 2002 and from 2004 to 2006 respectively. It can be observed that the average level of satisfaction is quite high, ranging from 7 to 8 for the majority of the items. Only the items Bus shelters and Ticket availability present some values below 6.

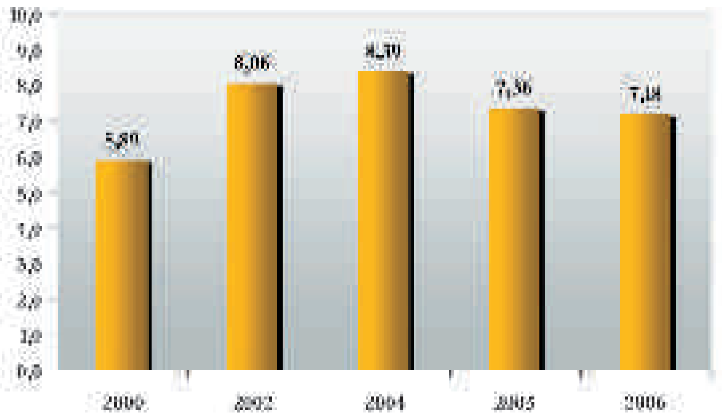
Moreover, the mean values of all the items do seem to change relevantly between time occasions, supporting the hypothesis that customers satisfaction evolves over time.

Tab. 1 – Items descriptions, mean values and standard errors (in brackets)

Item	*Operationalization*	*2000*	*2002*	*2004*	*2005*	*2006*
Convenience	Bus-Routes and Stops Convenience	6.97 (1.43)	8.27 (1.21)	8.32 (0.76)	7.41 (2.03)	7.35 (1.82)
Punctuality	Bus Punctuality	6.67 (1.24)	6.83 (1.52)	7.69 (0.93)	6.39 (2.18)	6.44 (2.18)
Frequency	Bus Frequency	6.78 (1.30)	7.26 (1.36)	[illegible] (1.05)	6.29 (2.24)	6.72 (2.17)
Speed	Transfer Speed	7.18 (1.03)	7.51 (1.20)	8.76 (0.8[illegible])	7.47 (1.53)	7.08 (1.89)
Cleanliness	Bus Cleanliness	6.32 (1.79)	8.04 (1.20)	7.96 (0.89)	7.21 (2.08)	6.60 (2.28)
Comfort	Travel Comfort: Seats, Crowding, Air-Condit	6.01 (1.29)	7.95 (1.16)	8.07 (0.87)	7.05 (2.17)	6.66 (2.23)
Safety	Travellers Personal and Property Safety	6.96 (1.[illegible])	8.71 (1.14)	8.22 (0.8[illegible])	7.93 (1.92)	7.38 (2.06)
Bus Shelters	Bus Shelters:	4.95	6.12	7.18	6.38	6.77

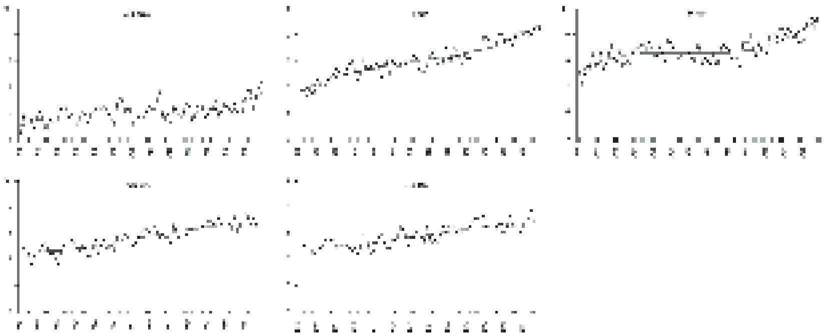
The analysis of the Global Satisfaction Index (GSI), calculated as the average of the Global Satisfaction scores on all the respondents, shows a growth until 2004 when it reaches the highest value (8.39). In the following years, the GSI decreases and in 2004 its value is 7.18. The decline in the Global evaluation is only partially explained by some modifications of the sampling design. In the last two years, the Survey included the suburban routes, largely used by students, who are less satisfied than other customers of the public transport services.

Fig. 2 The Global Satisfaction Index (2000-2006)



The analysis of satisfaction with respect to customer age shows other interesting findings. Satisfaction has an increasing age profile: satisfaction grows slowly in initial phases of life-cycle and peaks its maximum value in correspondence to older customers (Fig. 3). As we see below, this result reveals to be useful for analysing quality evaluations over time.

Fig. 3 – The profiles of Global Satisfaction Index (GSI) in years from 2000 to 2006



4. Lisrel dynamic models

4.1 The simplex and the extended simplex models

In order to evaluate the main factors influencing the TRAM customers overall satisfaction and its dynamic over time, the LISREL approach is used (Jöreskog, 1970). The first extension of the LISREL model to longitudinal studies has been proposed by Jöreskog (1970, 1979 and 2001; *et al.* 1977; and Sörbom, 1977). Jöreskog introduced dynamics in the classic LISREL model only by the latent endogenous variables. This model, called SIMPLEX model, is based on the hypothesis that the latent endogenous variable is generated by an AR(1) non-stationary process.

Bernini and Lubisco (2005b) propose an original Extension of the SIMPLEX model (ESM) according to which latent exogenous variables affect, in the structural model, as in the classic LISREL specification. Thus, in the ESM specification the latent endogenous variables, η at time t, depend on the latent endogenous variables η at time *t-1* and on the exogenous latent variables ξ at the current time *t:*

$$\eta_t = B_t \eta_{t-1} + \Gamma_t \xi_t + \zeta_t \qquad (1)$$

$$\mathbf{y}_t = \Lambda_{yt}\eta_t + \varepsilon_t \qquad (2)$$

$$\mathbf{x}_t = \Lambda_{xt}\xi_t + \delta_t \qquad (3)$$

with $\eta_1 = \Gamma_1\xi_1 + \zeta_1$ ($\mathbf{B}_1 = \mathbf{0}$), being η_t, ξ_t vectors of endogenous and exogenous variables, of dimensions m_t and n_t, $\mathbf{y}_t$ and $\mathbf{x}_t$ vectors of observable variables, of dimensions p_t and q_t, $\mathbf{B}_t$ a regression matrix of order $m_t \times m_{t-1}$, Γ_t a coefficient matrix of order $m_t \times n_t$; Λ_{yt} and Λ_{xt} are matrices of factor loadings of dimensions $p_t \times m_t$ and $q_t \times n_t$, ε_t and δ_t the corresponding vectors of unique factors. ε_t are assumed to be uncorrelated with all η_t, δ_t uncorrelated among themselves and between occasions, and ζ_{t+1} uncorrelated with η_t.

4.2 The ESM in the observed form solution

As regards estimation, structural equation modeling (SEM) with latent variables is still dominated by full information estimators such as Maximum Likelihood (ML) and Weighted Least Squares (WLS), but recently the instrumental variable estimators in general and the 2SLS estimators in particular have received more attention. Full information estimator simultaneously estimates all parameters, using information from the whole system of equations. Limited information estimators are less common in these general SEM, though they are more common in simultaneous equations without latent variables.

The full estimators dominant position in SEM is due to several factors. One practical reason is that full estimators are the default estimator in SEM software. Another reason for their popularity is that under correct model specification and with observed variables that come from distributions with no excess multivariate kurtosis, the full estimators are consistent, asymptotically unbiased, asymptotically efficient, and asymptotically normal, and they allow to estimate the asymptotic covariance matrix of the parameter estimator. The 2SLS estimator has the same properties except that it is asymptotically efficient among limited information estimators rather than full information estimators. Then, some large-sample efficiency advantage are expected for the full estimator when the underlying assumptions are met (Bollen 1996).

The problem with this analytical description of the ML/WLS and 2SLS estimators is that they assume ideal conditions that commonly fail in practice. In real applications the observed variables typically come from nonnormal distributions with excess kurtosis; the sample size might not be large; and the model is nearly certain to have structural specification errors. The 2SLS estimator is asymptotically "distribution free" so that in large samples its asymptotic standard errors and significance tests do not depend on normality. Furthermore, the econometric literature on 2SLS in multiequation models with observed variables suggests that the 2SLS estimator is less sensitive to structurally misspecified models than are full information estimators.

Bollen (1996, 2001) proposed a limited information 2SLS estimator for latent variable SEM. The Author considers the latent variable model

$$\eta = \alpha_\eta + B\,\eta + \Gamma\xi + \zeta \tag{4}$$

and the two equations for the measurement model

$$\mathbf{y} = \alpha_y + \Lambda_y\eta + \varepsilon \tag{5}$$

$$\mathbf{x} = \alpha_x + \Lambda_{\mathbf{x}}\xi + \delta \tag{6}$$

where the notation is similar to that used in eq. (1)-(3), except for the subscript t. α_η, α_y and α_x are intercept vectors. ε and δ have expected values of zero and are uncorrelated with each other and with ζ and ξ.

To apply the 2SLS estimator to equations (4)-(6) requires that each latent variable has a single observed variable to scale it, such that

$$\mathbf{y}_1 = \eta + \varepsilon_1 \tag{7}$$

$$\mathbf{x}_1 = \xi + \delta_1 \tag{8}$$

where $\mathbf{y_1}$ and $\mathbf{x_1}$ are the vectors of scaling indicators. Then, eq. (7) and (8) can be reexpressed as

$$\eta = \mathbf{y}_1 - \varepsilon_{\mathbf{1}} \tag{9}$$

$$\xi = \mathbf{x}_1 - \delta_1 \tag{10}$$

and following Bollen (1996, 2001), the latent variable and measurement models can be rewritten as

$$\mathbf{y}_1 = \alpha_\eta + \mathbf{B}\mathbf{y}_1 + \Gamma\mathbf{x}_1 + \mathbf{u} \tag{11}$$

where $\mathbf{u} = \varepsilon_1 - \mathbf{B}\varepsilon_1 - \Gamma\delta_1 + \zeta$. Note that this manipulation recast the latent variable model into a simultaneous equation model where all variables are observed except for the composite disturbance error. Eq. (11) appears to match the classical econometric simultaneous equation model but un important difference is that we cannot assume that **u** and $\mathbf{x_1}$ are uncorrelated. Recall that **u** contains δ_1 and that δ_1 correlates with those x's that are measured with error. That is in general $\mathbf{x_1}$ is not a vector of exogenous (predetermined) variables, though some x's in $\mathbf{x_1}$ that are error-free could be exogenous.

Let consider a single equation from (11). The i-th equation from $\mathbf{y_1}$ is:

$$\mathbf{y}_i = \alpha_i + \mathbf{B}_i\mathbf{y}_1 + \Gamma_i\mathbf{x}_1 + \mathbf{u}_i \tag{12}$$

where y_i is the i-th y from $\mathbf{y_1}$, α_i is the corresponding intercept, $\mathbf{B_i}$ is the i-th row from **B**, Γ_i is the i-th row from Γ and u_i is the i-th element from **u.** Defining $\mathbf{A_i}$ to be a column vector that contains α_i and all the nonzero elements of $\mathbf{B_i}$ and Γ_i strung together in a column. Let N equal the number of cases and $\mathbf{Z_i}$ be a N row matrix that contains 1's in the first column and the N rows of elements from $\mathbf{y_1}$ and $\mathbf{x_1}$ that have nonzero coefficients associated with them in the remaining columns. The N*1 vector $\mathbf{y_i}$ contains the N values of y_i in the sample and $\mathbf{u_i}$ is a N*1 vector of the values of u_i. Then we can rewrite eq. (12) as:

$$\mathbf{y}_1 = \mathbf{Z}_i\mathbf{A}_i + \mathbf{u}_i \tag{13}$$

OLS is inappropriate estimator for eq. (13) because in most cases at least some of the variables in $\mathbf{Z}_i$ will be correlated with $\mathbf{u}_i$. The 2SLS provides an alternative consistent estimator for of $\mathbf{A}_i$.

The 2SLS estimator requires instrumental variables (IVs) for $\mathbf{Z}_i$. The

IVs must be: (1) correlated with $\mathbf{Z}_{i;}$ (2) uncorrelated with $\mathbf{u}_{i;}$ (3) sufficient in number so that there are at least as many IVs as the number of explanatory variables on the right-hand side of the equation. The squared multiple correlation coefficient from the regression of the variables to be replaced on the IVs provides a check of condition (1). In equations in which there are more IVs than required, we refer to the equation as overidentified. In these cases, we can test the suitability of the IVs, allowing to verify condition (2). There are several tests of the overidentifying conditions, among others in the empirical study we use the Hansen's J statistics.

4.3 How constructing pseudo-panel data for dynamic customer satisfaction analysis?

The dynamic analysis of customer satisfaction requires using panel data but unfortunately surveys repeated over time on the same customers are not usual. More frequently, the data are collected from Independent Repeated Surveys (IRS), in which different samples of customers from the same population are observed in several time occasions. The application of the *average cohort techniques* on IRS data leads to the so called *pseudo-panel*: the use of variables that do not change over time (for example the year of birth of the customers) let us both to define groups of homogenous individuals and to follow them as a panel (Deaton, 1985; Browing *et al*., 1985). Empirical evidence suggests that customer satisfaction changes with respect to customer age, and this represents an interesting feature in our analysis.

As we see before (Fig. 3), satisfaction and quality evaluation change over customer age and therefore the average cohort technique can usefully used to explore the dynamics in the quality evaluation of the public transport satisfaction. Then, TRAM surveys can be used to construct pseudo panel data and to estimate dynamic LISREL models of satisfaction by cohort average technique. Customers are grouped in 65 cohorts, defined with respect to their year of birth, excluding those born before 1920 and after 1984. Cohort representative values are the mean values of variables calculated on customers belonging to that cohort in each time period. Repeating for all the cohorts and linking them, satisfaction data over customer life-cycle are obtained. Then the procedure overcomes the lack of panel data and enables estimation of dynamic models.

The use of a synthetic measure, such as the mean, for ordinal variables is largely discussed in the literature: the most commonly solution consists in treating measures in an ordinal scale as they are on an interval scale. This hypothesis is justified in presence of 10 scores (Zanella, 2001).

5. Estimate results

5.1 Factor analysis

In order to find the latent structure underlying the TRAM customer satisfaction evaluation system and to specify the structural equation model, we conducted an exploratory factor analysis on the data for the 2004 year, using Principal Axis Factoring and Varimax Rotation. A 2-factor structure, with more than 65% of total variance explained, is the best solution (Tab. 2). For each of the two sets of items the Cronbach's Alpha was calculated in order to evaluate how well these items measure the latent constructs they refer to. The results, 0.810 and 0.891 respectively for the first and second factor, confirm that the two sets of items are adequate.

Tab. 2 – Latent factors description

Factor 1 (T.T.) Travel Time	*Factor 2 (T.C.) Travel Comfortableness*
Punctuality	Convenience
Frequency	Cleanliness
Speed	Comfort
	Safety
	Bus Shelters
	Tickets Availability
Cronbach's Alpha: 0.810	Cronbach's Alpha: 0.891

The confirmatory analysis also indicates that the measurement model defined for 2004 has a good fit as indicated by a chi-square of 106.38

with 63 degrees of freedom (χ^2/df=1.69) and Adjusted Goodness of Fit Index of 0.92. The same measurement model applied to data referred to the other years shows a good fit (Tab. 3).

Tab. 3 – Confirmatory analysis indices

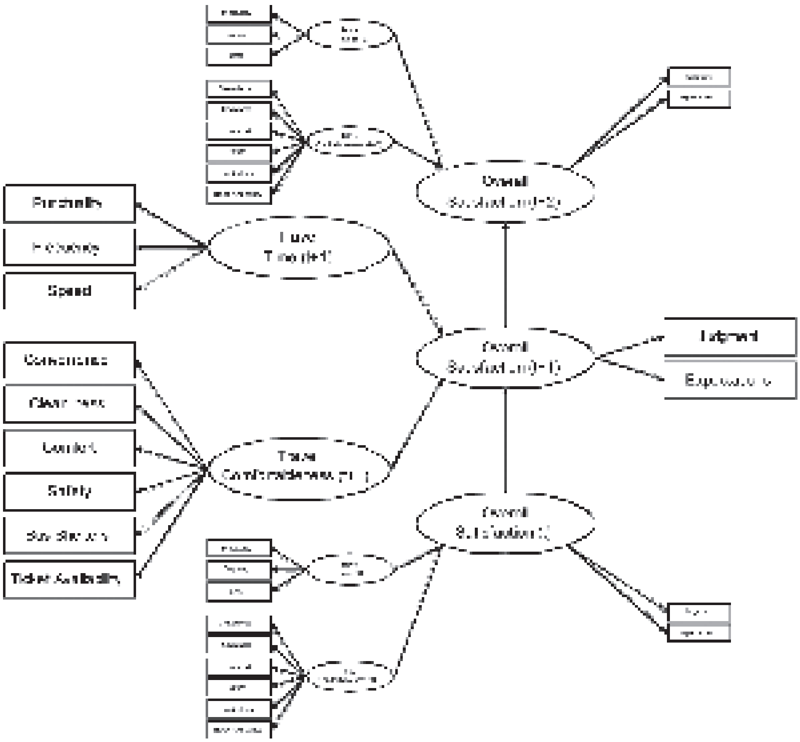
Year	χ^2/df	AGFI
2000	1.76	0.88
2002	1.74	0.89
2004	1.69	0.92
2005	1.73	0.87
2006	1.77	0.88

5.2 Model estimates

A structural equation model with the aim to verify the hypotheses previously defined is specified. The two latent exogenous variables, Travel Time and Travel Comfortableness, measured by the items reported in Tab. 2, affect the latent endogenous variable Overall Satisfaction, measured by the observed variables Judgment and Expectations. The ES model introduces dynamics in the structural model through time (causal) dependence between overall satisfaction latent variables (Fig. 4).

