

10. How urban environment affects travel behaviour: integrated choice and latent variable model for travel schedules

**Lissy La Paix, Michel Bierlaire,
Elisabetta Cherchi and Andrés Monzón**

1. INTRODUCTION

Since 1954 (Mitchell and Rapkin 1954) many researchers have studied the important but highly complex relation between travel demand and land use. One of the major problems discussed in the past literature was the relative importance of urban form characteristics versus socioeconomic characteristics to explain trip frequency. The conclusion from these early studies (Kitamura et al. 1997; Hanson 1982) was that the total number of trips is largely determined by demographic and socioeconomic factors but it is not strongly associated with land use characteristics.

Trip frequency has been the major focus also in the recent literature, mainly with reference to specific purposes (Limanond and Niemeier 2004) or to specific categories of people (Schmöcker et al. 2005; Roorda et al. 2009; Morency et al. 2011; Paez et al. 2007). These studies mainly found that trip frequency decreases with age, but this behaviour exhibits a large degree of variability depending on spatial analysis.

A few authors have studied the relationship between land use characteristics and tours, but they focused only on specific types of tours such as shopping tours (Agyemang-Duah et al. 1995) or specific types of activities (Naess 2006). The studies found longer total travelling distances among suburbanites than among inner-city residents, and fewer trips by households located in the Central Business District (CBD) than those living in the suburbs. Bhat (1999) used instead land use variables to model specifically the number of stops in the tours. More recently, Shiftan (2000, 2008) analyzed the effect of land use policies using an activity-based approach and developed an integrated residential choice

activity-based framework, which incorporates dwelling type and urban structure variables.

In this chapter we aim to study the effect of land use and socioeconomic characteristics in the discrete choice among different tour structures, where tour structures are defined by the main activities performed in the tour and the number of stops realized during the tour for other purposes than the main activity. At the same time, following the recent literature on the effect of latent factors on the discrete choice (Ben-Akiva et al. 1999; Walker and Ben-Akiva 2002; Atasoy et al. 2010; Hurtubia et al. 2010; Yáñez et al. 2010) we believe that the choice among type of tours can also be affected by unobservable attitude towards travel that is not reflected in the explanatory variables. To study this effect we use a hybrid model where the latent variable measures the propensity to travel of each individual, while the discrete choice is the type of tours. Propensity to travel (PT) refers to the natural or acquired tendency, inclination, or habit in a person to carry out a number of trips during a day. This might be thought of as a general or inherent willingness to travel, due to individual characteristics and commitments, which influence how frequently a person travels (daily trip frequency).¹ Following this definition, in our chapter, the PT is defined as the individual tendency to travel, and it is measured by the daily trip frequency. As for the discrete choice, we use the type of tour because it incorporates an explicit representation of temporal-spatial constraints among activity stops within a tour. Both the latent variable and the tour choice are a function of the individual socioeconomic and land use characteristics, but the tour choice is also a function of travel characteristics and a function of the latent variable. In this way we are able to measure the effect of land use and socioeconomic characteristics on the choice of tours directly and indirectly through the propensity to travel.

Accounting for latent factors explaining attitude toward lifestyle is important to control for self-selection, because attitude allows accounting for “the tendency of people to choose locations based on their travel abilities, needs and preferences” (Litman 2005). Only few authors have analyzed the individual activity-travel decision process including psychological factors. Handy et al. (2005) used a quasi-longitudinal design to account for the role of travel and neighbourhood preferences in explaining this relationship between neighbourhood characteristics and travel behaviour. They found that differences in travel behaviour between suburban and traditional neighbourhoods are largely explained by attitudes. They also found a self-selection issue in vehicle miles driven per week between different neighbourhoods. Walker and Li (2007) and Kitrinou et al. (2010) used instead an integrated choice and latent variable model for the residential location choice. In particular, Walker and Li (2007) used a latent

class choice model, where the latent model represents two latent life-styles (auto-oriented households and transit-oriented households). They found an explicit connection between demographics, lifestyle and residential location. On the other hand, Kitrinou et al. (2010) analyzed residential relocation decision in island areas and found that this decision is significantly affected by the following latent variables: satisfaction of the current commuting trip, geographical mobility, satisfaction with the current residential area and perceptions about the quality of life.

