A General Framework to Testing the Effect of Transport Policy Measures to Achieve a Modal Shift: A Sequential Hybrid Model

Julio Comendador
Investigador, Centro de Investigación del Transporte TRANSyT,
Universidad Politécnica de Madrid, España
Andrés Monzón
Profesor, Departamento de Transportes, Universidad Politécnica de Madrid, España
María E. López-Lambas
Profesora, Departamento de Transportes, Universidad Politécnica de Madrid, España

ABSTRACT

In order to achieve to minimize car-based trips, transport planners have been particularly interested in understanding the factors that explain modal choices. In the transport modelling literature there has been an increasing awareness that socioeconomic attributes and quantitative variables are not sufficient to characterize travelers and forecast their travel behavior. Recent studies have also recognized that users' social interactions and land use patterns influence travel behavior, especially when changes to transport systems are introduced, but links between international and Spanish perspectives are rarely deal. In this paper, factorial and path analyses through a Multiple-Indicator Multiple-Cause (MIMIC) model are used to understand and describe the relationship between the different psychological and environmental constructs with social influence and socioeconomic variables. The MIMIC model generates Latent Variables (LVs) to be incorporated sequentially into Discrete Choice Models (DCM) where the levels of service and cost attributes of travel modes are also included directly to measure the effect of the transport policies that have been introduced in Madrid during the last three years in the context of the economic crisis. The data used for this paper are collected from a two panel smartphone-based survey (n=255 and 190 respondents, respectively) of Madrid.

1. INTRODUCTION

To forecast the impacts of new transport policies and investments (new transport infrastructure, fare policies, congestion pricing, etc.), transport planners are particularly interested in analyzing transport demand at specific moments and also in understanding the factors that explain modal choices. Ortúzar and Willumsen (2011) indicate that the choice of a transport mode is the key role played by public transport making. Discrete Choice Models (DCM) is the methodology used to analyze and predict travel choices based on Random Utility Maximization (RUM) principles.

Weis and Axhausen (2009) indicate that the travel behavior observed reflects equilibrium conditions both behaviorally, as well as the underlying network flows and travel times, individuals can adapt their travel behavior after a change in the generalized cost on several levels: (*i*) the decision to leave home (i.e., to participate in out-of-home activities on a given day); (*ii*) the adaptation of these activities; (*iii*) the scheduling of activities; (*iv*) the choice of locations for carrying out activities (destination choice); and (*v*) the choice of an origin-destination connection (mode and route choice). Thus, a behavioral DCM model deals the latter two levels, which effectively constitute the second to fourth steps in the classic four step model. The upper levels of the demand generation process (from *i* to *iii* above) are been dealt by the literature with an aggregate methodology that handles indirect and multiple relationships between all the variables that influence a travel behavior process (Golob, 2003; Moore et al., 2013): Structural Equation Modeling (SEM). Reviewing many case studies on travel behavior knowledge, it is difficult to have a brainstorm of how to integrate the five levels above to evaluate the effect of transport policy measures to achieve a modal shift.

Madrid is the capital and major financial and business center of Spain where a number of transport policies have been introduced in the last four years in the context of the economic crisis: five new metro stations, a 25% average public transport fare rise, one hour extension of the on-street parking rate, all systems have modernized their vehicles and stations, real-time information at most metro/train stops, some bus stops and on-board most metro/train and buses, most residential streets are traffic-calmed at 30 km/h or less, etc. To evaluate accurately the willingness to change patterns of urban mobility in Madrid before and after the introduction of the transport policies, it has already used the methodologies discussed above: Comendador et al. (2014a) and Di Ciommo et al. (2014) using DCMs; and Comendador et al. (2014b) with SEM approaches. The general conclusion of these studies is transport policy actions are more likely to be effective when car use has been disrupted first. But as these studies are based on different methodologies, it is necessary to confirm their results using a complete methodology on a general framework: a hybrid DCM.

Inclusion of psychological and land-use factors through a Multiple-Indicator Multiple-Cause (MIMIC) following a SEM approach helped to improve the fitness level of DCM and to provide an understanding of the role of Level Of Service (LOS) and cost attributes in the decision making process when new transport policy measures are included (Ben-Akiva et. al, 2002; Raveau et. al, 2010; Yoon and Goulias, 2010). The inclusion of Latent Variables (LVs) in DCM has reemerged as an analysis and discussion topic after losing some of the importance that made it an interesting subject in the 1980s. Furthermore, evidence is mainly based on US data and North-European countries. Therefore, this paper adds some new evidence from a Spanish perspective to the research debate. The authors use data from a two panel smartphone-based survey (n=255 and 190 respondents, respectively) of Madrid.

In this paper, hybrid DCMs are estimated for car, public transport and walk modes of transport. These models are formulated on the basis of LOS and cost attributes depend on social influence and socioeconomic variables (Comendador et al., 2014a; Di Ciommo et al., 2014). These hybrid models are compared with specifications, including psychological and built environmental factors. Results show that inclusion of these factors through a MIMIC approach indeed helped to improve the fitness level of the DCMs.