Because of the reduced number of available observations whit respect to the high number of parameters, we estimate a parsimonious model in which only the first three years of observation are considered. The specified model is estimated by Weighted Least Squares (WLS) method and almost all the parameters have positive sign with high t-values (Tab. 4). The model shows a good fit (CFI=0.98). The value of the RMR (0,27) is not satisfactory but it is consistent with a parsimonious model has indicated by the PNFI and PGFI (respectively 0.92 and 0,87) (Browne and Cudeck, 1993).

Fig. 4. Path model



Some of the conditions for corretly applying WLS are non satisfied by our data. We refer to normal distribution of the variables, large sample size, and absence of misspecification errors. Then, we propose to estimate the ESM model using 2SLS estimators.

The equations for the latent variable model are:

$$OS_{00} = \gamma_{10}TT_{00} + \gamma_{20}TC_{00} + \zeta_{00}$$

$$OS_{0t} = \beta_{tt-1}OS_{0t-1} + \gamma_{1t}TT_{0t} + \gamma_{2t}TC_{0t} + \zeta_{00} \qquad t=2, 4, 5, 6$$

Following Bollen's approach, we rewrite the ES model in the OFS substituting ($Jud_{0t} - \varepsilon_{0t}$) in for OS_{0t}, ($Punt_{0t} - \delta_{0t}$) in for TT_{0t}, ($Comf_{0t} - \delta_{0t}$)

in for TC_{0t}. Rearranging terms leads to:

$$Judg_{00} = \gamma_{10}Punct_{00} + \gamma_{20}Comf_{00} + u_{00}$$

$$Judg_{0t} = \beta_{tt-1}Judg_{0t-1} + \gamma_{1t}Punct_{0t} + \gamma_{2t}Comf_{0t} + u_{0t} \quad t=2, 4, 5, 6$$

where u_{00} and u_{0t} represent the composite disturbance terms. Following the rules for selection of IVs, the IV for $Judg_{0t}$ is Exp_{0t}; $Punct_{0t}$ is instrumented by $Freq_t$, $Freq_{t-1}$, $Speed_t$ and $Speed_{t-1}$. Eligible IVs for $Comf_{0t}$ are Con_t, Con_{t-1}, Saf_t, Saf_{t-1}, $Clean_t$, $Clean_{t-1}$, Bus_{t-1}, Bus_{t-1}, $Tick_t$ and $Tick_{t-1}$.

Tab. 4 Structural model coefficient estimates (standard errors in brackets)

	ESM (WLS)			ESM - OPS (2SLS)		
	T.T.	T.C.	O.S.	T.T	T.C.	O.S.
γ [illegible]	0.04 *(0.15)*	1.61 *(0.15)*		0.60 *(0.12)*	0.33 *(0.14)*	
γ [illegible]	0.59 *(0.04)*	0.35 *(0.05)*		0.43 *(0.08)*	0.46 *(0.11)*	
γ [illegible]	0.54 *(0.14)*	0.57 *(0.14)*		0.40 *(0.15)*	0.56 *(0.14)*	
γ [illegible]	-	-		0.16 *(0.07)*	0.57 *(0.07)*	
γ [illegible]	-	-		0.52 *(0.07)*	0.24 *(0.07)*	
β_{21}			0.05 *(0.03)*			0.24 *(0.71)*
β_{42}			0.21 *(0.03)*			0.16 *(0.58)*
β_{54}						0.25 *(0.58)*
β_{65}						0.26 *(0.10)*

In the analysis a robust version of 2SLS estimator to heteroskedasticity and small sample is used. The diagnostics for the IVs in each equation for the 2SLS estimator are good; the R2 of the regressions of Judg0t, Punct0t and Comf0t on their respective IVs are always high. The overidentification tests are also favourable: the tests of the joint null hypothesis that the excluded instruments are valid instruments (uncorrelated with the error term and correctly excluded from the estimated equation) are always not rejected. Then, test results support the validity of choice of IVs.

Tab. 4 reports the 2SLS estimates and standard errors for the latent variable model and contrasts them with the WLS estimates from LISREL. Estimates and standard errors are quite different. Differently from WLS, 2SLS estimates are always significant and sTab. over time. Standard errors from 2SLS estimators are at least equal or greater than the WLS ones, even if the difference is far from dramatic (similar results are found in Bollen *et al.*, 2007). Furthermore, 2SLS approach estimates separately each equation of the model and then it is not limited by the number of available observations, as WLS approach. This properties is particularly relevant in dynamic analysis, when time points observations become large.

5.3 Economic results

Model estimates show some interesting results in order to verify research hypotheses and evaluate customers satisfaction dynamics.

Overall satisfaction significantly depends both on Travel Time and Travel Comfortableness. Coefficients with positive sign indicate that higher levels in Travel Time and in Travel Comfortableness satisfaction lead to higher level in Overall Satisfaction. Then Hypotheses H.1 is accepted.

The Hypotheses H.2 is also supported by the data: overall customer satisfaction depends on previous overall satisfaction and the role of Time and in Travel Comfortableness in affecting satisfaction changes over time. With the exception of 2006, the influence of Travel Comfortableness increases over time (it passes from 0.33 in 2000 to 0.55 in 2005), while Travel Time shows a decreasing influence over the period (the coefficient is 0.60 in 2000 and declines to 0.16 in 2005). Therefore, customer conditions on the bus and the easiness of the public transport use are the most affecting aspects of the customer satisfaction and their relevance increases over time. In addition, parameter estimates evidence that overall customer satisfaction is positively affected by satisfaction at the previous year, confirming the presence of dynamics in the customer satisfaction evaluations. Being the public transport choice mainly habitual, customer evaluations are largely determined by (past) evaluations.

Finally, model estimates on cohort data give significant results, supporting the assumption that public transport evaluations change over customer age (Hypothesis 3).

6. Conclusions and future researches

This study develops a dynamic customer satisfaction model for the urban transport services with the aim to investigate whether customer satisfaction changes over time and what are the main determinants. The conceptual model used in the analysis relates the Overall customer Satisfaction to customer evaluations toward Travel Time and Travel Comfortableness, investigating how these relations change over time.

In absence of panel data on customer evaluation we propose to use the average cohort technique in generating pseudo panel data. Satisfaction changes over customer age and the average cohort technique can be usefully used to explore the dynamics in the quality evaluation of the public transport. Then, pseudo panel data overcomes the typical problem of the lack of longitudinal observations on customers allowing dynamic analysis by LISREL Model.

In the paper, we propose to estimate the Extended Simplex Model using instrumental variable approach (2SLS estimator). As suggested by Bollen (1996, 2001), instrumental estimators may be superior to full information estimators when normality and misspecification assumptions are not longer satisfied, as in our case. Compared to WLS, 2SLS estimates appear more stable over time and evidence appreciable statistical properties in the analysis of dynamic satisfaction models.

Structural model estimates show that satisfaction changes over time, evidencing the presence of temporal dependence in the evaluation scheme used by customers. The influence of some characteristics of the transport service, such time and comfortableness of the travel, on customer satisfaction is also detected. Finally, the analysis supports the hypothesis that customer characteristics, and in particularly customer's age, affect the evaluation scheme.

ESM-OFS model on pseudo panel data is an interesting framework to analyze customer satisfaction. Model estimates from SIMPLEX and ESM are however derived under the weak assumption of time varying coefficients. Further researches will be conducted to overcome this hypothesis, investigating dynamic model in the form of autoregressive distributed lag structural equation models (ADL) with latent variables.

References

Bagozzi R.P., Warshaw P.R. (1990), *Trying to consume*. Journal of Consumer Research, 17, 127-140.

Bagozzi R.P. (1994), *Structural Equation Models in Marketing research: basic principles*. In Bagozzi R.P., *Principles of marketing Research*, Blackwell, Cambridge, 317-385.

Bernini C., Lubisco A. (2005a), *Customer Satisfaction dinamica e indagini sezionali ripetute nel tempo: un'applicazione ai servizi di trasporto urbano*, Atti della Riunione Scientifica "Valutazione e Customer Satisfaction per la qualità dei servizi", Rome (8-9/09).

Bernini C., Lubisco A. (2005b), *A new extension of dynamic Simplex Model for the public transport Customer Satisfaction*, Statistica, 4, 355-374.

Bernini C., Lubisco A. (2006), *L'analisi dinamica della Customer Satisfaction: il caso dei servizi di trasporto urbano*, Statistica & Applicazioni, III, 29-36.

Bernini C., Lubisco A. (2007), *The evaluation of age effects on the dynamics of public transport satisfaction*, Atti della Riunione Scientifica "Valutazione e Customer Satisfaction per la qualità dei servizi", Rom3 (12-13/04)

Bollen K.A. (1996), *An alternative Two Stage Least Squares (2SLS) estimator for latent variables equations*. Psychometrica, 61(1), 109-121.

Bollen K.A. (2001), *Two Stage Least Squares and latent variables models: simultaneous estimation and robustness to misspecifactions.* In R.Cudeck, S. Du Toit and D. Sorböm (eds), *Structural equations modelling: present and future*, a festschrift in honor of Karl Jöreskog. Lincoln, Scientific

Software, 119-138.

Bollen K.A., J. Kirby, P. Curran, P. Paxton, F. Chen (2007), *Latent Variable Models Under Misspecification: Two-Stage Least Squares (2SLS) and Maximum Likelihood (ML) Estimators*, Sociological Methods Research, 36, 48-86.

Browing M., Deaton A., Irish M. (1985), *A profiTab. approach to labor supply and commodity demands over the life-cycle*, Econometria, 53, 503-544.

Browne M., W Cudeck R. (1993), *Altenative ways of assigning model fit*, in Bollen K.A., Scott Long J. Testing, Structural equation models, Sage Publications, Newbury Park, 136-162.

Cagnone S., Lubisco A., Mignani S. (2003), *Modelli a variabili latenti per l'analisi della Customer Satisfation degli utenti dei servizi di trasporto pubblico*, Collana Appunti e ricerche, 2, Alma Mater Studiorum Università di Bologna Polo scientifico didattico di Rimini.

Deaton A. (1985), *Panel data from time series of cross-section*, Journal of Econometrics, 30, 109-126.

Fujii S., Gärling T., Kitamura R. (2001), *Changes in divers' perceptions and use of public transport during a freeway closure: effects of temporary structural changes on cooperation in a real life social dilemma*, Environment and behaviour, 33(6), 796-808.

Gärling T., Boe O., Fujii S. (2001), *Empirical tests of a model of determinants of script based driving choice*, Transportation research F, 4, 89-102.

Golob T.F. (2003), *Structural equation modelling for travel research*, Transportation research Pat B, Methodological, 37, 1-25.

Jakobsson C., Fujii S., Gärling T. (2002), *Effects of economic disincentives on private car use*, Transportation, 29(4), 349-370.

Jöreskog K.G. (1970), *Estimation and testing of simplex models*. British Journal of Mathematical and Statistical Psychology, 23, 121-145.

Jöreskog K.G. (1979), *Statistical estimation of Structural Model in longitudinal-developmental investigation*. In Longitudinal research in the study of behaviour and development (Edited by J. R. Nesselroade), 303-352.

Jöreskog K. and Sörbom D., (1988), *Structural equation modeling with ordinal variables using*, Scientific Software international.

Jöreskog K.G. (2001). *Analysis of Ordinal Variables: Longitudinal Data*.

In Scientific Software International.

Jöreskog K.G. Rock D.A. Werts C.E. and Linn R.L. (1977), *A maximum likelihood solution to the errors in variables and errors in equation models*, Journal of Multivariate Behavioral Research, 12, 187-197.

Jöreskog K.G., Sörbom D. (1977), *Statistical models and methods for analysis of longitudinal data*, in D.J. Aigner and A.S. Goldberger (Eds.): Latent Variables in Socioeconomic Models. Amsterdam: North-Holland Publishing Co.

Peter J.P., Olson J.C., Grunert K.G (1999), *Consumer behaviour and marketing strategies.* McGraw Hill, London.

Thögersen J. (2001), *Structural and psychological determinants of the use of public transport*. Presented at the TRIP colloquium at Horsholm (14-15/11).

Zanella A. (2001),*Valutazione e modelli interpretative di Customer satisfaction: ricerca di un quadro complessivo.* In *Processi e Metodi statistici di Valutazione*, Atti del Convegno Intermedio SIS, 113-120.

G. & L. E. R., Vol. 13 No. 2, 2009

Pietro Rostirolla[1], Oriana Romano[2]

TAXI SERVICE AND SUSTAINABLE MOBILITY: A PEOPLE SATISFACTION SURVEY

Abstract

This paper presents a new management and pricing model for taxi services, in order to provide an important contribution to sustainable urban mobility.

The main objective is to move the mobility demand from private to public transport for individual use, in order to obtain economical and environmental benefits. The evaluation of the project is done through statistical analysis, consisting of a people satisfaction survey and economic-financial analysis, based on the calculation of benefits and costs of several subjects involved in the project. A people satisfaction survey, specifically, allows to define intervention modality and to provide parameters for the simulation model; moreover, the survey allows to test ex-ante the feasibility of the project, involving a large number of stakeholders: users, not users, taxi drivers, service operators and experts.

JEL classification: L91

Keywords: Taxi service; People satisfaction; Feasibility analysis

1. Taxi service as a sustainable mobility tool

In order to achieve a sustainable transport system[3], local authorities can implement several actions that, according to a reasonable and also

[1] Department of Social Science University of Naples "L'Orientale" rostirol@unior.it

[2] Department of Social Science University of Naples "L'Orientale" oromano@unior.it

[3] Sustainability applied to the transport sector has to implement actions that will reduce the damage caused to the ecosystem, but also consider other aspects such as the social one (to ensure fairness, accessibility and safety) and the economic one (to limit the costs arising from the negative effects produced by the same transport as much as possible).

necessary activity of transportation of people and goods and with an evaluation of financial constraints and decision-maker priorities, affect directly or indirectly users' behaviour. In some cases these actions present high costs, with a strong impact on public budget.

Local governments, with substantial financial and human resources, are looking for appropriate solutions finalized to minimize the role and the privilege of private vehicles, suggesting measures distinguished in "hard" and "soft" ones.

"Hard" measures involve huge investment costs, because they require infrastructural work; whereas examples of "soft" measures are the following: rationalization of private vehicle flows, creation of new bus lanes, traffic limitation policy; charging policy for using public areas (park and road pricing) and distribution of tickets "park & ride". With regards to sustainable mobility policies defined by the government, however, cost is not the only problem. It is necessary to consider the implementation time - sometimes too long for environmental emergencies – and the intervention effectiveness.

In recent years, in order to promote the spread of a sustainable mobility culture, it has been opted for new modes of transport, new technologies, new forms of organization of private journeys and new systems of fares.

However, starting with an analysis of objectives, sometimes conflicting with each other, it is reasonable to sustain that only a mix of actions can achieve sustainable mobility.

This work presents the improvements that could be provided in sustainable mobility by the implementation of a new management and pricing model for taxi services. This model overcomes the problem of prohibitive fares and intends to cater for an higher share of mobility demand, given that at the moment this is mostly dependent on private transportation.

According to the national law[4] taxi service belongs to the category of public transport. Compared with other means of transport, it is a high quality service, because it is faster and more flexible in time, availability and routes.

As The White Paper on Transport (European Commission, 2001)

[4] Law n. 21/1992.

suggests, taxis can play a significant role in achieving the goal of re-organizing the distribution modes of transport demand.