As far as we know, our research is the first attempt to study the relationship between neighbourhood type and tour structure with hybrid choice models. Our study was carried out in the context of Madrid, where data on trips and activities performed by a sample of families living in three neighbourhoods with different characteristics (in terms of urban density, land use and transport supply) were available. From the two approaches now available on the estimation of hybrid choice models, we used the advantage of simultaneous estimation, because it leads to more efficient estimates (Ben-Akiva et al. 1999). Other studies compare simultaneous and sequential estimations. (Bolduc and Álvarez-Daziano 2009; Bolduc et al. 2008; Ben-Akiva et al. 2002b; Raveau et al. 2010).

The chapter is organized as follows: Section 2 is a description of the general framework of integrated discrete choice and latent variable model and the model specification used for this specific application. Section 3 is a description of the case study of Madrid and the variables used for the model estimation. Section 4 includes the empirical results of the hybrid choice model estimated. Section 5 is the conclusion of the chapter, and there we also identify future research.

2. INTEGRATED CHOICE AND LATENT VARIABLE MODEL

The present work follows the general framework and methodology provided by Ben-Akiva et al. (2002a), and Walker and Ben-Akiva (2002), for incorporating latent variables into choice models via the integration of choice and latent variable models. This general framework is based on the assumption that the individual utility for each alternative depends on a vector of measurable attributes, as with any random utility choice model, but it also depends on unknown (latent) variables. The modeller can only measure the manifestations of the latent constructs through indicators, which allow the identification of the latent constructs.

The integrated choice and latent variable structure explicitly models the latent variables that influence the choice process. To specify both

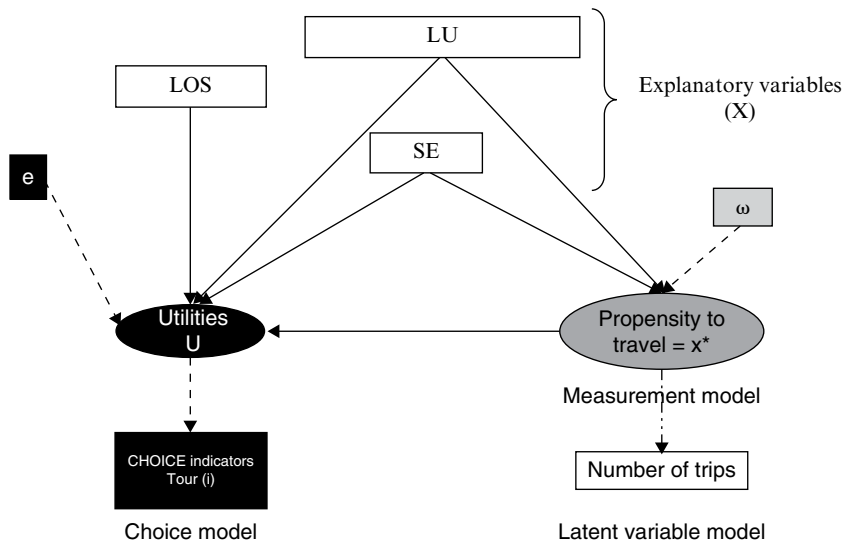


Figure 10.1 Framework for the integrated choice and latent variable model

the discrete choice and the latent variables, two types of equations are required: a measurement equation that links the unobservable latent variable to its observable indicator, and a structural equation that links the observable to the latent variables and models the behavioural process by which the latent variables are formed.