The paper is organized as follows. First, we give a review of the literature, focusing on methodology aspects. We then present the database of our research, followed by the estimation hybrid models results and some conclusions and policy implications.

2. METHODOLOGY

The aim of this paper is to study the role of psychological and environmental LVs on the mode choice process. These factors are measured by mean of psychometric and geographical tools, respectively, which are fit into DCMs through a MIMIC approach. In this chapter, first the MIMIC formulation, LVs and their explanatory variables are defined; and finally the hybrid discrete choice model formulation is presented.

2.1 MIMIC

2.1.1 Formulation

LVs are factors that, although they influence individual behavior and perceptions, cannot be quantified in practice because of their intangibility, since these variables do not have a measurement scale, or because of their intrinsic subjectivity, since different persons may perceive them differently. Identification of LVs requires supplementing a standard preference survey, either revealed or stated, with questions that capture users' perceptions about some aspects of the alternatives (and the choice context). The answers to these questions generate perception indicators that serve for identifying the LVs. Otherwise, these LVs could not be measured (Galdames et. al, 2011). Moreover, since Cervero and Kockelman (1997) some studies have treated measurable land use variables to define some environmental LVs with satisfactory results (Edwing and Cervero, 2010; De Abreu e Silva et al., 2012, Yoon and Goulias, 2010).

To use LVs, a Multiple-Indicator Multiple-Cause (MIMIC) following a Structural Equation Modeling approach model must be estimated, in which the LVs (η_{ijl}) are explained by characteristics from the users through structural equations (1); at the same time, the LVs explain the perception indicators or measurable land use variables (y_{ijp}), which are observed by the modeler from the survey through measurement equations such as (2):

$$\eta_{ijl} = \sum_r \alpha_{ilr} s_{ijr} + \nu_{ijl} \qquad (1)$$

$$y_{ijp} = \sum_l \gamma_{ipl} \eta_{ijl} + \zeta_{ijp} \qquad (2)$$

where: i (alternative); j (individual); l (latent variable); r (explanatory variable); p (indicator).

2.1.2 Determining Latent Variables

At first step it is necessary to determine/label LVs based on literature. Galdames et al. (2011) demonstrated that the inclusion of psychological factors through a LVs approach helped to improve the fitness level of Revealed Preference (RP) models and to provide and understanding of the role of Level Of Service (LOS) and cost transport mode attributes in the decision-making process. Comendador et al. (2014b) identified a LV labeled ‘comfort perception’ on a SEM approach that represented the perceptions towards public transport for all the respondents. An important conclusion about this aspect was that the as more pro public transport intention as less ‘comfort perception’ towards public transport is. Thus, there are other factors that affect in a positive way on public transport intention.

But to achieve a complete framework to explain travel behavior may be the inclusion of built environmental dimension. Applications involving travel behavior from the perspective of land use are dating from the 1990s. Usually, four important factors are distinguished: density, diversity and design (3D’s of Cervero and Kockelman, 1997) and accessibility (introduced by Geurs and van Wee, 2004). There is not a general agreement on how to measure each of those 4 components (see review of Comendador et al., 2014c). To define new evidence from the Spanish perspective to the LVs treatment on travel behavior, Comendador et al., (2014c) used a factor analysis to identify six LVs: ‘street network design’, ‘urban block diversity’, ‘nonresidential diversity’, ‘job accessibility’, ‘center accessibility’ and ‘density’.

2.2 Hybrid Discrete Choice Model

Traditionally, in the DCM it is assumed that people j are rational decision makers maximizing their own utility U_{ij}; the modeler, who is an observer, defines a representative utility V_{ij} and (since he does not have perfect information) an error term ε_{ij} associated with each alternative:

$$U_{iq} = V_{iq} + \varepsilon_{iq} \tag{3}$$

The representative utility V_{ij} is a function of the objective attributes X_{ijq} where q refers to a particular attribute (i.e., travel time or fare, as well as socioeconomic characteristics of the individual); if LVs are included, a utility function is obtained:

$$V_{ij} = \sum_q \beta_{iq} X_{ijq} + \sum_l \theta_{il} \eta_{ijl} \tag{4}$$

where β_{iq} and θ_{il} are parameters to be estimated associated with the tangible attributes and the LVs, respectively. Some studies (Cherchi and Ortúzar, 2011) include several interactions among the attributes X_{ijq} and decision makers’ characteristics that allow accounting for systematic heterogeneity (or taste variations) in the individual preference.

Since the η_{ilj} variables are unknown, the DCM must be estimated jointly with the MIMIC model's structural and measurement equations. Finally, to characterize the individual's decisions over the set of available alternatives $A(j)$, binary variables d_{ij} are defined, which take values according to:

$$d_{ij} = \begin{cases} 1 & \text{if } U_{ij} \geq U_{ik} \quad \forall k \in A(j) \\ 0 & \text{in other cases} \end{cases} \tag{5}$$

Hybrid DCMs can be estimated simultaneously, as well as sequential approach. The former consists of estimating jointly DCM and MIMIC models. While simultaneous is an accurate approach to estimate a hybrid DCM because the estimators from the simultaneous approach are both consistent and efficient (Raveau et al., 2010), the current state of the art does not allow the flexibility to do it correctly in the general case. Moreover, it has been demonstrated that the sequential approach allows for unbiased estimators (Ben-Akiva et al., 2002; Raveau et al., 2010). Therefore, in addition to minor difficulties in its application and interpretation, the sequential approach is chosen by the authors to achieve the objectives set in this paper.