Taxis provide several advantages: support for population mobility; reduction in use of private vehicles in metropolitan areas and satisfaction of different needs and time demands. In relation to its great flexibility, taxi service plays a central role in urban mobility, representing a compromise between public and private means of transport.

In recent times liberalization policies[5] has been carried out in order to favourite competition for taxi service market and to solve the situation in which the demand is higher than the supply[6].

This situation, however, is extendible to several major Italian metropolitan areas, but not to all. For example, the city of Naples is an exception and it has been taken as a model for the simulation that follows, in which liberalization (favouring a greater number of licenses), will not represent an attractive perspective. This is because the supply of 2,370 taxis is higher than demand. Given that Naples is particularly plagued by congestion problems, traffic and high levels of emissions, targeted actions which discourage the use of private cars in favour of a more efficient distribution of mobility are desirable.

This paper is structured as follows: a brief introduction to present the project followed by a presentation of the results and the effects obtained is done (Section 2 and Section 3). An important contribution is provided by the use of surveys. In this case, considering the complexity of the project, a people satisfaction survey was used, with *ad hoc* interviews prepared for different stakeholders involved in the project proposal.

[5] Law n. 248/2006.

[6] However, the debate between regulation and liberalization, public and free market is still in progress, because in many countries, taxis are an integral part of public transport services. In fact, according to the experiences of several countries that have opted for one or another form of governance neither regulation, nor liberalization have led to fully satisfactory results, bringing out a picture of the taxi service as a case of failure of the State but also of the market. On the regulation on taxi service see also: Barrett S. D. (2003), "Regulatory Capture, Property Rights and Taxi Deregulation: a Case Study", in Economic Affairs, vol. 23, December, pp. 34-40; Beesley, M. E. (1973), "Regulation of Taxis", in The Economic Journal, vol. 83, n. 329, March, pp. 150-72; Beesley, M. E., S. Glaister (1983), "Information for Regulating: the Case of Taxis", in The Economic Journal, vol. 93, n. 371, September, pp. 594-615; Cairns R.D., C. Liston-Heyes (1996), "Competition and Regulation in the Taxi Industry", in Journal of Public Economics, vol. 59, pp. 1-15; Daly J. (2004), "Taxi Deregulation: Three Years on", in Student Economic Review, vol. 18, pp. 225-233; Douglas G.W. (1972), "Price Regulation and Optimal Service Standards: The Taxicab Industry", Journal of Transport Economics and Policy, vol. 6, n. 2 pp. 116-127; OCSE (2007), (De)regulation of the Taxi Industry, Paris. Stigler, G.J. (1971).

Then the description of people satisfaction survey applied to this case study is illustrated (Section 4).

2. The proposal

The Taxi Service Management model, proposed here, wants to show taxi service potential and demonstrate that this kind of intervention could solve not only traffic problems, but it could also improve other issues of the local government, such as environmental impact and parking.

Tariff modification is the basic starting point for the project. Taxis would be more used if prices of individual journeys diminished. Otherwise people would not have any reason to choose taxis instead of private vehicles.

Mainly the project suggests the purchase of tickets[7] entitling to a discount on individual journeys: the higher is the cost of the card, the greater the discount.

Three types of cards are considered:

- Card "A" is suiTab. for the start up phase. Indeed, it is characterized by a promotional low cost, although the total price is determined in addition to the subscription to public transport, so that this allows to use taxis together with other public means of transport;
- Card "B" is destined to users who mainly get taxis for work or business. The price of this card is higher than card "A", giving that subscribers will have an elevated number of journeys per month;
- Card "C" presents a higher cost and a more effective reduction in relation to a planned increased use of taxi services. All family members can use it.

Different users have been considered:

- frequent users who already use taxis and would like to increase the journeys thanks to the card;
- new users who actually use the car, but thanks to the card could use

[7] Cards can be compared to those that are usually purchased to become a "club member", therefore the card holder, according to the card validity, has the right to receive discount on a single event and in a certain period of time. Discount cards issued by taxi operators in Europe are different from the kind of card mentioned here as they are directed exclusively to specific categories of users: people with disabilities or mobility problems.

taxis for short journeys in the city centre and private vehicles for longer journeys;
- new users who decide to use exclusively taxis instead of private vehicles. This is the most cost -effective hypothesis because it reduces car fixed costs (fee service, insurance, parking).

To verify the suitability of the cards, a number of minimum rides[8] per category are calculated:

Tab. 1 - Value of Cards

	Card A	Card B	Card C
Cost for users (at month)	€ [illegible]	€ [illegible]	€ [illegible]
Value for taxi services purchased (at month)	€ [illegible]	€ [illegible]	€ [illegible]
% of discount	40%	50%	50%
Num. of equilibrium rides per card			
regular users	4	7	8
- new users that use cars as well	[illegible]	[illegible]	[illegible]
- new users that don't use car anymore	Always cost-effective	Always cost-effective	Always cost-effective
from Card A to C[9]			28

Thanks to cost saving strategies, it is possible to maintain and increase the use of taxis for habitual users, but also to attract new users.

It is important to highlight that the proposal does not provide investment costs for local authorities, but merely an indication of promotion costs. Moreover risks related to incorrect estimates of demand are limited, given that it is reasonable to assume that the supply can adapt quickly to changes in users'demand; so it should be easy to achieve and maintain a supply/demand balance.

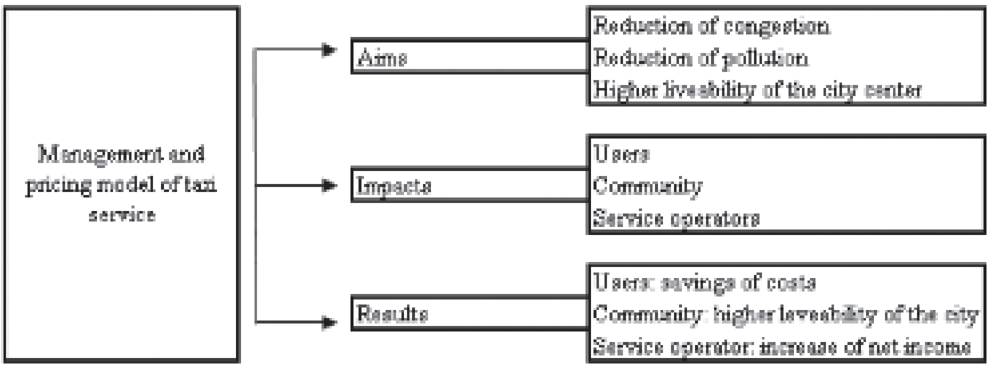
The starting point could be the testing of the project by a taxi company, in order to test the project's benefits.

To sum up, anticipating the conclusions, the project implementation includes:
- substantial savings for users, thanks to the introduction of discounted fares;
- a reduction of waiting time for the drivers;

[8] Below a certain number of journeys purchasing of cards is not cost-effective.

- an increase in earnings for taxi-drivers;
- the possibility of new jobs to face the growing demand;
- a reduced need of public space for parking;
- a decrease in traffic and pollution and a better urban environment;
- less pollution thanks to the public administration's participation in co-financing the purchase of vehicles with low emissions.

Fig. 1:Management and pricing model of taxi service



3. The management and tariff model

The model formulation is determined through several stages:

1. simulation considers initially a draft of 600 taxis, a kind of company created *ad hoc* for the project and thanks to the cost saving due to the purchase of the cards, it considers an increase of journeys as a consequence of the greater use of taxis;
2. starting from the number of taxis in service in Naples, the main parameters of the mathematical model are defined, such as: working days in a month, the average number of daily races, the average number of monthly courses split between regular and occasional users, the average cost of races and supplements. After this, fixed costs of taxis are set out;
3. monthly costs of the three types of discount-cards are set out. A monetary evaluation of service costs in relation to the taxi fleet and a

monetary evaluation of benefit-costs[9] per users, public operator and community are defined;
4. outcomes per pricing scenario are simulated.

In order to verify the actor's individual outcome in the project implementation it is necessary to proceed with the definition of a benefit-cost balance; it is also required to give flexibility to the project, taking into account changes and additions: for instance in case of economic and regulatory constraints. The fulfilment of these conditions can produce all those effects that determine social convenience in a market economy where the decision-making process is highly decentralized.

The project suggests three scenarios: the first one concerns a group of 600 taxis participating in the project, the second one concerns the participation of the whole current taxi fleet, consisting of 2,370 cars and finally the third one consists of smaller groups of 100 taxis. This differentiation is in place because the proposal can be gradually implemented and the number of participating taxis can be increased. Therefore it could be possible to implement the project with a starting module of 100 taxis. Tab. n. 2 shows the annual incremental values for each hypothesis.

Tab. 2: Results per scenario

Incremental imports	Basic draft	All taxis draft	Module draft
Taxi	600	2370	100
Users with card	9819	38.785	1636
A) Gross profit and service company	7.412.117	29.277.862	1.235.353
B) Gross profit taxi driver	12.354	12.354	12.354
Incremental time of drive in hours	441	247	514
Private vehicles Km	3.407.040	13.457.808	567.840
Private vehicles avoid Km	11.072.880	43.737.876	1.845.480
Taxi incremental Km	10.221.120	40.373.424	1.703.520
Vehicles avoid Km	851.760	3.364.452	41.960
Saving of fuel (litres)	70.980	280.371	11.830
Saving of fuel (euro)	99.372	392.519	16.562
Saving of time for users	10,992,155	45,098,797	1,788,856

9 Card A is cost –effective until a maximum of 29 rides at month. After that is more convenient to buy Card C.

Total benefit for users	11.081.327	46.491.247	1.505.392
Avoided costs for parking made by private vehicles	6.8[illegible]	26.915.616	1.[illegible]
Avoided spaces for parking	32.760	128.422	5.460
Avoided costs for public parking	54.600.000	215.670.000	9.100.000
Amortization parking	2.730.000	10.783.500	455.000
Benefits for local authority	2.730.000	10.783.500	455.000
C) Total benefits	21.293.644	86.357.608	3.403.745
CO2 costs company	55.373	36.520	54.937
C/A	2,9	3,0	2,4
Saving of time for other vehicles	22.496.786	149.903.068	622.045
Saving of fuel for other vehicles	3.924.565	61.233.026	109.016
D) Total benefits for non users	26.350.651	411.136.094	731.062
E) Total social benefits (C+D)	47.573.295	497.608.613	4.225.205
F) Present value of E (five years)	225.971.801	2.154.751.369	18.295.139
Financing ecologic vehicles	5.760.000	22.752.000	960.000
Financing/F in %	3%	1%	5%
G) Financed ecologic vehicles at 40%	480	1.896	80
Co-financing local government (40% benefits)	3.600.000	14.220.000	600.000
Cost local government/benefits for local government %	30%	30%	25%
Ecologic vehicles producers	2.160.000	8.532.000	360.000
Discount	15%	5%	15%

Benefits arising from this tariff model do not only affect users and non-users in terms of lower costs and better quality of life in urban centres, but also concern taxi drivers and associations. This is easily understood from the basic assumptions of the project, aimed at increasing the journeys and reducing waiting times. Evaluation of the project is done by comparing the alternative situations “without” the project and “with” the project so as to highlight the differences.

Regarding service delivery, the largest daily mileage due to the increasing number of journeys represents a raise of revenues: in actual fact the result is a sharp growth in net income for the reason that the fares, although discounted, are nevertheless higher than the marginal costs of production, represented almost exclusively by the consumption fuel.

Tab. 3: Results per taxi-service company

	Without card			With card			Incremental value		
	Day	Month	Year	Day	Month	Year	Day	Month	Year
N. of rides	13	338	4056	26	676	8112	13	338	4056
Gross revenue	91,0	2.366,0	28.392,0	182,0	4.732,0	56.784,0	91,0	2.366,0	28.392,0
Km made by taxi	32,5	845,0	10.140,0	65,0	1.690,0	20.280,0	32,5	845,0	10.140,0
Savings of fuel (euro)	4,4	113,4	1.360,5	8,7	226,7	2.720,8	4,4	113,4	1.360,5
Total time (hours)	1,7	44,2	530,0	3,1	80,9	970,7	1,4	36,7	440,7
Gross operating revenue	86,6	2.252,6	27.031,6	173,3	4.505,5	54.065,1	86,6	2.252,6	27.031,6
Discount on revenue	-	-	-	61,2	1.592,4	19.109,2	61,2	1.592,4	19.109,2
Net revenue (with discount)	86,6	2.252,6	27.031,6	112,0	2.913,8	34.955,9	25,4	660,2	7.922,3
Revenue from cards				14,7	383,4	4.600,2	14,7	383,4	4.600,2
Total revenue	86,6	2.252,6	27.031,6	126,8	3.296,2	39.554,1	40,1	1.043,6	12.522,5
Fixed costs									
- road tax	0,4	11,7	140,0	0,4	11,7	140,0	-		-
- insurance	7,7	200,0	2.400,0	7,7	200,0	2.400,0	-	-	-
taxi company fee	2,3	60,0	720,0	2,3	60,0	720,0			
- ordinary maintenance	0,5	14,1	169,0	1,1	28,2	338,0	0,5	14,1	169,0
- expenditure for reparation in case of accidents	1,3	33,3	400,0	1,3	33,3	400,0	-	-	-
garage	5,8	150,0	1.800,0	5,8	150,0	1.800,0			
- amortisation	6,4	166,7	2.000,0	6,4	166,7	2.000,0	-	-	-
Total fixed costs	24,5	635,8	7.629,0	25,0	649,8	7.798,0	0,5	14,1	169,0
Gross profits	62,2	1.616,9	19.402,6	101,8	2.646,3	31.756,1	39,6	1.029,5	12.353,5
Tax	12,4	323,4	3.880,5	20,4	529,3	6.351,2	7,9	205,9	2.470,7
Net profit	49,8	1.293,5	15.522,0	81,4	2.117,1	25.404,9	31,7	823,6	9.882,8

In terms of opportunity costs, the project sets out benefits for the local administration. The use of taxis, in fact, provides for the disposal of cars: fewer cars mean less demand of parking and therefore a saving by the administration in building car parks and using public land. Thanks to this savings local authorities can finance the purchase of eco- friendly taxis, in order to reduce pollution[10].

[10] About costs-benefits analysis see Rostirolla (1998).

The implementation of a new price model for taxi services could have a positive effect on employment: a greater demand for taxis, would imply an increasing demand for licences and thus new jobs.

4. People satisfaction survey

A people satisfaction survey is the link between the project idea and its implementation in practice; it is a useful tool to minimize the uncertainties of the event, through the identification of an average type of customer (which could represent specific targets) in order to identify specific measures of supply. Moreover it allows the calculation of benefits-costs of those directly involved in order to assess the feasibility of the project. Nevertheless it is a starting point to correct possible wrong parameters defined in the economic model. A People satisfaction survey goes beyond the concept of customer satisfaction[11], involving a larger number of subjects to obtain more points of view concerning the phenomenon under observation.

Generally, customer satisfaction in public transport is applied in order to measure service quality attributes (Eboli and Mazzulla, 2007) or increase customer expectation (Puvadolpitak, 2003). In this light Brog and Kahn (2003, p.1) affirm: "*the results* of the surveys are an important source of information for quality assurance and for acquiring new customers […]. Customer satisfaction surveys are an increasingly popular instrument for identifying the expectations of both existing and potential customers".

The general aim of the analysis, in this work, is to better understand strengths and weaknesses of the taxi service, nevertheless the survey allows to quantify taxi driver's workload, characteristics of service demand, waiting time and assessments on the current normative rules and

[10] Given the high cost, eco-friendly cars such as electric or hybrid, despite the market, are still not really taken into consideration by consumers. The project, however, requires that taxi drivers are encouraged to purchase eco-friendly cars, thanks to the contribution in economic terms, that both the local administration and the manufacturers of these cars would be willing to provide. In fact by reducing parking place demand, for example, the local government could re-invests that saving in sponsoring the purchase of environmentally friendly taxis. On the other hand, manufactures could be interested in sponsoring this purchase, experiencing an ineviTab. increase in demand and therefore profits.
The use of environmentally friendly vehicles could contribute significantly in the improvement of air quality and reduction of noise levels.

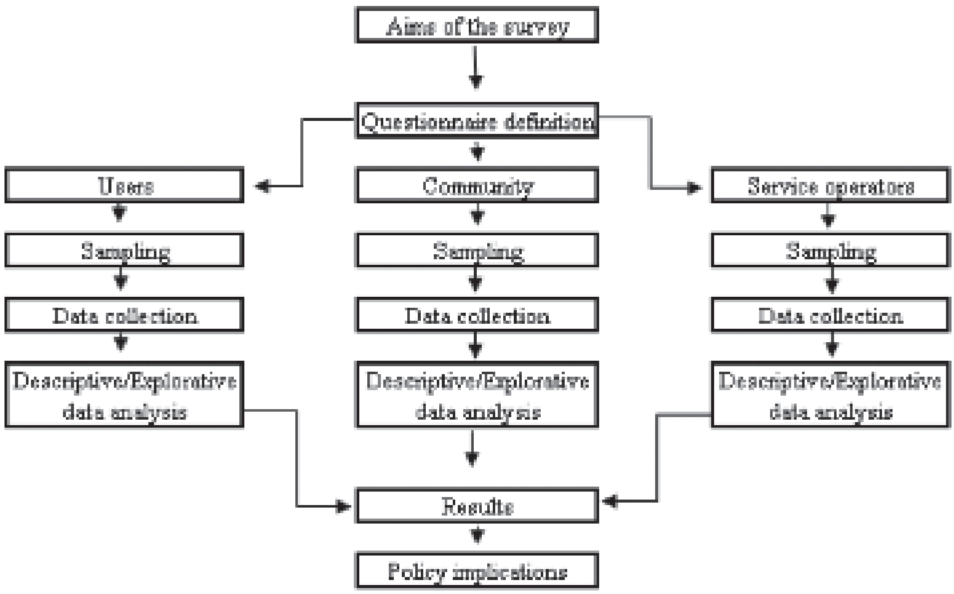
changing proposals. The survey wants to estimate the potential demand and the degree of acceptance of the proposal. An important role in the project is played by the definition of *ad hoc* questionnaires distributed to those involved and interested in tariff and regulatory changes, like users, not users, public operators, taxi drivers and citizens.

Users have been interviewed at the end of the taxi journey, while non-users have been interviewed in the main car parks.

Then the answers of the questionnaires were subjected to multidimensional analysis.

The survey was conducted through different steps:

Fig. 2: People satisfaction survey



4.1 Taxi drivers survey

With the aim at investigating the views of those directly involved in the project ad hoc surveys were carried out. A sample of 50 taxi drivers have been involved. The questionnaires distributed, consisting of 9 questions, basically want to show the weaknesses of the taxi service, test the importance of some measures that the administration might decide to implement, in order to encourage the use of the taxi service, and check the level of acceptance of discount –cards on single fares.

The questionnaire shows the characteristics of the taxi service in Naples. Waiting times are quite long: 58% of taxi drivers said that waiting time between two journeys goes from a minimum of one hour to a maximum of 90 minutes. Taxi drivers work hard in two time slots: up until 8:30 in the morning and in the evening from 6:30 to 8:30. This means that during the rest of the day taxis are not used at full and therefore they are used effectively for four hours a day.

According to the findings from the survey, the average number of journeys per day are: 5 short journeys (from € 4.50), 4 medium ones (from € 4.50 to € 7, 00), 2 medium-long journeys (from € 7.00 to € 10.00), no journeys or perhaps one long journey only (more than € 10.00).

Taxi drivers say, in 41% of cases, that the Administration should take measures to control fast lanes: respecting bus lanes allows taxis to increase speed and reduce the time taken especially for shorter journeys; the 26% and 25% respectively believe that it is necessary to control illegal parking and parking areas. The proposal to revise taxi shifts, to set fixed rates and to increase the parking charges is taken into account by small percentages. Nobody wants any measures concerning the increase in licensing and the reduction of taxi fares, confirming the strong opposition to any form of liberalization of the service.

The 56% of drivers would welcome the introduction of discount cards, even if this entails keeping track of the journeys made.

4.2 User Survey

In the questionnaire of taxi service users there are 13 multiple choice questions designed to capture information related to patterns of service users and to detect the negative factors in the service and to assess a positive attitude towards the project.

The questions basically cover the following items:

- satisfaction toward service provided and individuation of its critical

aspects;
- percentage of taxi use, compared to other means of transport;
- reasons for travelling;
- interviewee's personal information: employment, household composition;
- will to use the discount card.

In total 400 people were interviewed. The results show positive feedback on the great potential the taxi service has as a strategic means of transport for mobility in the city centre, but due to some limitations, especially economic ones, it is still not fully exploited by certain categories of people and therefore it continues to be considered as an occasional means of transport to use.

The 50.5% of users are satisfied with the service. This data can be analysed further if the degree of satisfaction and the frequency of use are taken into account: the connection between a joint analysis of these two factors shows that those who use taxis as an exception or up to 5 times in a month, believe to be dissatisfied with the service. Satisfaction increases when the number of journeys grows: in fact the people interviewed who used taxis from 6 to 15 trips in a month or more give positive feedback.

According to users, the negative aspects of the taxi service is the price (32%), followed by lack of speed (23%), availability (17%) and comfort (17%)[12]. Most people interviewed consider taxis as being an expensive type of transport: fares are not competitive compared to other public means of transport and with private cars. Mainly lack of speed is due to traffic. Availability of the taxi service is due to the current shift system, since it is difficult to catch a taxi on certain hours of the day (mainly at night). These factors appear to be crucial, but not impossible to handle: a different pricing system could solve the problem of high costs; reduction of cars around and different regulatory shifts could help with the problem of lack of speed and availability.

Regarding the reasons that encourage people to use taxis, 43% of people answered that they primarily travel for work.

The survey has been further developed through the distribution of the

[12] Customer satisfaction is defined as the consequence of comparing expectation with perceptions of performance (Oliver, 1981).

questionnaire to a panel of users previously informed of the project.

Regarding the question about discount, informed users would be much more inclined to use the cards, compared to non-informed users.

These results lead to some considerations about the importance of communication and participation of people on projects concerning the majority of people: well-spread information about certain initiatives, is a key factor in guiding preferences and generating consensus.

4.3 Non-user Survey

Questionnaires seek to obtain information on: private driver routes, reasons for travelling; frequency of car use in comparison with other means of transport; reasons for which other means of transport are not favoured by private drivers. Moreover, throughout the survey, it is possible to notice positive feedback on the use discount cards.

49.3% of people interviewed travels for work. 53% of the journeys are not made by taxis. Reasons for non-using them are mainly price (47%), traffic (22,8%), the car being considered more comforTab. (19,4%) and non-availability of taxis in certain hours and places (10,9%). If we analyze the connection between frequency of taxi use and the working categories considered, it is possible to notice that most of the people using taxis are self-employed and entrepreneurs, while students, housewives, employees and unemployed people do not use them.

44.3% gets cars for more than fifteen times a month, while the percentage of those who use private means of transport "rarely" or "only at weekends" is exactly the same: 28%. Among motorists 38% stated that they often get public means of transport. 27%, however, rarely use public transport: 43% say they do not choose public transport mainly because it is often necessary to get more than one means of transport to reach the destination, while 35% believe that cars are more convenient. 41% believe that cars still have lower costs than taxi fares; 34.3% believe the opposite, while 22% believe they are substantially equal. Regarding the discount initiatives, motorists provide answers that do not differ from each other in an evident way: 35% answered positively, 33% negatively and 32% were in doubt.

The professional group that answers "yes" to the proposal of using a card-off, is that of self-employed people and pensioners. Students, workers, housewives and unemployed people respond essentially “no”.

As in the case of taxi users, and for non-users it was decided to hand out the questionnaire to a group of individuals previously informed of the project. In this case it is possible to compare results and test the importance of dissemination of information, particularly, for public utility projects.

For both groups of drivers, discount-cards would be a good incentive to increase taxi journeys; however the percentage of informed motorists that expresses a positive opinion is 60%, while the percentage those not informed is 35%.

5. Conclusions

The particular attention paid to the taxi service in recent years and the regulatory changes that have affected the service, have stressed its importance, like a sustainable mobility tool in towns.

Starting with an analysis of objectives, it seems that the proposal, consisting in the implementation of a new management and pricing model for taxi services, can achieve sustainable mobility, respecting local government’s financial constrains

The project presented in this work wanted to show how a new model of tariff can benefits to various stakeholders directly or indirectly involved. Evaluation of the project is done by comparing the alternative situations “without” the project and “with” the project so as to highlight the differences.

Economic analysis has defined the benefits and costs from users, not users and service operators and it has been accompanied by a people satisfaction survey. This survey is different from a customer satisfaction survey, because it involve a larger number of subjects in order to obtain more points of view concerning the phenomenon under observation. The people satisfaction survey has been made not only to test the initiative feasibility, but also to provide a support to easily understand which parameters can be changed to ensure the initiative is to succeed.

References

Brog W., Kahn T. (2003), *Customer Satisfaction Surveys for Public Transport Companies - Greater Efficiency through more Demand-Orientated Methods* –Workshop 1d: Public transport: How can public transport meet the increased mobility demand, Institute for Traffic and Infrastructure Research *www.epomm.org/ecomm2003/ecomm/papers/WernerBrog.pdf*

Eboli L., Mazzulla G. (2007), *Service Quality Attributes Affecting Customer Satisfaction for Bus Transit* - Journal of Public Transportation, 10(3), 21-34, 2007.

European Commission (2001), *White paper – 'European transport policy for 2010: time to decide*, Office for Official Publications of the European Communities, Luxembourg.

Oliver R.L., (1981), *Measurement and Evaluation of Satisfaction Processes in Retail Settings*, *Journal of Retailing*, 57, 25-48.

Puvadolpitak K. (2003), *A study of customer satisfaction with express section of Linfox transport (Thailand) Ltd*., Research paper submitted in Partial Fulfillment of the Requirements for BC 7000 Indipendent study of Business Communication, The University of the Thai Chamber of Commerce, Summer Session http://library.utcc.ac.th/onlinethesis/onlinethesis/M0217673/M0217673.pdf

Rostirolla P. (1998), *La fattibilità economico-finanziaria – Metodi e applicazioni,* Liguori editore, Naples.

G. & L. E. R., Vol. 13 No. 2, 2009

Antonio Lucadamo*, Luigi D'Ambra, Francesca Pagliara

MULTINOMIAL LOGIT MODELS: MULTICOLLINEARITY AND ITS CONSEQUENCES ON PARAMETER ESTIMATION AND ON INDEPENDENCE FROM IRRELEVANT ALTERNATIVES TESTS

Abstract

The Multinomial Logit Model is one of the most used Discrete Choice Models and it has been widely used to study the transportation system. This model has a closed form and so it is computationally cheap, but it needs some restrictive assumptions. The Independence from Irrelevant Alternatives (IIA) property for example, is not always respected and the tests used to verify it can lead to opposed results almost when there is multicollinearity between explicative variables. Furthermore, the presence of multicollinearity can generate problems in the parameter estimation. In this paper we study the effects of the multicollinearity on the results of the IIA tests and we propose a new method to estimate the parameters of the multinomial logit model. The data used for the analysis concern the choice of the residence in the Zurich area.

JEL classification: C25; C51

Keywords: Discrete choice; Multinomial Logit; Principal Component Analysis; Multicollinearity

1. Introduction

Urban mobility is a vital component of any city, often influencing its physical shape as well as its level of economic and social development. However the governance of urban mobility is very complex since it is

[1] TEDASS - University of Sannio, Italy - alucadam@unina.it.

the result of the interactions of several elements, i.e. of the functioning of a socio-technical complex system (Cascetta *et al.*, 2008).

More specifically, a *transportation system* can be defined as the combination of elements and their interactions, which produce the demand for travel within a given area and the supply of transportation services to satisfy this demand. This definition is general and flexible enough to be applied to different contexts. The specific structure of the system is defined by the problem itself (or class of problems) for whose solution it is system employed.

Almost all of the components of a social and economic in a given geographical area interact with different levels of intensity. However, it is practically impossible to take into account every interacting element to solve a transportation engineering problem. The typical system engineering approach is to isolate those elements, which are most relevant to the problem. These elements, and the relationships among them, make up the *analysis system* (Cascetta, 2009). The remaining elements belong to the *external environment* and are taken into account only in terms of their interactions with the analysis system. In general, the analysis system includes the elements and the interactions that are expected to be significantly affected by the projects under consideration. It follows that there is a strict *interdependence between the identification of the analysis system and the problem to be solved*. The transportation system of a given area can also be seen as a sub-system of a wider territorial system with which it strongly interacts. The extent to which these interactions are included in the analysis system, or else in the external environment, depends on the specific problem (Manheim, 1979).

The *transportation system* can be split into two main components: demand and supply.

The distribution of households and activities in the area is the determinant of *transportation demand* deriving from the need to use different urban functions in different places. Household members are the users of the transportation supply system and make "mobility choices" (holding a driving license, number of cars, etc.) and "travel choices" (trip frequency, time, destination, mode, path, etc.) in order to undertake activities (work, study, shopping, etc.) in different locations. The result of these choices is the transportation demand; i.e. the number of trips made among the different zones of the city, for different purposes, in different periods, by means of the different available transportation modes. Similarly, economic activities transport goods that are consumed by the households or by other economic activities. Goods movements make up the freight transportation demand.

Both mobility and travel choices are influenced by some characteristics of the transportation services offered by the different travel modes (individual car, transit, walking). These characteristics are known as level-of-service or performance attributes and include travel times, monetary costs, service reliability, riding comfort, etc. Thus, the choice of destination may be influenced by the travel time and cost needed to reach each destination. The choice of departure time depends on the travel time to the destination. The choice of transportation mode is influenced by times, costs, reliability of the available modes.

The characteristics of transportation services depend on the *transportation supply*, i.e., the set of facilities (roads, parking spaces, railway lines, etc.), services (transit lines and timetables), regulations (road circulation and parking regulations), and prices (transit fares, parking prices, road tolls, etc.) producing travel opportunities. The physical elements of the transportation supply system have a finite capacity; i.e. a maximum number of users that can be served in a given time interval.

Individual trips can be aggregated into users flows, i.e. the number of users on the physical elements of the supply system in a given time interval. Examples are automobile and truck flows on road sections, passenger flows on transit lines, and so on.

When flow approaches capacity, the interactions among users increase and *congestion* effects are triggered. Congestion can significantly deteriorate the performances of transportation services for the users, e.g. travel times, service delays, fuel consumptions all increase with congestion. Congestion can also have other "external" negative effects (such as noise, air pollution and visual impacts in the case of road traffic). Congestion can have cross-modal effects; e.g. road congestion can influence the performances of surface transit services.

2. Modeling transportation systems

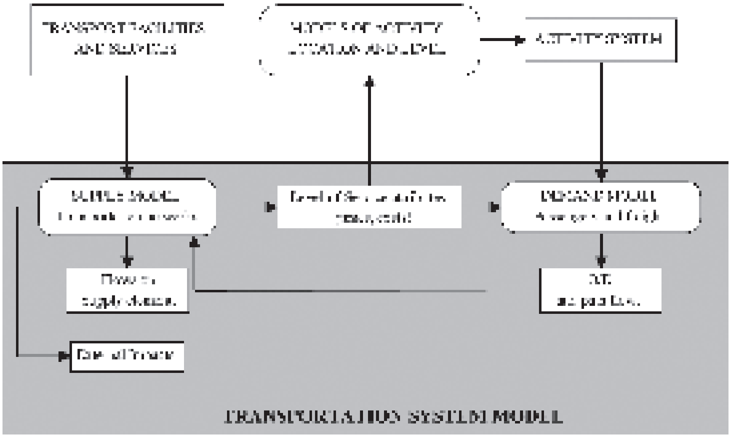
The relevant interactions among the various elements of a transportation system can be simulated with mathematical models; the models and their relationships are described in Fig. 1 (Cascetta, 2009).

Supply models simulate the transportation services available among the different zones with flow network models. More specifically, supply models simulate the performance of transportation infrastructures and services for the users, as well as the main external effects of transport (pollution, energy consumption, accidents). The level-of-service attributes, such as travel time and cost, will be input variables for the demand models. To simulate the performance of single elements (facilities) and the effects of congestion, especially for road systems, supply models use the results of traffic flow theory (Sheffi, 1985). For the transit system the hyper path approach is generally considered.

Demand models simulate the relevant aspects of travel demand as a function of the activity system and of the supply performances. Typically, the characteristics of travel demand simulated include the number of trips in the reference period (demand level) and their distribution among the different zones, the different transport modes, and the different paths. Other components of travel demand are simulated in specific applications such as the distribution between different time intervals within the reference period. Travel demand models are usually derived from random utility theory.

Assignment models (or network demand-supply interaction models) simulate how O-D demand and path flows load the various elements of the supply system. Assignment models allow the calculation of link flows, i.e. the number of users loading each link of the network representing the transportation supply in the reference period (Ortuzar and Willumsen, 2001). Furthermore, link flows may affect the transportation supply performances through congestion and therefore may affect the input to demand models. The mutual interdependencies of demand, flows and costs are simulated by assignment models.