In sequential estimation the problem is treated in two stages, separating the LV and DCM interactions. First, the MIMIC model is solved to obtain parameter estimators for the equations relating the LVs with the explanatory variables and perception indicators. With these parameters in the structural equation (1), expected values for the LVs of each individual and alternative are obtained. In this way, the LVs can be added to the set of variables of the DCM, and their parameters are estimated together with those of the traditional variables in a second stage.

For the purposes of this paper, besides the traditional sequential approach, the authors consider using a Mixed Logit (ML) to run the DCM, which allows to consider not only the expected value of the LVs, but also its variability (from MIMIC model), which represents a better reproduction of reality. The ML model is in particular suitable to account for repeated choices by individual decision-makers (between week-days or/and different *waves*). This model specification treats the utility coefficients as varying between individuals but remaining through all choice situations for each person (Train, 2009).

3. PANEL DATA INCLUDING SUBJECTIVE ELEMENTS: CASE OF MADRID

Public transport policy in the Madrid Metropolitan Area (MMA) is often deemed as a success. Most of it can be attributed to the CRTM, created in 1985. Fifteen years ago, the GDP per capita in the MMA was about 30% above the average of the E.U. (27,279); this means that the MMA is a relatively wealthy area, even inside the current economic crisis context. However, there are differences in terms of GDP per capita among some cities inside the MMA. Most of the wealthiest municipalities are placed in the northwest whereas the poorest are located in the south east. The GDP per capita of the wealthiest municipalities and neighborhoods is around twice the GDP per capita of the poorest (Vasallo and Perez de Villar, 2008).

The public transport system in the MMA is made up of four modes. Two of them are typically urban modes (underground and urban buses), and the other two are mostly metropolitan modes (commuter rail and interurban buses). Beside, since 2007 a new mode of transport (light underground) had been set up in order to increase public transport offer in the outskirts. Inside the economic crisis context, other pro-public transport policies have been introduced: five new metro stations, by 25% average public transport fare rise, on-street parking fee rises one hour more, all systems have modernized their vehicles and stations, real-time information at most metro/train stops, some bus stops and on-board most metro/train and buses, most residential streets are traffic-calmed at 30 km/h or less, etc.

However, it must be stressed that the aim of this paper is not only to provide a consistent framework on how to integrate some research topics already developed in previous studies and emerging methods on travel behavior, when assessing the introduction of any transport policy, but also to add new evidences from a Spanish perspective to the research debate on how to change patterns of urban mobility.

3.1. Sample design

The data used in this study come from a two wave survey carried out in the HABIT project (Habit and Inertia in mode choice behavior: a data panel for Madrid). After a process of screening and achieving consistency of both waves, the final sample size reached n=255 and 190 respondents, respectively. During fall 2011 and winter of 2012 (wave 1) and fall 2012 (wave 2), a smartphone with a panel-survey application was delivered for one week among two focus groups in order to capture a portion of the population of Madrid affected by recent changes in transport policy: (1) workers of Regional Health Department; and (2) workers of the Universidad Politécnica de Madrid, taking advantages of their close relation to the authors, which helped to easily achieve a random sample of a population of 5774 workers (2011 Census data). Since, high costs in terms of time and money are one of the biggest constraints when building data panels (Yáñez et al., 2010a), authors discarded the most common sampling unit in transport survey (i.e. the household), and panel survey used is based on a sample of a worker subpopulation.

Three different transport modes are considered: car, public transport and walking. Data for people who used no common transport modes are excluded from the analysis. For each mode, GPS information was available about travel times, number of transfers, distances recover; and then estimation costs. Regarding the users, the panel gathered information about socioeconomic variables. To generate a LV factor about public transport perceptions, respondents were asked about space, safety, cleanliness, access time, and etcetera. Lastly, land use variables were calculated with the GPS information of each travel. Table 1 contains a selection of individual and socioeconomic characteristics of this sample. Despite these restrictions, the sample well represents the Madrid worker population in many aspects as Table 1 shows.

Variables		Worker Population*	Wave 1 (n=255)		Wave 2 (n=190)	
			Average	Standard Deviation	Average	Standard Deviation
Travel behavior variables (daily)	# trips	2.6	2.4	0.3	2.1	0.4
	Travel time (min)	28.6	32.7	5.8	30.5	6.1
	Commuting distance (km)	6.0	7.9	3.7	6.5	2.6
	Car use (%)	45.0	57.3	1.8	55.4	1.9
	Pub. Transport use (%)	40.0	34.4	2.9	34.8	3.2
	Walking use (%)	12.0	8.1	2.2	9.5	1.7
Explanatory socio-economic variables	Male (%)	51	52	0.8	51	1.0
	Age	40	43	9.2	44	11.6
	Income	2500	2100	410	2220	320
	# family members	2.7	3.1	0.5	3.0	0.3

**Source: INE (2011) and Monzón et al. (2013)*

Table 1 - Sample travel behavior and socioeconomic characteristics

Each panel-based survey considered two main phases. The first phase consisted in a face-to-face interview registering personal data about the respondent. In the second phase, authors gave the smartphone to the people and asked them to register the daily trips they made during the five workdays (Monday to Friday). The trips recorded were monitored in real time and respondents were eventually contacted at the end of the day to correct or clarify the information. A chart was also given to the participants to manually register those trips not registered by the smartphone. The complete registration of daily trips, took about 20 seconds for a trip by car or by walking and one minute for a public transport journey. At the end of the trip, the data were automatically sent to a server accessible by the monitor of the survey. Both the face-to-face interview and the smartphone trip diary were based on the palm-based Santiago Panel used for evaluating the TranSantiago system in Chile (Yáñez et al., 2010b), also covering a wide variety of socio-economic variables.

3.2. Definition of the variables

3.2.1. Explanatory variables

Based on the findings of previous studies and the data availability aforementioned, following variables are included as explanatory variables for the MIMIC model: SEX (male = ‘1’; female = ‘0’); AGE; INCOME (monthly household income); CHILD (presence of child = ‘1’, no = ‘0’); FAMILY (number of family members); HELPED (if assisted for child-care or for housekeeping = ‘1’, no = ‘0’); VOLUNTARY (voluntary participation in some non-compulsory meetings or activities= ‘1’, no = ‘0’).

3.2.2. Indicators

For each trip, perceptions towards public transport were measured using 5-point Likert scale to generate a LV for the MIMIC model about the ‘comfort perception’ (Comendador et al., 2014b): ACCESS_TIME, PHYSICAL_ACCESS, TRAVEL_TIME, WAITING_TIME, TRANSFER, SPACE, SAFETY, and CLEANLINESS. Given that perception towards public transport was asked for each trip, the average perception as arithmetic mean is used assuming that the Likert scale can be treated as a continuous variable (Wang and Law, 2007).

The indicators about built environment characteristics within a service area have been calculated within the station service area (SSAs). Service areas were obtained using Geographical Information System (GIS), and are based on distances across the transport/road network. The distance threshold considered was 900 meters, which is the most suitable distance for people to walk for accessing the Metro network in Madrid (García-Palomares et al., 2013). Once the SSAs were defined with a GIS, they have been intersected with various environment indicators that that hypothetically favor transit use: density, diversity and design (Cervero, R. and Kockelman, 1997); and accessibility dimensions (Geurs and van Wee, 2004) for the MIMIC model. At Comendador et al. (2014c) appears a complete description about which land use indicators defined each built environmental LV applying a factor analysis methodology. A list of these indicators is: POP_DEN (population density); EMP_DEN (employment density); JOB_RATIO (ratio of employment per inhabitant); MIX (variation coefficient of the area covered by different land uses within SSAs); EQUIPMENT (hectares (ha) of commercial and educational land use); SINGLE_RES (ha of single-family residential); MULTI_RES (ha of multifamily residential); INDUSTRY (ha of industry); OFFICES (ha of offices); INF (ha of infrastructure that promotes economic activity, such as roads, highways, railroads, airports, electricity, telecommunications, water supply and sanitation); GREEN (ha of parks and recreations); DIST_CBD (distance of each SSAs to Center Business District); BROWNFIELD (ha of land available for building); and four centrality measures: REMOTENESS, BETWEENESS, STRAIGHTNESS and REACH centrality.

4. ESTIMATION

4.1 MIMIC model: identification and results

In a previous study (Comendador et al., 2014c) some of the MIMIC model *measurement relations* such as (2) were studied by the authors using factor analysis to guarantee their correct specification. Using the land use indicators described above, six environmental LVs were justified (‘street network design’, ‘urban block diversity’, ‘nonresidential diversity’, ‘job accessibility’, ‘center accessibility’ and ‘density’); but to capture comfort perceptions of public transport it is necessary to include an additional psychological LV:

‘comfort perception’. Therefore, seven explanatory variables are included in the MIMIC model as exogenous variables through *structural equations* (1): SEX, CHILD, AGE, FAMILY, INCOME, HELPED and VOLUNTARY.

The MIMIC model is fitted using AMOS 20.0 software package (Byrne, 2001). This model fitting is done using a covariance based structural analysis, also referred to as method of moments, consisting in minimizing the difference between actual sample covariances and those implied by the model parameters (Bollen, 1989). In the AMOS software package, computing intercepts for the endogenous variables is only feasible when using the maximum likelihood approach. The model assumes direct causal relationships between certain dependent variables, and thus goes further than merely capturing these relationships via error correlations. For reasons of space, the authors only present the details for relationships (measurement and structural equations, Figure 1) of the best model estimated according to the best goodness of fit indices (Schumacker and Lomax, 2004).