Fig. 1 - Structure of transportation systems simulation models ***Source: Cascetta (2009).***



In this paper the focus is on demand models.

2.1 The random utility theory for modeling transportation users' behaviour

Random utility theory is based on the hypothesis that every individual is a *rational decision-maker*, maximizing utility relative to his/her choices. Specifically, the theory is based on the following assumptions:

a. the generic decision-maker i, in making a choice, considers m_i mutually exclusive alternatives that constitute his/her *choice set* C. The choice set may differ according to the decision-maker (for example, in the choice of transport mode, the choice set of an individual without a driver's license or car obviously does not include the alternative "car as a driver");
b. decision-maker i assigns to each alternative c in his/her choice set a perceived utility or "attractiveness" U_{ic} and selects the alternative that maximizes this utility;
c. the utility assigned to each choice alternative depends on a number of measurable characteristics, or *attributes*, of the alternative itself and of

the decision-maker: $U_{ic} = U_{ic}(\boldsymbol{X}_{ic})$, where $\boldsymbol{X}_{ic}$ is the vector of attributes relative to alternative *c* and to decision-maker *i*;

d. because of various factors that will be described later, the utility assigned by decision-maker *i* to alternative *c* is not known with certainty by an external observer (analyst) wishing to model the decision-maker's choice behavior, so U_{ic} must be represented in general by a random variable.

From the above assumptions, it is not usually possible to predict with certainty the alternative that the generic decision-maker will select. However, it is possible to express the probability that the decision-maker will select alternative *c* conditional on his/her choice set *C*; this is the probability that the perceived utility of alternative *c* is greater than that of all the other available alternatives:

$$p_i\,[\,c/C\,] = \boldsymbol{Pr}\,[\,U_{ic} > U_{ib} \;\; \forall\; b \neq c,\, b \in C\,] \tag{1}$$

The perceived utility U_{ic} can be expressed as the sum of two terms: a systematic utility and a random residual. The *systematic utility* V_{ic} represents the mean (expected value) utility perceived by all decision-makers having the same choice context (alternatives and attributes) as decision-maker *i*. The *random residual* ε_{ic} 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:

$$U_{ic} = V_{ic} + \xi_{\mathrm{ic}} \;\; \forall\; c \in C \tag{2}$$

The probability of choosing an alternative depends on the systematic utilities of all competing (available) alternatives, and on the joint probability law of the random residuals ε_{ic}. It follows:

$$p_i(c \mid C) = \boldsymbol{Pr}\,[V_{ic} + \varepsilon_{ic} > V_{ib} + \varepsilon_{ib} \quad \forall b \neq c,\, b \in C] \tag{3}$$

Systematic utility is expressed as a function $V_{ic}(X_{ic})$ of attributes X_{ic} relative to the alternatives and the decision-maker. Although in principle the function $V_{ic}(X_{ic})$ may be of any type, it is usually assumed for analytical and statistical convenience that the systematic utility V_{ic} is a linear function, with coefficients β, of the attributes X_{ic} or of functional transformations of them:

$$V_{ic}(\boldsymbol{X}_{ic}) = \beta^{\mathrm{T}} \boldsymbol{X}_{ic} \tag{4}$$

3. The Multinomial Logit Model

As shown in the previous section, the probability that any element *c* in C is chosen by the decision maker *i* can be expressed according to the equation (3).

Any particular multinomial choice model can be derived using the previous equation given specific assumptions on the joint distribution of the disturbances. It can be shown, as in Domencich and McFadden (1975), that if the error terms are:

- Independently distributed;
- Identically distributed;
- Gumbel distributed with a location parameter η and a scale parameter $\mu \geq 0$;

then the probability that alternative *c* will be chosen is

$$p_i(c \mid C) = \frac{\exp(\mu V_{ic})}{\sum_{b \in C} \exp(\mu V_{ib})} \tag{5}$$

If we consider that the utility is linear in parameters and we fix the scale parameter to 1, we can write:

$$p_i(c \mid C) = \frac{\exp(x_{ij}\beta_{jc})}{\sum_{b \in C} \exp(x_{ij}\beta_{jb})} \tag{6}$$

where C is the choice set of alternatives, x_{ij} are the values that the individual *i* assumes for the variable *j* and β are the parameters to be estimated (Ben-Akiva & Lerman, 1985).

The logit model has some special properties that under certain circumstances greatly simplify estimation of the parameters. Most of these theories can be addressed to McFadden (1974). The technique generally used to estimate a logit model is the Maximum Likelihood and for parameter estimation it is necessary to use some iterative algorithm. One of the most used is the Newton-Raphson algorithm that considers the first and second derivatives of the previous equation.

The first derivative is also called gradient vector and it is given by

$$g = \frac{\partial \log L(\beta)}{\partial \beta_c} = \sum_i^n x_i^{'} \left[y_{ic} - P_{ic} \right] \tag{7}$$

The second derivative, or Hessian matrix is:

$$H = \frac{\partial^2 \log L(\beta)}{\partial \beta_c \partial \beta_q^{'}} = -\sum_i^n P_{iq} \left[\delta_{cq} - P_{iq} \right] \boldsymbol{x}_i{'} \boldsymbol{x}_i \tag{8}$$

or in matrix form:

$$H = -X^{'} Z_{cq} X \tag{9}$$

Where Z_{cq} is a diagonal matrix with generic term $P_{iq}\left[\delta_{cq} - P_{ic}\right]$, $\delta_{cq} = 1$ if q=c and $\delta_{cq} = 0$ otherwise.

The parameter estimate according to Newton-Raphson is then:

$$\beta^{t+1} = \beta^{t} - (H)^{-1} g \tag{10}$$

It's important to consider the Hessian matrix, because its inversion could lead some problems in the estimation of the Multinomial Logit Model (D'Ambra, 2008).

4. Logit properties

One of the most discussed aspects of the multinomial logit is the

Independece from Irrelevant Alternatives property (IIA). This property states that for any two alternatives c and *d* the ratio of the logit probabilities

$$\frac{P_i(c)}{P_i(d)} = \frac{\exp(V_{ic})}{\exp(V_{id})} = \exp(V_{ic} - V_{id}) \tag{11}$$

does not depend on any alternatives other than c and *d.* Therefore the relative odds of choosing *c* over *d* are the same no matter what other alternatives are available or what the attributes of the other alternatives are. Since the ratio is independent from alternative other than *c* and *d,* it is said to be Independent from Irrelevant Alternatives. This assumption is realistic in some choice situation, but sometimes it can be clearly inappropriate. One of the classical example is the famous red-bus blue-bus problem, but there are many other situation in which the IIA assumption is not respected.

In this context we must search for a better model specification:

- Finding alternatives with missing or mis-specified variables;
- Point toward an acceptable nested logit structure.

To verify if the IIA property holds many tests exist. We can divide them in two groups:

- Estimate a model with a subset of the choice set. Reject IIA if the parameter estimates differ from the full choice set estimates.

  1) Hausman and McFadden (1984)
  2) McFadden, Tye and Train (1976)
  3) Small-Hsiao (1985)

- Implement a Lagrange multiplier test of IIA with the full set of alternatives.

  1) McFadden test (1987)

4.1 Estimating model with choice subsets

If we suppose IIA holds, then:

$$p_i(c \mid C) = \frac{\exp(x_{ij}\beta_{jc})}{\sum_{b\in C}\exp(x_{ij}\beta_{jb})} \tag{12}$$

and

$$p_i(c \mid \tilde{C} \subseteq C) = \frac{\exp(x_{ic}\beta_{jc})}{\sum_{b\in \tilde{C}}\exp(x_{ib}\beta_{jb})} \tag{13}$$

where $\tilde{C}$ is a subset of the full set of alternatives. It should provide similar estimates, since under IIA, exclusion of alternatives does not affect the consistency of estimators. As it was said, it is possible to use:

- Hausman-McFadden test (HM).

We have to build the following statistic:

$$\left(\hat{\beta}_{\tilde{C}} - \hat{\beta}_C\right)'\left(\Sigma_{\hat{\beta}_{\tilde{C}}} - \Sigma_{\hat{\beta}_C}\right)^{-1}\left(\hat{\beta}_{\tilde{C}} - \hat{\beta}_C\right) \tag{14}$$

That is asymptotically χ^2 distributed with $\tilde{C}$ degrees of freedom, where C is the number of elements in the subvector of coefficients that is identifiable from the restricted choice set model. So the null hypothesis that IIA holds is rejected if the value that comes from the equation 14 is bigger than the tabulated value of χ^2.

- McFadden, Tye and Train test (MTT).

In this case it's possible to build an approximate likelihood ratio test statistic with C degrees of freedom:

$-2\left[l_{\tilde{C}}\left(\hat{\beta}_C\right) - l_{\tilde{C}}\left(\hat{\beta}_{\tilde{C}}\right)\right]$ where the two log likelihood values are calculated on the estimation sample for the restricted choice set model. This statistic is not a proper likelihood ratio test because β_C is not a vector of constants. For this reason we can consider the following correction:

• Small and Hsiao (SH). To remove the bias they proposed to use:

$$\frac{1}{1-C_1(\alpha C)}\left\{-2\left[l_{\tilde{C}}\left(\hat{\beta}_C\right)-l_{\tilde{C}}\left(\hat{\beta}_{\tilde{C}}\right)\right]\right\} \tag{15}$$

Where C is the number of observations in the unrestricted choice set estimation, C_1 is the number of observations in the restricted choice set estimation ($C_1 < C$) since those observations with chosen alternatives not in the restricted choice set are omitted, and $\alpha \geq 1$ is a scalar. Asymptotically this corrected likelihood ratio statistic actually is χ^2 distributed with C degrees of freedom. Sometimes the assumption made by this correction that a scalar difference between the covariance matrices exists is not defensible, so it was proposed an exact test for the IIA assumption. To perform the test Small and Hsiao randomly divided the full estimation data set into two parts, denoted A and B. On sample A, using the restricted choice sets, they estimated $\hat{\beta}_C^A$, the subvector of coefficients corresponding to the parameters that are identifiable when the restricted set of alternatives are used; next, on sample B, using the restricted choice set, they estimated $\hat{\beta}_C^B$ and the corresponding log likelihood, $l_{\tilde{C}}^B\left(\hat{\beta}_C^B\right)$, finally again on sample B, but now based on the unrestricted choice sets, they obtained $\hat{\beta}_C^B$. They showed that if we form the following convex combination:

$$\hat{\beta}_{\tilde{C}}^{AB} = \left(1/\sqrt{2}\right)\hat{\beta}_C^A + \left(1\text{-}1/\sqrt{2}\right)\hat{\beta}_C^B \tag{16}$$

And use it to evaluate the log likelihood of the sample B with the restricted choice set, denoted as $l_C^B\left(\hat{\beta}_{\hat{AB}}^{B}\right)$, then the statistic $-2\left[l_{\tilde{C}}^B\left(\hat{\beta}_C^{AB}\right)-l_{\tilde{C}}^B\left(\hat{\beta}_C^B\right)\right]$ is asymptotically χ^2 distributed with $\tilde{C}$ degrees of freedom, $\tilde{C}$ being the common dimension of the $\hat{\beta}_C^A, \hat{\beta}_C^B, \hat{\beta}_{\tilde{C}}^B, \hat{\beta}_{\tilde{AB}}^B$ parameter vectors. This test is more computationally and time-consuming. Generally it's better if the simpler corrected approximate likelihood ratio test is carried out first, and then, only if its underlying assumption is violated, should the exact test procedure be used.

4.2 Lagrange multiplier test of IIA

This test checks if cross-alternative variables enter the model. If so, IIA assumption is violated. The test is composed of 3 steps:

- **Step 1**: We estimate the systematic utilities $\hat{V}_{ic}$ and fitted choice probabilities $\hat{p}_i(c \mid C)$ using all N observations: $\hat{V}_{ic} = \beta' x_{ic} \forall c \in C$;

$$\hat{P}_i(c \mid C) = \frac{\exp(x_{ij}\hat{\beta}_{jc})}{\sum_{b \in C} \exp(x_{ij}\hat{\beta}_{jb})}.$$

- **Step 2**: For a given $A \subset C$ we calculate auxiliary variables in the following way:

$$\hat{V}_{iA} = \frac{\sum_{b \in A} \hat{V}_{ib}\, \hat{p}_i(b \mid C)}{\sum_{b \in A} \hat{p}_i(b \mid C)} \qquad i = 1, \ldots, I \tag{17}$$

$$Z_{icA} = \begin{cases} V_{ic} - \hat{V}_{iA} & f \quad c \in A \\ 0 & otherwise \end{cases} \qquad i = 1, \ldots, I \tag{18}$$

Since Z is non zero only for the alternatives in the set A, it contains information regarding the other alternatives in A. The spirit of the proposed test is to verify the presence of cross alternative variables.

- **Step 3:** We can now estimate:

$$\hat{p}_i(c \mid C) = \frac{\exp\left(x_{ij}\hat{\beta}_{jc} + \upsilon^A Z_{icA}\right)}{\sum_{b \in C} \exp\left(x_{ij}\hat{\beta}_{jb} + \upsilon^A Z_{ibA}\right)} \tag{19}$$

The hypotheses are the following:

$$H_0 : \upsilon^A = 0$$
$$H_1 : \upsilon^A \neq 0$$

υ^A is distributed as a χ^2.

If we reject H_0 we reject the IIA assumption. If H_0 is not rejected nest A is considered to satisfy IIA.

5. Multicollinearity

Both in linear regression model and in logistic regression model a very desirable condition is that there is no multicollinearity among the regressors. The term multicollinearity is due to Ragnar Frisch (1934). Originally it meant the existence of a exact linear relationship among some or all explanatory variables of a regression model. For p-variables regression involving explanatory variables $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$ (where $\mathbf{x}_0 = 1$ for all observations to allow for the intercept term), an exact linear relationship is said to exist if the following condition is satisfied:

$$a_0\mathbf{x}_0 + a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_p\mathbf{x}_p = 0 \tag{20}$$

where $a_0, a_1, a_2, \dots, a_p$ are constants such that not all of them are zero simultaneously. The chances of one's obtaining a sample of values where the regressors are related in this fashion are indeed very small, in practise except when the number of observations is smaller than the number of regressors, so today the term multicollinearity is used in a broader sense to include the case of perfect multicollinearity, as shown by the previous formula as well as the case where the $\mathbf{x}$ variables are intercorrelated but not perfectly so, as follows:

$$a_0\mathbf{x}_0 + a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_p\mathbf{x}_p + \boldsymbol{v}_i = 0 \tag{21}$$

where v_i is a stochastic error term.

5.1 Multicollinearity and its links with test to identify the IIA property

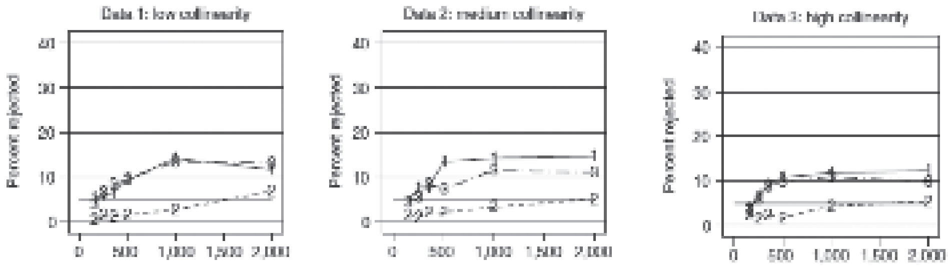
The presence of multicollinearity in the regressor matrix can lead to different results if we compute the tests to detect if the IIA property is respected or not. Cheng and Long (2007) showed in fact how the degree of collinearity, the choice of the category that is deleted in the restricted estimation and the sample sizes can change the properties of the tests. They run simulations for sample sizes of 150, 250, 350, 500, 1000 and 2000. The simulations involved the following steps:

- Draw a random sample size of N with replacement from the population.
- For this sample, estimate the Multinomial Logit Model for a data set with three predictors and the outcome variable with three categories.
- Using estimates from the previous step, compute three variations of the Hausman-McFadden test and of the Small-Hsiao test, excluding the first category for the restricted estimation, then the second and at least the third.

These steps were repeated 500 times for each sample size in each data set. To determine the empirical size for each test, they computed the percentage of times that each test rejected the null hypothesis that IIA held in the population at the 0,05 level of significance.

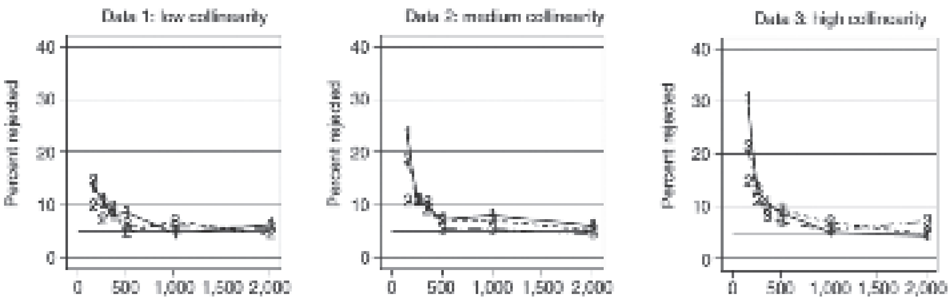
Fig. 2 shows the percentage of times the HM test rejected the null hypothesis of no violation of the IIA assumption using the 0,05 level. There are the graphs relative to three degrees of collinearity and with the elimination of one of the 3 categories for the outcome variables.

Fig. 2 - Size properties of the Hausman-McFadden test of the Independence of Irrelevant Alternatives



The results illustrate that the HM test does not reliably converge to its appropriate size even when the sample is 2000. Second, the properties of the test depend on which outcome category is deleted in the restricted estimation. In this case the degree of collinearity doesn't seem to have a particular influence on the results. But if we consider the Fig. 3 we can see that there are some differences.

Fig. 3 - Size properties of the Small-Hsiao test of the Independence of Irrelevant Alternatives



It is possible in fact to note that the SH test approximates its nominal size as the sample increases to 500 or 1000. The magnitude of departures from the nominal size and the sample size at which these distortions are largely removed depends on the degree of collinearity in the data. For example, with high collinearity, the size properties are quite poor with samples smaller than 500 and require a sample of at least 1000 before they are nearly eliminated.

The previous results show that tests of the IIA assumption that are based on the estimation of a restricted choice set are unsatisfactory for applied work. The Hausman-McFadden test shows substantial size distortion that is unaffected by sample size in the simulations. The Small-Hsiao test has reasonable size properties in some data sets but has severe size distortion even in large samples when there is collinearity in the regressors. For this reason it's important to be careful when a researcher wants to test the IIA assumption and it's important to consider not only the results of the tests, but the structure of the data too.

5.2 Multicollinearity and parameter estimation

As previously stated the presence of multicollinearity may indicate that some explanatory variables are linear combinations of the other ones. When this situation exists the logistic model estimations become unstable. To solve this problem in the binary case Marx (1992) introduced iteratively reweighted partial least squares algorithm, Bastien et al (2005) proposed partial least squares logit regression, Aguilera et al (2006) presented Principal Component Logistic Regression (PCLR), Vágó and Kemény (2006) developed the ridge logistic regression. A method to provide an accurate estimation of the model parameters in Multinomial Logit Model has instead been proposed by Camminatiello and Lucadamo (2008). They introduced in fact an extension of PCLR model, called Principal Component Multinomial Regression (PCMR). They proposed to use as covariates of the multinomial model a reduced number of Principal Components (PC) of the predictor variables.

PCMR creates at first step the PC's of the regressors as linear combination of the original variables $Z = XV$, where $X = \left[x_1, x_2, ..., x_p\right]$ is the set of p quantitative independent variables and $V = \left[v_1, v_2, ..., v_p\right]$ matrix., built by the eigenvectors of the correlation matrix R, whose elements are the correlation coefficients among the regressors. At second step the multinomial model is carried out on the subset of *p* PC's. The probability for the individual *i*, to choose the alternative *c* can be expressed in terms of all PC's as

$$p_i(c) = \frac{\exp\left\{\sum_{j=1}^{p}\sum_{k=1}^{p} z_{ik} v_{kj} \beta_{jc}\right\}}{\left\{\sum_{b=1}^{s}\exp\sum_{j=1}^{p}\sum_{k=1}^{p} z_{ik} v_{kj} \beta_{jb}\right\}} = \frac{\exp\left\{\sum_{k=1}^{p} z_{ik}\gamma_{kc}\right\}}{\left\{\sum_{b=1}^{s}\exp\sum_{k=1}^{p} z_{ik}\gamma_{kb}\right\}} \tag{22}$$

where $z_{ik}, (i = 1,\ldots,n; g = 1,\ldots,p)$ are the elements of the PC matrix, $v_{kj}, (k = 1,\ldots,p)$ are the elements of the transposed matrix $\mathbf{V}^T$, $\gamma_{kb} = \sum_{k=1}^{p} v_{kj}\beta_{jb}, (b = 1,\ldots,s)$ are the coefficients to be estimated, β_{jb} are the parameters expressed in function of original variables and *s* is the number of alternatives of the data set.

At third step only some of the components are chosen and the multinomial logit model is carried out only on this subset and finally the multinomial model parameters can be expressed in function of original variables: $\hat{\mathbf{a}}^{(a)} = \mathbf{V}^{(a)}\tilde{\mathbf{a}}^{(a)}$. The problem is to decide which components to retain in the analysis. One way is to choose only the components that have an eigenvalue bigger than one, that is one of the criteria used in Principal Component Analysis, but in this way we don't take in consideration the relation with the dependent variable. For this reason, the proposal is to use only the components that influence in statistical significant manner the response variable, without taking in account the eigenvalue.

The authors showed, via a bootstrap resampling, that this leads to lower variance estimates of model parameters comparing to classical multinomial model. The data they used for the analysis concern the choice of the residence in the Zurich area (Burgle, 2006) and the analysis are carried out by BIOGEME (Bierlaire, 2003). For this analysis they showed what's happen about the bootstrap estimate of variance of the estimated parameters. The results are reported in table 1 and they are relative only to the variables that are significant on the original dataset.

Tab. 1 – Variance of estimated parameters for different methods

Variables	Basic value	MNL	PCMR(6)	PCMR(2)
Distance from working place	0.795633	0.034862	0.01947	0.00264
Rent in CHF	0.101965	0.08315	0.00262	0.000545
Density of open space (500 m radius)	0.000354	0.016107	0.0026	0.00191
Density of young people (1 km radius)	4.390690	0.002317	0.01596	0.00296
Sun index	0.119155	0.026059	0.00974	0.00593
Driving time to city center	2.917186	0.456245	0.00444	0.00417

In the first column there are the names of the variables, in the second column there are the coefficient values, then, from the third to the fifth column, the variances of the estimated parameters, calculated for different datasets: the original matrix; the six PC's with eigenvalue bigger than one; the two components that are significant to explain the dependent variable. It is easy to observe that the variance of the estimated parameters computed on the original variables, is always bigger than the estimated variance calculated using the PCMR and that lower variance estimates of model parameters (last column) are obtained retaining in the model only the significant components.

6. Conclusions and further perspectives

Transport planning, infrastructure project evaluation and policy making, particularly at the urban level, continue to be important issues in the 21st century. Transport modelling requires mathematical techniques in order to make predictions, which can then be utilised in planning and design. This is the basis for improved decision-making and planning in the transport arena.

This paper has focused on some of the issues connecting with discrete choice models, which are the most commonly used tools to model users' behaviour. By far the model specification which is used most often is the Multinomial Logit Model which provides a convenient closed form for the underlying choice probabilities without any requirement of multivariate integration. Therefore, choice situations characterized by many alternatives can be treated in a computationally convenient manner. The ease of computation and the existence of a number of computer programs has led to many applications of the Logit model. Yet it is widely known that a potentially important drawback of the MNL model is the independence from irrelevant alternatives property. This property states that the ratio of the probability of choosing any two alternatives is independent of the attributes of any other alternative in the choice set. But we showed that, in many circumstances, this property is not respected and the application of the Multinomial Logit Model can lead to some problems. Furthermore, the different tests used to verify if the IIA property has been violated can give opposite results when there is multicollinearity between the explicative variables. In these circumstances, if at least one of the tests leads to a rejection of the IIA property, it is better to use another model, for example a Nested Logit model. In this case, we need to build the nests for the model, but there is not a standard way to construct the classes. Therefore, when we have many alternatives, there are different possibilities and it is not so easy to find the right combination. For this reason, we think that the use of a cluster analysis combined with a multivariate method can be useful to provide information about the construction of the nests.

Furthermore we showed that the use of PCMR can improve parameters estimation in Multinomial Logit Model. Obviously further research is necessary on the new proposed method as, for example, an extensive simulation study to verify the results and the method for selecting the PC's, but we think that it can be a first solution to the problem of multicollinearity. At the same time a generalization of other models proposed in literature for solving the problem of high-dimensional multicollinear data in the binary logit model could be interesting.

References

Aguilera A.M., Escabias M., Valderrama M.J. (2006), *Using principal components for estimating logistic regression with high-dimensional multicollinear data*, Computational Statistics & Data Analysis, 50, 1905-1924.

Bastien P., Esposito V., Tenenhaus M. (2005), *PLS generalised linear regression*, Computational Statistics & Data Analysis, 48 (1), 17-46.

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

Bierlaire M. (2003), *BIOGEME: A free package for the estimation of discrete choice models*, Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Burgle M. (2006), *Residential location choice model for the greater Zurich area*, in '6th Swiss Transport Research Conference', Monte Verità, Ascona.

Camminatiello I., Lucadamo A. (2008), *Estimating multinomial logit model with multicollinear data,* Proceedings of conference "Metodi Modelli e Tecnologie dell'informazione a Supporto delle Decisioni" – Lecce, Italy.

Cascetta E. (2009), *Transportation Systems Analysis: Models and Application*, Springer New York.

Cascetta E., Pagliara F., Papola A. (2008), *Governance of urban mobility: complex systems and integrated policies*, Advances in Complex Systems, 10, 339-354.

Cheng S., Long J.S. (2007), *Testing for IIA in the multinomial logit model*. Sociological Methods & Research 35 (4), 583-600.

D'Ambra L. (2008), *Modelli di regressione a risposta discreta*. Appunti del corso "Modelli statistici per le decisioni economiche" A.A. 2008-2009.

Domencich T., McFadden D. (1975), *Urban Travel Demand – A behavioral Analysis*, North Holland, Amsterdam.

Hausman J., McFadden D. (1984), *Specification tests for the multinomial logit model*, Econometrica 46, 403–426.

Manheim M. (1979), *Fundamentals of transportation systems analysis*, MIT Press, Cambridge, Mass.

Marx B.D. (1992), *A continuum of principal component generalized linear regression*, Computational Statistics & Data Analysis, 13, 385-393.

McFadden D. (1974), *Conditional logit analysis of qualitative choice behavior*, in P. Zarembka, ed., 'Frontiers in Econometrics', New York: Academic Press, 105–142.

McFadden D., Tye W., Train K. (1976), *Diagnostic Tests for the Independence from Irrelevant Alternatives Property of the Multinomial Logit Model*, Institute of Transportation Studies, Univ. of California, Berkeley, Working Paper No. 7616, Aug.

McFadden D. (1987), *Regression-based specification tests for the multinomial logit model*. Journal of Econometrics. 34, 63-82.

Ortuzar J.de D., Willumsen L.G. (2001), *Modelling Transport*, John Wiley and Sons, New York, 3rd edition.

Sheffi Y. (1985). *Urban transportation networks*. Prentice Hall, Englewood Cliff, NJ.Wachs.

Small K., Hsiao C. (1985), *Multinomial Logit Specification Tests*, Technical report, depatment of Economics Princeton University, Princeton, N. J.

Vágó E., Kemény S. (2006), *Logistic ridge regression for clinical data analysis (a case study)*, Applied ecology and environmental research 4(2), 171-179.

G. & L. E. R., Vol. 13 No. 2, 2009

Luisa Scaccia[*]

RANDOM PARAMETERS LOGIT MODELS APPLIED TO PUBLIC TRANSPORT DEMAND

Abstract

The multinomial logit model has been used widely as a fundamental tool for the analysis of discrete choices and has found large application in transport studies. However, its restrictive assumptions, such as independence from irrelevant alternatives (IIA) and preference homogeneity across respondents, have motivated the development of more flexible model structures that allow for an increasingly realistic representation of travel behaviour. Among these, a primary role is played by random parameter models. This paper proposes a comparison between two different specifications of random parameter models, namely the mixed logit and the discrete mixture model. An application to public transport demand is illustrated.

JEL CLASSIFICATION: C25; C23; C51

KEYWORDS: DISCRETE CHOICE; DISCRETE AND CONTINUOUS DISTRIBUTIONS; MIXTURE MODELS; RANDOM PARAMETERS.

1. Introduction

Over the past thirty years, the area of travel behaviour research has made vast usage of discrete choice models belonging to the family of Random Utility Models (RUM). For a long time, the high cost of estimating advanced models meant that most applications were limited to the use of the most basic model structures, such as multinomial

[*] Dip. di Istituzioni Economiche e Finanziarie, Università di Macerata, via Crescimbeni 20, 62100 Macerata, Italy, Tel. +39 0733 258 3242, e-mail: scaccia@unimc.it.

and nested Logit. In particular, the Multinomial Logit (MNL) model, or conditional logit model (Daniel McFadden, 1974), possesses many advantages in terms of closed-form solution and simplicity of interpretation and use.

However, the MNL model makes some restrictive assumptions such as independence from irrelevant alternatives (IIA) and preference homogeneity across respondents. Such drawbacks have led to increasing dissatisfaction with the MNL approach. Partly as a response to the perceived weaknesses of the MNL model, partly as a result of the gains in computing power and estimation techniques, Random Parameters Logit (RPL) models have grown in popularity with discrete choice modellers (Kenneth Train, 1998, 2003; McFadden and Train, 2000) over the past ten years. In this approach, the utility of each individual is a function of transport attributes (and, eventually, of individual socioeconomic characteristics) with attribute coefficients that are random and reflect individual preferences. The distribution of the coefficients is generally supposed to be continuous, leading to the so called Mixed Multinomial Logit (MMNL) model. Less frequently, a discrete distribution is used to approximate the real underlying distribution of the random parameters. The resulting Discrete Mixture (DM) model, or mass point mixed logit model, is a particular case of the Latent Class (LC) model (Wagner Kamakura and Gary Russell, 1989; Peter Boxall and Wiktor Adamowicz, 2002). LC models, which have been used in the context of transport studies by, for example, William Greene and David Hensher (2003) and Back Jin Lee et al. (2003), capture taste heterogeneity by assuming that the underlying distribution of tastes can be represented by a discrete distribution, with a small number of mass points that can be interpreted as different classes or segments of individuals. The probability of an individual being assigned to a specific class is modelled as a function of attributes of the respondent and possibly of the alternatives. In DM models, instead, these allocation probabilities are independent of explanatory variables and are simply given by constants that are to be estimated.

Compared to the MMNL model specification, the DM model has the advantage of being relatively simple, reasonably plausible, and computationally appealing. However, it is somewhat less flexible than the MMNL since the attribute parameters in each class are fixed. In contrast, the main disadvantage of the MMNL is that the distribution of parameters should be specified by the analyst. The recognition of the fact that each model has its virtues and limitations has motivated a flourishing literature attempting to compare these two approaches. Greene and Hensher (2003) compare the MMNL model with the LC model in the context of a real application. Stephane Hess et al. (2007) propose a systematic comparison of continuous and discrete mixture models making use of both real and simulated data. Other applications comparing MMNL and DM specifications are those of Sergio Colombo et al. (2008), Junyi Shen et al. (2006), and Luisa Scaccia (2005), while a theoretical discussion can be found in Michel Wedel et al. (1999). Greater efforts are, however, advocated by Greene and Hensher (2003) to further compare and contrast such advanced models.

The aim of this paper, hence, is to re-explore the potential advantages and limits of MMNL and DM models both from a theoretical point of view and through an application to data on public transport demand. The hope is also to encourage a more widespread use of these models, since the vast majority of large-scale real-world applications still rely mainly on the use of MNL and nested Logit.

The paper is organised as follows. Section 2 briefly resumes the MNL, with special emphasis on its limitations. Section 3 describes MMNL and DM models and the way in which these models overcome most of the limits of MNL model. In Section 4, we apply these two specifications of RPL models to the analysis of some stated preference data on public transport demand. Finally, conclusions are given in Section 5.