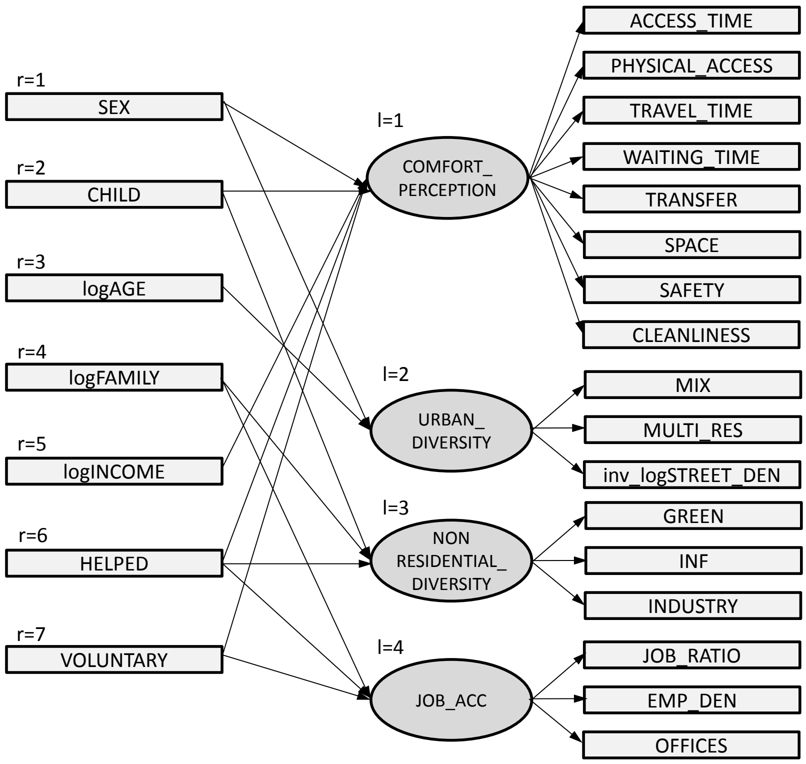

Figure 1 – Significant relationships (*p < 0.1*) of the best MIMIC model examined (fit indices: *RMSEA=0.037; CFI=0.982; AGFI=0.864*)

Then, based on this MIMIC identification and authors’ hypotheses from their own previous studies, it can be concluded that this model structure is expected, except for the following two findings: (i) the effect of the ‘street network design’ LV is not significant; (ii) the accepted general recommendation (Anderson and Gerbing, 1988) for running a MIMIC model is to have at least three or four indicators per LV, i.e., ‘center accessibility’ and ‘density’ LVs have been excluded (see Comendador et. al, 2014c), so this might imply a flaw in the survey design to fully capture the LVs’ variability. Moreover, to improve the value of standardized regression weights (α_{lr}) it is necessary to apply the logarithm for explanatory variables with scale dimension: AGE, FAMILY and INCOME.

Regarding the structural equations results for all the alternatives, the 'comfort perception' towards public transport is measured by using five explanatory variables: SEX (α_{11}=−0.100), CHILD (α_{12}=−0.067), logINCOME (α_{15}=0.199), HELPED (α_{16}=0.241) and VOLUNTARY (α_{17}=−0.260). This psychological LV is very influenced by the novel social influence variables considered: HELPED in a positive way and VOLUNTARY negatively. The 'urban diversity' is measured by SEX (α_{21}=0.129) and logAGE (α_{23}=−0.263). Higher age is associated with living in an area covered by not many land uses. The other *diversity* LV that explains the hectares of parks, infrastructures and industries -'nonresidential diversity'- is measured by three explanatory variables related to *family children*: CHILD (α_{32}=0.129), logFAMILY (α_{34}=0.082) and HELPED (α_{36}=0.114). Finally, the 'job accessibility' is measured negatively by logFAMILY (α_{44}=0.120) and by the novel social influence variables: HELPED (α_{46}=−0.076) and VOLUNTARY (α_{47}=−0.072).

4.2 Hybrid DCM: identification and results

Once the expected values of the LVs are calculated for each respondent from the MIMIC model, DCMs are estimated. These models are formulated on the basis of LOS and other socioeconomic variables (Di Ciommo et al., 2014): 'travel time', 'personal travel cost', 'gender=male', 'car ownership', 'purpose=work'. Moreover, the authors include the known *social influence* variables ('voluntary' and 'helped') to evaluate the true importance of this social aspect on a general framework.

Table 2 presents the estimation results. For every estimator the respective *t*-value follows between parentheses. The models' log-likelihood at convergence and ρ^2 index (calculated with respect to the constants-only model) are shown as well. To evaluate the importance of the LVs inclusion and to facilitate the conclusions, all the DCMs estimated below have a ML model structure with panel correlation accounted for by the 'travel time' and 'personal travel cost' random parameters.