2. The Multinomial logit model

2.1 Model specification

Random utility models assume that the decision maker has a perfect discrimination capability and thus, when faced with J possible alternatives, he will choose the one that maximizes his own utility. Therefore an individual n will choose alternative i if and only if $U_{ni} > U_{nj}$, $\forall i \neq j$, where

U_{ni} is the utility that individual n is associating with alternative i. While this utility is known to the decision maker, the researcher is, instead, supposed to have incomplete information and he can only observe a portion of the utility, V_{ni}, which is called *representative utility*. Therefore, utility is decomposed as $U_{ni} = V_{ni} + \varepsilon_{ni}$ where ε_{ni} captures the factors that affect utility but are not observable and, thus, are treated as random by the researcher. Assuming a particular density, $f(\varepsilon_n)$ for the vector $\varepsilon_n = (\varepsilon_{n1}, \ldots, \varepsilon_{nj})$ the researcher can make probabilistic statements about the individual's choice, i.e. the probability that individual n chooses alternative i is $P_n(i) = \Pr(U_{ni} > U_{nj}, \forall i \neq j) = \int_{\varepsilon_n} I(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}, \forall i \neq j) f(\varepsilon_n) d\varepsilon_n$ where $I(\cdot)$ is the indicator function. Different discrete choice models are obtained from different specifications of the distribution of the unobserved portion of utility.

The MNL model assumes that the ε_{ni} are independently, identically distributed (i.i.d.) type I extreme value, i.e. $f(\varepsilon_{ni}) = e^{-\varepsilon_{ni}} e^{-e^{-\varepsilon_{ni}}}$. Under this assumption the probability that individual n will choose alternative i is simply obtained, after some algebraic manipulation, as:

$$P_n(i) = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}} \tag{1}$$

When using a linear in parameters specification for the representative utility, equation (1) becomes:

$$P_n(i) = \frac{e^{\beta' x_{ni}}}{\sum_{j=1}^{J} e^{\beta' x_{nj}}} \tag{2}$$

where x_{ni} is a vector of observed variables relating to alternative i and individual n and β is a vector of unknown parameters.

2.2 Properties of the multinomial logit model

Under the assumptions of the MNL model, the choice probabilities have a simple closed form and are analytically differentiable, and this makes it possible to apply the traditional maximum likelihood procedures for parameter estimation. Moreover, the log-likelihood function

$$\boldsymbol{L}(\beta) = \sum_{n=1}^{N} \sum_{i=1}^{J} y_{ni} \ln P_n(i),$$

in which N is the number of individuals in the sample, y_{ni} is a dichotomic variable equal to 1 if person n chooses alternative i and to 0 otherwise, and in which $P_n(i)$ is given in equation (2), is globally concave in parameters β (McFadden, 1974), which helps in the numerical maximization procedures. The simplicity of the MNL model and its computational attractiveness have made it the most widely used method for discrete choice analysis.

However, the assumption that the disturbances are i.i.d. represents an important restriction. First of all, this assumption implies that the utilities associated to different alternatives are uncorrelated, i.e.:

$$Cov(U_{ni}, U_{nj}) = Cov(\varepsilon_{nj}, \varepsilon_{nj}) = 0$$

and thus the MNL model is not capable of accounting for unobserved similarities among alternatives. Strictly related to this, is the well known property of independence from irrelevant alternatives (IIA), which states that the ratio between the probabilities of choosing two different alternatives i and k is independent from alternatives other than i and k. In fact, for the MNL logit we have:

$$\frac{P_n(i)}{P_n(k)} = \frac{e^{V_{ni}} \Big/ \sum_j e^{V_{nj}}}{e^{V_{nk}} \Big/ \sum_j e^{V_{nj}}} = e^{V_{ni} - V_{nk}}, \quad \forall i, k$$

This ratio is the same no matter what other alternatives are available or what the attributes of the other alternatives are. The same issue can be expressed in terms of cross-elasticity. It can be easily shown that, according to the MNL logit, an improvement in the attributes of an alternative reduces the probabilities for all the other alternatives by the same percentage. This pattern of substitution between alternatives is clearly unrealistic in most situations.

Another shortcoming of the MNL model is its lack of flexibility as a representation of behaviour, implying that all individuals have the same tastes. More precisely, the MNL model can accommodate for systematic taste variation: socioeconomic variables can be included as interactions with attributes or as interactions with alternative-specific constants, or different models can be estimated for different subsets of data. Hence, the MNL model can accommodate for variation that relates to observed characteristics of the individuals, but not for random taste variation. Finding better ways to represent heterogeneity in choice modelling is important to improve understanding of the factors underlying consumer behaviour and willingness to pay, and how the benefits and costs of policies are distributed across recipients. Finally, the MNL logit model can be extended to the analysis of panel data only if the unobserved factors that affect choices can be considered independent over the repeated choices. In this case, the MNL model can be used to examine panel data in the same way as purely cross-sectional data, assuming that the error components are independent over individuals, choices and time. Then, the probability that individual n will chose the set of alternatives $\mathbf{i} = \{i_1, \ldots, i_T\}$ is:

$$P_n(\mathbf{i}) = \prod_{t=1}^{T} \frac{e^{\beta' x_{ni_t t}}}{\sum_{j=1}^{J} e^{\beta' x_{njt}}}$$

The model can accommodate for dynamics linked to observed factors, such as a person's past choices influencing current choices or lagged response to changes in attributes. However the independence of errors over repeated choices makes it impossible to handle dynamics associated with unobserved factors.

3. Random parameters logit models

3.1 Model specification

The RPL models assume, as well as the MNL model, that the ε_{ni} are i.i.d. type I extreme value. However, the parameters β are no longer considered as fixed, but are now assumed to vary across the population, according to a certain probability distribution. The result is a mixture of models, where the underlying choice probability, conditional on the value of the parameters β, is simply the logit probability, i.e.:

$$P_n(i \mid \beta) = \frac{e^{\beta' x_{ni}}}{\sum_{j=1}^{J} e^{\beta' x_{nj}}}.$$

When the parameters β are assumed to vary in a continuous way in the population, according to a probability density function $f(\cdot)$, the mixture of models is defined as:

$$P_n(i) = \int_{\beta} P_n(i \mid \beta) f(\beta)\, d\beta \tag{3}$$

and it is generally referred to as MMNL model.

As an alternative, it can be assumed that the number of possible values for the taste parameters is finite and thus the parameter β_q of a generic q-th attribute, for $q = 1, \ldots, Q$, has a discrete distribution, with m_q mass points labelled β_q^l, for $l = 1, \ldots, m_q$, each of them associated with probability π_q^l satisfying the conditions that $0 \le \pi_q^l \le 1, \forall q, l$, and $\sum_{l=1}^{m_q} \pi_q^l = 1, \forall q$. In this case, the mixture of models, generally referred to as DM model, is defined as:

$$P_n(i) = \sum_{l_1=1}^{m_1} \cdots \sum_{l_Q=1}^{m_Q} P_n\left(i \mid \beta_1^{l_1}, \ldots, \beta_Q^{l_Q}\right) \pi_1^{l_1} \cdot \ldots \cdot \pi_Q^{l_Q},$$

where the conditional probabilities $P_n\left(i \mid \beta_1^{l_1}, \ldots, \beta_Q^{l_Q}\right)$ are simply the logit probabilities.

2.2 Properties of random parameter logit models

The MNL model is clearly a particular case of both the MMNL and the DM models, which is obtained when the mixing distribution of β is degenerate at fixed values. Apart from this trivial case, RPL models generalize the MNL model and allow to overcome its limits.

First of all, in the MMNL and the DM models, the parameters β are random and represent the tastes of individual decision makers, thus allowing for heterogeneous tastes in the population. In this framework, unlike the MNL model, both systematic and random taste variations can be accommodated. Variations related to observed attributes of the individual can be captured through specification of explanatory variables, as in MNL models, and/or the mixing distribution. For example, cost of transport may be divided by the individual's income to allow the relative importance of cost to decline as income rises. The random parameter of this variable, then, represents the unobserved variation over people with the same income on the value that they place on cost. This unobserved taste variation cannot be captured under the MNL model.

Secondly, under RPL models, the utilities of different alternatives can be correlated even if the errors are independent over alternatives. For example, for the MMNL model:

$$Cov(U_{ni}, U_{nj}) = Cov(\beta' x_{ni} + \varepsilon_{ni}, \beta' x_{nj} + \varepsilon_{nj}) = x_{ni}' x B_{nj}$$

where B is the variance-covariance matrix of β. Thus various correlation patterns, and hence substitution patterns, can be obtained. For example, a situation in which an improvement in alternative i draws proportionally more from alternative j than from alternative k can be easily represented, simply specifying an element of x that is positively correlated between i and j but negatively correlated or uncorrelated between i and k, and allowing the parameter of this variable to be random. This flexibility in representing various substitution patterns obviously breaks the undesirable IIA property, which characterizes the MNL model.

Finally, RPL models can be easily generalized to allow for repeated choices by each sampled individual. A simple way to do it consists in treating the random parameters as varying over people but being constant over choice situations for each person. Utility from alternative i in choice situation t by person n can then be written as $U_{nit} = \beta_n' x_{nit} + \varepsilon_{nit}$, with ε_{nit} being i.i.d. extreme value over time, individuals and alternatives, and β_n being the parameter vector specific to subject n. Hence, the conditional probability of subject n choosing the set of alternatives $\mathbf{i} = \{i_1, \ldots, i_T\}$ is:

$$P_n(\mathbf{i} \mid \beta_n) = \prod_{t=1}^{T} \frac{e^{\beta_n' x_{n i_t t}}}{\sum_{j=1}^{J} e^{\beta_n' x_{njt}}} \tag{4}$$

and the unconditional probability under the MMNL model is then obtained integrating the conditional probabilities with respect to the mixing distribution of β. Lagged dependent variables can also be added without changing the conditional probabilities in equation (4), since lagged variables entering U_{nit} will be uncorrelated with the error terms for period t. A similar generalization can be easily obtained also for the DM model. In both cases, the randomness of the parameters allows the inclusion in the model of dynamics associated with unobserved factors.

The flexibility in representing taste heterogeneity, substitution patterns among alternatives and temporal dynamics in panel data, offered by RPL models, comes at the cost of complex specification, estimation and application issues related to these models (see Hensher and Greene, 2003). In terms of model specification, the first relevant issue is that of the choice of the parameters that are to be random. This choice is particularly important since the random parameters are the basis for accommodating correlation across alternatives and defining the degree of unobserved heterogeneity. McFadden and Train (2000) suggest a Lagrange multiplier test for testing the presence of random components against the null hypothesis of fixed-value parameters. This test statistic has the advantage that its asymptotic distribution under the null hypothesis does not depend on the parameterization of the mixing distribution under the alternative. However, it is recognized to have a low power in most situations and hence further research in this field is advisable in order to develop more powerful procedures.

Once the parameters that are to be random have been chosen, under the MMNL the problem arises of selecting an appropriate distribution for these parameters. Most popular specifications have been the normal, triangular, uniform and lognormal distributions. However, in practical applications, any of them has shown its deficiencies, generally related to sign and length of tails – see Hess et al. (2005) for an interesting discussion of this issue. Clearly, the assumptions made during model specification have a direct influence on model results, and an inappropriate choice of mixture distribution for a given taste coefficient can lead to problems in interpretation and potentially misguided policy-decisions (Cinzia Cirillo and Kay Axhausen, 2006; Mogens Fosgerau, 2006; Hensher, 2006; Hess and Axhausen, 2004). Hensher and Greene (2003) suggest an empirical procedure based on a kernel density estimator to parameter estimates after applying a jackknife procedure to a multinomial logit model. This method allows one to visually inspect the distribution of parameters. Mogens Fosgerau and Michel Bierlaire (2007) proposed a method based on a seminonparametric specification to test if a random parameter of a discrete choice model indeed follows a given distribution. Fosgerau (2008) also describes a nonparametric test procedure which uses a combination of smoothed residual plots and a test statistic able to detect general misspecification.

The problem of selecting an appropriate distribution does not arise in the DM model. The use of a discrete distribution may be seen as a nonparametric estimator of the random distribution and the researcher is not required to make any prior assumption on the shape of this distribution. However, in this case, the issue of selecting the number of support points arises. The likelihood ratio test cannot be used to choose between models with different numbers of support points, even if models are nested, and it is necessary to resort to information criteria like the AIC (Hirotugu Akaike, 1974) or the BIC (Gideon Schwarz, 1978).

The issue of parameter estimates is also more complicated in RPL models than in the MNL model. For the MMNL, the choice probabilities in equation (3) do not have a closed form and simulation methods are required for parameters estimation. In practice, M values $\beta_{(1)},\mathrm{K}\ ,\beta_{(M)}$ are drawn from $f(\hat{a})$ and used to calculate the simulated probabilities

$$\tilde{P}_n(i)=\frac{1}{M}\sum_{m=1}^{M}\frac{e^{\beta_{(m)}'x_{ni}}}{\sum_j e^{\beta_{(m)}'x_{nj}}}$$

which are then used to evaluate the simulated log-likelihood

$$SLL(\beta) = \sum_{n=1}^{N} \sum_{i=1}^{J} y_{ni} \ln \widetilde{P}_n(i)$$

The number of draws required to secure a stable set of parameter estimates varies enormously, according to the complexity of the model specified, and estimation can be particularly time consuming. The use of Halton draws (John Halton, 1960) can sensibly reduce computational time, however some authors stress the need for further investigation of their properties in simulation-based estimation (see Zsolt Sándor and Kenneth Train, 2004).

An obvious advantage of the DM approach compared to MMNL models is that, if the model for the conditional choice probabilities used inside the mixture has a closed form, as the MNL model does, then the DM has itself a closed form. However, the non-concavity of the log-likelihood function does not allow the identification of a global maximum, even for discrete mixtures of MNL. Given the potential presence of a high number of local maxima, performing several estimations from various starting points is advisable. Moreover, constrained maximum likelihood must be used to account for constraints on the weights π_q^l. Finally, parameter estimation in DM models frequently suffers from clustering of mass points, which causes models with more support points to collapse back to more parsimonious specifications.

All these specification, estimation and application issues of the RPL models, as well as the advantages they offer over the MNL model, will be highlighted in the next section through an application to real data.

4. Application

4.1 Data set description

The data set refers to a study carried out between January and Mars 2004 on the bus service which links the centre of Urbino (Italy), where the University is located, to Sogesta, a residential location with more than 100 students (Edoardo Marcucci and Luisa Scaccia, 2005). The distance between the two locations is about 2 km and the bus took about

9 minutes to cover it. The aim was to analyse the attributes of the local public transport and to investigate possible interventions to improve the service. The quality of the service was, in fact, considered as unsatisfactory and many students were used to hitchhike along the road.

In order to identify the attributes that characterize the quality of the service, a focus group of 30 students was interviewed. The group resulted particularly sensitive to five attributes of the service: cost of monthly ticket, headway, first and last run, real time information displays, bus shelters. Each attribute was further described by five levels. The attributes and corresponding levels, as used in the study, are shown in Tab. 1.

Tab. 1 - Attributes and levels characterizing the quality of the bus service

ATTRIBUTES	LEVELS
	12.30 €
	15.40 €
Cost of monthly ticket	16.00 € (actual situation)
	18.20 €
	20.40 €
	28 minutes
	31 minutes
Headway	45 minutes (actual situation)
	59 minutes
	64 minutes
First and last run	05:15 -- 03:00
	06:15 -- 02:00
	07:15 -- 01:00 (actual situation)
	08:15 -- 24:00
	09:15 -- 23:00
Information displays	Real time information displays
	Without real time information displays (actual situation)
Bus shelters	At Mercatale and Segesta
	Only at Mercatale (actual situation)

Using the software CBC (Choice - Based Conjoint) of Sawtooth Software (http://www.sawtoothsoftware.com), questionnaires were created, each containing 15 choice exercises, 11 of which were random, 2 aimed at testing the quality of the answers, and 2 aimed at testing preference stability. Each choice exercise contained four hypothetical alternatives as shown in Tab. 2. Respondents were approached randomly and face-to-face interviews were carried out at the bus stop. A total number of 50 respondents took part in the study, providing a data set of 750 observations.