Parameters	**ML1**	**ML2**	**HM1**	**HM2**
Total Travel Time (mean)	-0.074 (*-27.0*)*	-0.062 (*-17.1*)*	-0.083 (*-25.2*)*	-0.067 (*-16.7*)*
Total Travel Time (st.dev)	0.148 (*24.9*)*	0.148 (*24.8*)*	0.115 (*23.1*)*	0.116 (*23.0*)*
Personal Travel Cost (mean)	-0.142 (*-12.8*)*	-0.146 (*-13.3*)*	-0.097 (*-17.1*)*	-0.047 (*-4.1*)*
Personal Travel Cost (st.dev)	0.385 (*25.2*)*	0.386 (*25.0*)*	0.584 (*37.8*)*	0.472 (*24.3*)*
Systematic heterogeneity in Travel Time				
Voluntary	---	-0.002 (*-0.6*)	---	0.004 (*1.5*)
Helped	---	-0.015 (*-4.5*)*	---	-0.043 (*-10.9*)*
Attributes specific for Public Transport				
ASC	5.516 (*22.5*)*	5.363 (*21.2*)*	4.805 (*12.9*)*	6.200 (*15.6*)*
Gender: male	1.673 (*11.7*)*	1.641 (*11.5*)*	1.389 (*9.2*)*	1.098 (*7.2*)*
Car ownership	-4.424 (*-22.8*)*	-4.409 (*-22.7*)*	-3.487 (*-18.6*)*	-3.347 (*-17.6*)*
Purpose: Work	0.262 (*2.0*)*	0.284 (*2.1*)*	0.313 (*2.3*)*	0.157 (*1.2*)

Voluntary	-0.943 (*-7.0*)*	-0.976 (*-6.1*)*	-1.255 (*-8.8*)*	-0.894 (*-5.3*)*
Helped	0.420 (*3.0*)*	0.652 (*3.8*)*	0.787 (*5.2*)*	0.800 (*4.4*)*
Comfort perception	---	---	-0.408 (*-5.4*)*	-0.746 (*-10.0*)*
Urban diversity	---	---	-6.653 (*-2.1*)*	-7.815 (*-2.5*)*
Nonresidential diversity	---	---	-0.005 (*-5.6*)*	-0.006 (*-5.7*)*
Job accessibility	---	---	0.030 (*6.9*)*	0.019 (*4.4*)*
Attributes specific for Walking				
ASC	5.351 (*19.4*)*	5.179 (*17.9*)*	3.370 (*7.1*)*	4.586 (*9.6*)*
Gender: male	1.105 (*6.1*)*	1.078 (*5.9*)*	1.056 (*5.6*)*	0.690 (*3.6*)*
Car ownership	-2.699 (*-11.8*)*	-2.687 (*-11.7*)*	-1.699 (*-7.3*)*	-1.795 (*-7.7*)*
Purpose: Work	-0.892 (*-5.3*)*	-0.882 (*-5.2*)*	-0.827 (*-4.8*)*	-0.846 (*-4.9*)*
Voluntary	-0.263 (*-1.6*)	-0.319 (*-1.6*)	-0.434 (*-2.5*)*	0.033 (*0.2*)
Helped	-0.045 (*-0.3*)	0.248 (*1.2*)	0.432 (*2.4*)*	0.495 (*2.3*)*
Comfort perception	---	---	0.181 (*1.7*)	-0.034 (*-0.3*)
Urban diversity	---	---	-0.113 (*-2.9*)*	-0.116 (*-3.0*)*
Nonresidential diversity	---	---	-0.004 (*-3.3*)*	-0.004 (*-3.7*)*
Job accessibility	---	---	0.027 (*5.0*)*	0.007 (*1.4*)
Measures of fit				
Log-likelihood	-2068.1	-2067.0	-1971.4	-1959.4
$\rho^2(0)$	0.495	0.496	0.519	0.522

**Significant at 95% level of confidence*

Table 2 – DCMs results (N=12526 observations)

The first model (**ML1**) is the simplest specification, as it does not include the effect of the LVs. The LOS parameter signs are correct (consistent with microeconomic theory): lower for car mode. But when authors have examined the database it was found that many individuals made economically inconsistent choices, and that this behavior can only be explained through psycho-sociological aspects. For this reason is important to include on the simplest model the *social influence* variables (‘voluntary’ and ‘helped’): not significant in walking, but the public transport ‘voluntary’ and ‘helped’ constants presents a high significance of its parameters. There is an influence between ‘travel time’ and ‘voluntary’. It is quite trivial, because when I do more activities, I have less time, so the introduction of voluntary activity explains a part of the history explained by time variable. On the other hand, ‘travel time’ and ‘helped’ have also little colinearity because if I have some help or not in child care, I have more or less free time, respectively. At first ‘voluntary’ reduces the probability of choosing public transport, and ‘helped’ increases it. With this result is difficult to know if to be a ‘voluntary’/’helped’ person implies low/high probability of choosing public transport, respectively; or vice versa. The following model results will solve this hesitation.

Other study with the same database (Di Ciommo et al., 2014) justifies several interactions among some respondents’ attributes and decision makers’ characteristics (e.g. travel time), that allow accounting for systematic heterogeneity in the individual preference. The second model (**ML2**) includes this effect, but only on *social influence* variables (‘voluntary’ and ‘helped’) to facilitate the comparisons. Due to the colinearity between *social influence* variables and ‘total travel time’, the constant associate with travel time reduces its statistical significance. The systematic heterogeneity (i.e. taste) in travel time for ‘helped’ is negative. By that is meant that if I have some help for house and child care, I have some time that I can use for going to work by public transport; not for utilitarian reasons but I have no choice. The rest of the parameters do not change much their specifications.

The third model (**HM1**) is based on ML1, but is a hybrid model including the four LVs resulting from the MIMIC model above. When psychological and land use factors are added to the discrete choice model framework, models improve their fitness and statistical significance (Yoon and Goulias, 2010; Moore et al., 2013). With the inclusion of these LVs, it can be seen that mode specific constants (*ASC*) reduce their relative importance and statistical significance, and that LVs are statistically significant (except 'comfort perception' towards public transport for walking users). Apparently, this is showing that the observed behavior is not only ruled by the maximum utility criterion, but also by a strong psychological-environment concept, that develops without the mediation of cognitive processes during choice. It is probable that this fact was hidden behind the modal constants in the ML1 and ML2 models. Car users have more positive 'comfort perception' towards public transport to the users of bus, metro, etc. This statement corroborates the previous conclusion about having people who go to work by public transport not for utilitarian reasons, but sometimes they have no choice. Regarding environmental LVs, the more diversity place of residence is, more difficult is the use of public transport or walking. On the other hand, public transport and walking choices increase with the 'job accessibility'. The four LVs were measured by, among others, 'gender' variable, decreasing its significance in this hybrid model. And again, the *social influence* variables are significant (even for walking mode).

Model performance increases when LVs are properly included. Indeed, according to the LR-test last model (**HM2**), which allows for systematic taste variations in travel time for the *social influence* variables, is easily the best fit. It must be noted that, except travel time parameter, variables with the highest *t*-values are precisely those associated with a psycho-sociological construct: *systematic heterogeneity in travel time* for 'helped' respondents and 'comfort perception' for public transport users. As it occurred at ML2, the taste variation for 'voluntary' variable does not significant; and the *environmental* importance during mode choice is not affected.

5. CONCLUSION

There are a lot of perspectives to testing the effect of transport policy measures to achieve a modal shift: econometric and psychological theories, statistical methodologies, processing or not of social influence or/and land use variables, use of a database coming from a longitudinal or a cross sectional survey, etcetera. Moreover, during the last 40 years a great many case studies on travel behaviour have been analysed using some of the possible combinations of the perspectives above. Therefore, it is very difficult to find the midpoint of the most accurate and complete research in understanding the factors that explain modal choices to minimize negative externalities of car-based trips for a specific case study.

Using as a starting point the principle that the choice of a transport mode is the *key role* played by public transport planners in policy making (Ortúzar and Willumsen, 2011), this paper develops a sequential hybrid DCM methodology to analyze the effect of the transport policies introduced in Madrid during the last four years, in the context of the world economic crisis. This paper does not only to provide a framework on how to integrate some research topics already developed in previous studies and emerging methods on travel behavior, when assessing the introduction of any transport policy, but also to add new evidences from a Spanish perspective to the research debate on how to change patterns of urban mobility. The data used comes from a two panel smartphone-based survey (n=255 and 190 respondents, respectively) of Madrid.

On the basis of this analysis of an urban modal choice case, it is possible to affirm that hybrid DCMs are clearly superior in fit to an equivalent DCM without incorporating LVs. The LVs approach used in the hybrid DCMs is an advance over the typical dummy variables approach because it does not depend on the subjective definition of boundaries for generation of the values of 0 to 1 required in this last case. The consideration of LVs in a DCM allows capturing the real importance that the LOS and cost variables have on the individual decision-making process.

The signs of LOS attributes are consistent with microeconomic theory: time and cost parameters represent a disutility to individuals. Respect of environmental LVs included, higher diversity places of residence (higher mixture of land uses and street density) are associated more car use. On the other hand, public transport and walking choices increase with places close to job accessibility. Regarding the psychological perception towards public transport, car users have more positive perception to the users of public transport. Finally, with the inclusion for systematic taste variations in travel time for two novel social influence variables (voluntary and helped), models improved their fitness and statistical significance. It is recommended that future research uses a survey design to fully capture all the LVs' variability on travel behavior; not only other environmental (density and design) and psychological factors (e.g. expectancy and norms), but also other *social influence* aspects (e.g. number of friends at Facebook, Twitter and LinkedIn) very correlated with the free time to travel.

The results confirm previous authors' findings (Comendador et al., 2014b and Di Ciommo et al., 2014) – in a more general framework, certainly- on the theory that transport policy actions are more likely to be effective when car use has been first disrupted, because: (1) in areas with higher job accessibility, public transport is more likely to be used since the use of car is more restricted; (2) many people using public transport on their way to work do not do it for utilitarian reasons, but because no other choice is available; and (3) the more things to do for a whole day (household responsibilities, voluntary activities, etcetera), the more use of car.

ACKNOWLEDGEMENTS

The authors wish to thank the Secretaría de Estado de Ciencia e Innovación, which has funded the HABIT Project, (R+D Programme 2011-2013). They are also very grateful to Juan Carlos García Palomares and Srinath Ravulaparthy for their support in preparing the land use variables based on Geographical Information Systems software. The authors would also like to thank Rico Wittwer, Floridea Di Ciommo, Elisabetta Cherchi and Juan de Dios Ortúzar who provided methodological assistance.