Tab. 2 - An example of a stated choice experiment

ALTERNATIVE A	ALTERNATIVE B	ALTERNATIVE C	ALTERNATIVE D
Cost of monthly ticket [illegible]	Cost of monthly ticket [illegible]	Cost of monthly ticket [illegible]	None of previous alternatives
Headway [illegible] minutes	Headway 25 minutes	Headway [illegible] minutes	
First and last run [illegible]	First and last run [illegible]	First and last run [illegible]	
Without real time information displays	Real time information displays	Real time information displays	
[illegible] Macerata and Sforzacosta	[illegible] Macerata and Sforzacosta	[illegible] Macerata	

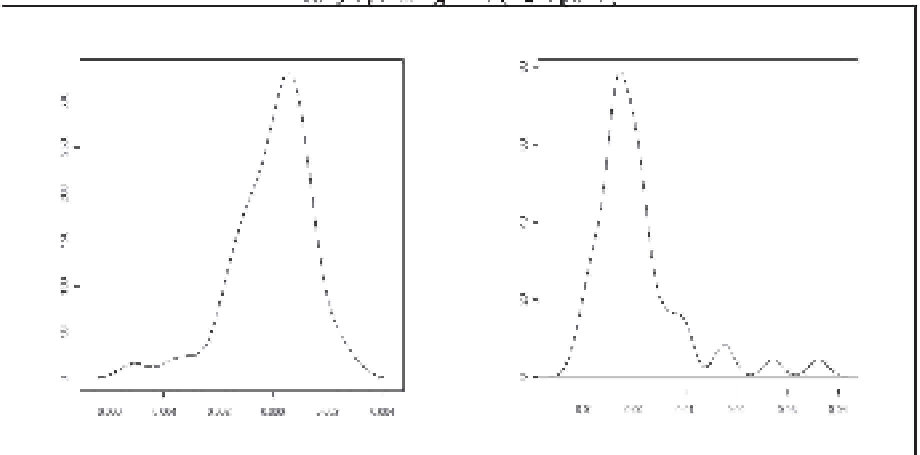
4.2 Model specification and results

Three different models were specified and estimated on the data: a simple MNL model, a MMNL model and a DM model. In the MMNL and DM models, the repeated choice nature of the data was taken into account by specifying the likelihood function with the integration and summation, respectively, outside the product over replications for the same respondent. The models were estimated in Biogeme (Bierlaire 2003, 2008). The utility of MNL model was specified as a linear function of all the attributes and also of alternative specific constants (ASC). The significance of an ASC related to an unlabelled alternative would imply that, after controlling for the effects of the modelled attributes, this alternative has been chosen more or less frequently than the others, revealing alternative ordering effects. However this was not the case with the present dataset and all the ASCs turned out to be not significantly different from zero and were thus removed from further analysis. Also the coefficient of the dummy for the presence of real time information displays turned out to be not significantly different from zero (p-value of 0.57) and the dummy variable was then removed from the specification of the utility function. Parameter estimates for the remaining attributes are given in Tab. 3. Notice that the first and last run times were specified in the utility function in terms of daily operating time (i.e. subtracting the first run time from the last run time).

To specify the RPL models, the Lagrange multiplier test suggested by McFadden and Train (2000) was used to decide which parameters are to be random. The null hypothesis of no mixing was rejected for the parameters of the attributes headway and daily operating time. The cost parameter was instead treated as non random because, in this way, the estimation of marginal willingness to pay (WTP) for an improvement in a certain attribute is simplified and its distribution is simply the distribution of that attribute's coefficient. Moreover, treating the parameter of the cost as fixed allows to restrict the cost variable to be non positive for all individuals.

Once the parameters of headway and daily operating time were chosen to be random, the jackknife procedure proposed by Hensher and Greene (2003) was used to obtain a kernel density estimator to parameters distribution. The results are shown in Fig. 1. Both of the densities seem to be unimodal and, even if some skweness can be noticed, the normal density was chosen to approximate the distribution of parameters of headway and daily operating time in the MMNL model. The model was then estimated using 1,000 random draws.

Fig. 1 Kernel density estimate for the parameters of headway (left panel) and daily operating time (right panel).



In the specification of the DM model, 2 support points were chosen for the distribution of both random parameters. When models with a larger number of support points were tried, the smallest probability happened to be further split, becoming not significantly different from zero, causing these models to collapse back to the one with 2 support points. Notice that many different starting values for parameter estimates were used to deal with the local maxima problem.

Parameter estimates for the three models specified are given in Tab. 3. It is observed that under any specification, the signs of the parameter estimates are as expected. The negative signs of the cost and headway parameters indicate that the utility of the trip maker decreases with an increase in the magnitude of the respective attributes. The utility increases instead as the daily running time does. For the qualitative attribute, bus shelters, the positive sign indicates that the presence of this attribute at both bus stations is considered utility.

Tab. 3 - Estimation results for the three models specified.

	MNL			MXL			DM(2)		
LL	-425.309			-393.456			-2[illegible].032		
Np	4			6			10		
Adj. ρ^2	0.456			0.507			0.498		
Es. time	00:00			10:21			00:02		
	Est.	Std.err.	p-value	Est.	Std.err.	p-value	Est.	Std.err.	p-value
β_{cost}	0.195	0.0367	0.00	0.255	0.0248	0.00	0.241	0.0221	0.00
$\beta_{shelter}$	0.227	0.149	0.13	0.307	0.176	0.08	0.333	0.170	0.04
$\beta_{headway}$	-0.0524	0.006	0.00	-	-	-	-	-	-
$\beta_{headway\ \mu}$	-	-	-	-0.0941	0.0149	0.00	-	-	-
$\beta_{headway\ \sigma}$	-	-	-	0.0666	0.0136	0.00	-	-	-
$\beta_{headway,1}$	-	-	-	-	-	-	-0.0351	0.00893	0.00
$\pi_{headway,1}$							0.296	0.121	0.00
$\beta_{headway,2}$	-	-	-	-	-	-	-0.132	0.0224	0.00
$\pi_{headway,2}$	-	-	-	-	-	-	0.484	0.121	0.00
$\beta_{running\ time}$	0.115	0.0025	0.00						
$\beta_{running\ time\ \mu}$	-	-	-	0.620	0.0515	0.00	-	-	-
$\beta_{running\ time\ \sigma}$	-	-	-	0.354	0.0762	0.00	-	-	-
$\beta_{running\ time,1}$	-	-	-	-	-	-	0.067	0.0714	0.00
$\pi_{running\ time,1}$	-	-	-	-	-	-	0.815	0.0775	0.00
$\beta_{running\ time,2}$							0.0551	0.0637	0.36
$\pi_{running\ time,2}$							0.185	0.0725	0.01

The interpretation of the coefficients is not meaningful except for significance and sign. Therefore, the marginal WTPs for the different attributes are calculated by taking minus the ratios between the coefficients of the attributes and the coefficient of the cost attribute. These values represent the marginal rates of substitution between the attributes and the cost and provide an idea of how much the travellers, on average, are willing to pay for a positive unit change in each quantitative attribute under consideration. The WTP for the bus shelter attribute indicates the willingness to pay to have bus shelters also at Sogesta and not only at Mercatale. In the MMNL model, with a fixed cost coefficient and normally distributed attributes, marginal WTPs are also normally distributed. Tab. 4 summarizes the marginal WTP estimates from the different models under consideration. For example, the MMNL shows an average WTP of 0.36 Euros more on the monthly ticket to obtain 1 minute less of headway. This value is very closed to the one obtained under the MNL model. Under the DM model, the population is divided into two groups of almost equal size with respect to their WTP for shorter headway. The first group shows a very small propensity to pay for a shorter headway, while the second group is much more sensitive to a long headway and hence is willing to pay more to shorten it. Looking at the WTP for operating time, it can be noticed that the DM model identifies a segment of the population, the 18.5% of it, which is not sensitive to operating time (the parameter $\beta_{run.time,\,2}$ in Tab. 3 is not significantly different from zero) and hence is not willing to pay for an improvement of this attribute, i.e. for an extension of the operating time.

Tab. 4 - Estimation of the WTP for the different attributes under the three models specified.

WTP	MNL	MMNL	DM(2)	
$\mu_{shelter}$	1.54	1.37	1.46	
$\sigma_{shelter}$	-	-	-	
$\mu_{headway}$	-0.34	-0.36	-0.16	-0.71
$\sigma_{headway}$	-	0.26	-	-
$\mu_{run.time}$	2.15	2.40	2.80	0.24
$\sigma_{run.time}$	-	1.49	-	-

Moving on to the comparison of the three different models, from Tab. 3 it can be noticed that both the RPL models offer a significant improvement in model fit over the MNL model. The MMNL offers the best performance, with an adjusted ρ^2 equal to 0.507. The DM model, even if improving over the MNL model, does not fit the data so well as the MMNL model. This probably depends on the fact that the inspection of the parameter distributions through the jackknife procedure seems to reveal unimodal distributions, for which the DM with 2 mass points might offer a not very good approximation. Probably a larger number of support points would be required, if the real parameter distributions were effectively unimodal and continuous, but this would determine difficulties in the estimation procedures, as alluded to in Section 3.2.

Another observation relates to the much lower estimation cost of the DM model, with an estimation time of 2 seconds, compared to the 10 minutes and 21 seconds required by the MMNL model. This much lower estimation time would give a significant advantage to the DM model in the case of larger data sets. However, the estimation time for the MMNL model is particularly high, since we decided to use random rather than Halton draws in the estimation.

5. Summary and conclusions

RPL models allow to overcome all the limits of the simple MNL model. As a drawback, several issues arise in both model specification and estimation. In this paper, we summarized the theoretical results concerning RPL models, considering both continuous distributions, leading to the MMNL model, and discrete distributions, leading to the DM model, for the parameters. We outlined the advantages and the limits of both these RPL models, also making use of an application to a real data set concerning the public transport demand.

The results from the application clearly show the major advantage of the DM approach in terms of estimation cost, due to the fact that parameter estimates have a closed form solution and does not rely on simulation processes. Moreover, this approach avoids the problem of choosing an adequate distribution for the random parameters. However, even if the DM model provides a considerable improvement over the MNL model, for the example at hand it does not seem to be competitive with the MMNL model in terms of fit to the data. The reason is probably that, in this case, the underlying distribution of the random parameters seems to be continuous and unimodal, and, thus, 2 support points are not enough to provide a reasonable approximation to it. Hess et al.

(2007), in a simulation study, find that DM models, with a sufficiently large number of support points, can offer a very good approximation to the normal distribution. However, with our data set, we experienced a clustering of mass points which did not allow us to estimate DM models with a number of support points larger than 2. This is probably due, also, to the fact that the data set is a relatively small one.

In this regard, the paper provides a different perspective to that of Hess et al. (2007), Shen et al. (2006) and Colombo et al. (2008) in the ongoing discussion about the comparison between MMNL and DM models. In small data sets, with continuous and unimodal underlying distribution of the random parameters, MMNL models, generally requiring a smaller number of parameters to account for heterogeneity in tastes, could perform better than DM models.

References

Akaike H. (1974), *A new look at the statistical model identification*, IEEE Transactions on Automatic Control, 19, 716-723.

Bierlaire M. (2003). *BIOGEME: A free package for the estimation of discrete choice models*, Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Bierlaire M. (2008). *An introduction to BIOGEME Version 1.7*, biogeme.epfl.ch.

Colombo S., Hanley N., Louviere J. (2008), *Modelling preference heterogeneity in stated choice data: an analysis for public goods generated by agriculture*, Stirling Economics Discussion Papers 2008-28, University of Stirling, Department of Economics.

Boxall P.C., Adamowicz W. (2002), *Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach*, Environmental and Resource Economics, 23, 421-446.

Cirillo C., Axhausen K.W. (2006), *Evidence on the distribution of values of travel-time savings from a six-week travel diary*, Transportation Research Part A, 40, 444-457.

Fosgerau M. (2006), *Investigating the distribution of the value of travel time savings*, Transportation Research Part B, 40, 688-707.

Fosgerau M. (2008), *Specification testing of discrete choice models: a note* *on the use of a nonparametric test*, Journal of Choice Modelling, 1, 26-39.

Fosgerau M., Bierlaire M. (2007), *A practical test for the choice of mixing distribution in discrete choice models*, Transportation Research Part B, 41, 784-794.

Greene W.H., Hensher D.A. (2003), *A latent class model for discrete choice analysis: contrasts with mixed logit*, Transportation Research Part B: Methodological, 37, 681-698.

Halton J. (1960), *On the efficiency of evaluating certain quasi-random sequences of points in evaluating multi-dimensional integrals*, Numerische Mathematik, 2, 84-90.

Hensher D.A. (2006), *The signs of the times: imposing a globally signed condition on willingness to pay distributions*, Transportation, 33, 205-222.

Hensher D.A., Greene W.H. (2003), *Mixed logit models: state of practice*, Transportation, 30, 133-176.

Hess S., Axhausen K.W. (2004), *Checking our assumptions in value-of-travel-time modelling: Recovering taste distributions*, CTS Working Paper, Centre for Transport Studies, Imperial College London.

Hess S., Bierlaire M., Polak J.W. (2005), *Estimation of value of travel-time savings using mixed logit models*, Transportation Research Part A: Policy and Practice, 39, 221-236.

Hess S., Bierlaire M., Polak J.W. (2007), *A systematic comparison of continuous and discrete mixture models*, European Transport \ Trasporti Europei, 37, 35-61.

Kamakura W., Russell G. (1989), *A probabilistic choice model for market segmentation and elasticity structure*, Journal of Marketing Research, 26, 379-390.

Lee B.J., Fujiwara A., Zhang J., Sugie Y. (2003), *Analysis of Mode Choice Behaviours based on Latent Class Models*, paper presented at the 10th International Conference on Travel Behaviour Research, Lucerne.

Marcucci E., Scaccia L. (2005), *Alcune applicazioni dei modelli a scelta discreta al settore dei trasporti*, In: E. Marcucci (Ed.), I modelli a scelta discreta nel settore dei trasporti. Teoria, metodologia e applicazioni, Carocci editore, Roma.

McFadden D. (1974), *Conditional logit analysis of qualitative choice behaviour*, In: P.C. Zarembka (Ed.), Frontiers in Econometrics, New York, Academic Press, 105-142.

McFadden D., Train K. (2000), *Mixed MNL models for discrete response*, Journal of Applied Econometrics, 15, 447-470.

Sándor Z., Train K. (2004), *Quasi-random simulation of discrete choice models,* Transportation Research Part B: Methodological, 38, 313-327.

Scaccia L. (2005), *L'eterogeneità delle preferenze nel trasporto merci: un confronto tra diversi metodi per catturarla*, In: G. Borruso and G. Polidori (Ed.), Riequilibrio ed integrazione modale nel trasporto delle merci. Gli attori ed i casi italiani, Franco Angeli, Milano, 83-95.

Schwarz G.E. (1978), *Estimating the dimension of a model*, Annals of Statistics, 6, 461-464.

Shen J., Sakata Y., Hashimoto Y. (2006), *A Comparison between Latent Class Model and Mixed Logit Model for Transport Mode Choice: Evidences from Two Datasets of Japan*, Discussion Papers in Economics and Business 06-05, Osaka University, Graduate School of Economics and Osaka School of International Public Policy (OSIPP).

Train K. (1998), *Recreation demand models with taste differences over people*, Land Economics, 74, 230-239.

Train K. (2003), *Discrete Choice Methods with Simulation*, Cambridge University Press.

Wedel M., Kamakura W., Arora N., Bemmaor A., Chiang J., Elrod T., Johnson R., Lenk P., Neslin S., Poulsen C.S. (1999), *Discrete and continuous representations of unobserved heterogeneity in choice modelling*, Marketing Letters, 10, 219-232.

Finito di stampare
nel mese di Dicembre 2009
dalla Litografia Brandolini - Sambuceto
per le Edizioni TRACCE
Via Eugenia Ravasco, 54
65123 PESCARA
Tel. 085/76658
www.tracce.org

In the latest centuries, generally speaking, history records the alternation of some important seasons which lend themselves to represent economic models, which are the bases of modern economic thought.
First of all, there is the age of *colonial economy* centered on the role of imperial states, together with the birth of monopolistic companies, in the management of trades with dominion areas.
Then, the age of *international economy* was lived, culminating in the second post war trade relation system. It was mainly founded on the functions of the national states and their authorities to support both national espansionary fiscal policy and exchange clearings, in their trade ratios with the rest of the world.
At last, in the latest years, *interglobal economy* took vehemently the lead through the modern electronic infrastructures of telematic and telecommunications.
As the former models, the *interglobal economy* too does not automatically assure either stable equilibrium or the overcoming of traditional crises.
It gives benefits and disadvantages too.
From the normative and positive points of view, one of the disadvantages which most drew the attention of researchers is the weakening and disappearance of national and subnational economic and monetary policy instruments.
Instead one of the benefits which most attracted interest might be located on the nature itself of the technological revolution in progress, foreboding new opportunities in the integration process of local economic systems, which might qualify themselves as network growth links (or growth poles?).
The Review has the aim to represent and to inquire the normative and positive profiles of the foundamentals which might characterize the thin and difficult frontier between globalization and economic localism.

FONDAZIONE
PESCARABRUZZO

Euro 12,91

ISSN 1722-424 1

9 771722 424009