REFERENCES

ANDERSON, J. C., and GERBING, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411-423.

BEN-AKIVA, M., WALKER, J., BERNARDINO, A. T., GOPINATH, D. A., MORIKAWA, T., and POLYDOROPOULOU, A. (2002). Integration of choice and latent variable models. *Perpetual motion: Travel behaviour research opportunities and application challenges*, 431-470.

BOLLEN, K.A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons, Inc.

BYRNE, B. (2001). *Structural equation modelling with AMOS – Basic concepts, applications, and programming*. East Sussex: Psychology Press.

CERVERO, R. and KOCKELMAN, K., (1997). Travel Demand and the 3Ds: Density, Diversity, and Design. *Transportation Research-Part D*, 2, 199-219.

CHERCHI, E., and ORTÚZAR, J. de D. (2011). On the use of mixed RP/SP models in prediction: Accounting for systematic and random taste heterogeneity. *Transportation Science*, 45, 98-108.

COMENDADOR, J., DI CIOMMO, F. and LÓPEZ-LAMBAS, M.E. (2014a). Assessing Subjectivity in Urban Mobility through a Revealed Preferences survey of New Stations Users. Submitted for publication in Case Studies on Transport Policy. Presented at 18th PANAM Conference.

COMENDADOR, J., WITTWER, R., LÓPEZ-LAMBAS, M.E. and MONZÓN, A.

(2014b) Understanding Travel Behavior to change Intentions towards Private Car and Public Transport. Submitted for publication in Transportation Research Part A.

COMENDADOR, J., LÓPEZ-LAMBAS, M.E. and MONZÓN, A. (2014c). Urban built environment analysis: Evidence from a mobility survey in Madrid. Presented at 11th CIT Conference.

DE ABREU E SILVA, J., MORENCY, C., and GOULIAS, K. G. (2012). Using structural equations modeling to unravel the influence of land use patterns on travel behavior of workers in Montreal. *Transportation Research Part A*, 4, 1252-1264.

DI CIOMMO, F., COMENDADOR, J., LÓPEZ-LAMBAS, M.E., CHERCHI, E., and ORTÚZAR, J. DE D. (2014). Exploring the role of social influence variables on travel behaviour. Submitted for publication in Transportation.

EWING, R. and CERVERO, R. (2010). Travel and the built environment. *Journal of the American Planning Association*, 76, 265-294.

GALDAMES, C., TUDELA, A. and CARRASCO, J. (2011). Exploring the Role of Psychological Factors in Mode Choice Models by a Latent Variables Approach. *Transportation Research Record*, 2230, 68–74.

GARCÍA-PALOMARES J. C., GUTIÉRREZ J., and CARDOZO O. D. (2013). Walking accessibility to public transport: an analysis based on microdata and GIS. *Environment and Planning B: Planning and Design 40*, 1087 – 1102.

GOLOB, T. F. (2003). Structural equation modelling for travel behaviour research. *Transportation Research Part B*, 37, 1–25.

GEURS K. T., and VAN WEE B. (2004). Accessibility evaluation of land-use and transport strategies: review and research directions. Journal of Transport Geography, 12, 127-140.

MOORE, J., CARRASCO, J.A., and TUDELA, A. (2013). Exploring the links between personal networks, time use, and the spatial distribution of social contacts. *Transportation*, 40, 773-788.

ORTUZAR, J. de D., and WILLUMSEN, L. G. (2011). *Modelling Transport.* Fourth Edition, John Wiley and Sons, Chichester.

RAVEAU, S., ÁLVAREZ-DAZIANO, R., YÁÑEZ, M. F., BOLDUC, D., and ORTÚZAR, J. D. (2010). Sequential and Simultaneous Estimation of Hybrid Discrete Choice Models. *Transportation Research Record*, 2156, 131-139.

SCHUMACKER, R.E., and LOMAX, R.G. (2004). *A beginner's guide to structural* *equation modeling*. 2nd Edition Lawrence Erlbaum Associates, Inc.; Mahwah, New Jersey.

TRAIN, K.E., (2009). Discrete Choice Methods with Simulation, 2nd edition. Cambridge University Press, Cambridge.

WANG, D., and LAW, F. Y. T. (2007). Impacts of Information and Communication Technologies (ICT) on time use and travel behavior: a structural equations analysis. *Transportation*, 34, 513-527.

WEIS, C., and AXHAUSEN, K. W., (2009). Induced travel demand: evidence from a pseudo panel data based structural equations model. *Research in Transportation Economics*, 25, 8-18.

YÁÑEZ, M. F., MANSILLA, P., and ORTÚZAR, J. de D. (2010a). The Santiago Panel: measuring the effects of implementing Transantiago. *Transportation*, 37, 125-149.

YÁÑEZ, M., RAVEAU, S., and ORTÚZAR, J. de D. (2010b). Inclusion of latent variables in mixed logit models: Modeling and forecasting, *Transportation Research Part A*, 44, 744–753.

YOON, S. Y., and GOULIAS, K. (2010). Constraint-based assessment of intra-household bargaining on time allocation to activities and travel using individual accessibility measures. Proceedings of the 89th Annual Meeting of the TRB, 2010.