Figure 10.1 shows the framework of the integrated latent variable and choice model we used in our formulation. As in the typical general framework of the hybrid models, the right portion of Figure 10.1 represents the latent variable model (PT) while the left part represents the discrete choice model. Terms in ellipse represent unobservable (latent) constructs, while those in rectangles represent observable variables. We assumed that socioeconomic (SE) and neighbourhood characteristics (LU) play a role in both the latent and discrete choice model, while the travel attributes (LOS) affect only the choice of the type of tour, because this information only exists once a tour has started. The solid arrow that links travel attributes and neighbourhood characteristics to the observed decision of type of tour represents the structural equations that control the decision making process. Similarly, another solid arrow links the PT with the observed decision. The dashed arrows represent instead the measurement equations.

2.1 Propensity to travel latent variable model

Following the above general framework, for the latent variable model we defined a structural equation that links the propensity to travel to the SE and LU characteristics and a measurement equation that links the propensity to travel to its indicator, that is, the frequency of trips.

The structural equation is specified as follows:

$$PT_n = \sum_k \lambda_k^s S_{kn} + \sum_m \lambda_m^z Z_{mn} + \omega_n \tag{10.1}$$

Where PT_n is the propensity to travel for individual n . S_n is a vector of SE characteristics with k elements, Z_n is a vector of neighbourhood attributes with m elements, λ^s and λ^z are two vectors of parameters associated with the SE and LU characteristics respectively, while ω_n is the error term distributed Normal with zero mean and standard deviation σ_ω .

The measurement equation is specified as follows:

$$I_n = \alpha PT_n + v_n \tag{10.2}$$

where I_n is the indicator of the latent variable, in our case the *number of trips*, α is the associated parameter to be estimated and v_n is the error term, Normal distributed with zero mean and standard deviation σ_v .

Since both the latent variable of (PT_n) and its indicator of (I_n) are assumed to be normally and independently distributed, their distributions (f_{PT} and f_I) indicator are given respectively by:

$$f_{PT}(PT_n | \sigma_\omega) = \frac{1}{\sigma_\omega} \phi \left(\frac{PT_n - \sum_k \lambda_k^s S_{nk} + \sum_m \lambda_m^z Z_{nm}}{\sigma_\omega} \right) \tag{10.3}$$

$$f_I(I_n | PT_n; \alpha, \sigma_v) = \frac{1}{\sigma_v} \phi \left(\frac{I_n - \alpha PT_n}{\sigma_v} \right) \tag{10.4}$$

2.2 Discrete type of tours choice model

Analogously to the latent variable model, for the discrete choice model we need the distribution of the utilities (U_{jn}) that individual n associates with each type of tour j and a measurement equation to identify the choice. The structural equation is specified as follows:

$$U_{jn} = ASC_j + \sum_i \beta_i^s S_{in} + \sum_r \beta_r^z Z_{nr} + \beta_j^{PT} PT_n + \epsilon_{jn} \tag{10.5}$$

where the utility function is expressed as a function of a vector of socioeconomic characteristics (S_n), a vector of neighbourhood characteristics (Z_n), and the propensity to travel (PT_n) of each individual n . Note that the effect of the PT is estimated specifically for each tour ($N-1$ alternatives). Finally, ϵ_{jn} is, as always, the error term extreme value distributed with mean zero and standard deviation σ_ϵ .

Note that the discrete choice and the latent variable models can include different attributes, hence the vectors S_n and Z_n can be different between equations (10.1) and (10.5). The parameters associated with these attributes in the discrete choice are of course different from those in the latent variable.

In our model, tours are defined as a sequence of trip segments in the full day activity pattern. There is one primary activity and one or more secondary activities, carried out at intermediate stops. Although the list of activities reported in the survey included many categories (home, work, work-related, studies at university or school, pick-up/drop-off to someone, shopping, leisure-social, entertaining trips and sports – and others), for estimation, purposes were grouped into three activities (home, work/study and shopping/others). However, in our formulation, we also distinguished tours based on the number of stops; then, the discrete alternatives j are represented by the following types of tours:

$j = 1$ is HOME: no trips during the day.

$j = 2$ is a simple tour (without any additional stops) between home and work/school and back home. We label it (HWH).

$j = 3$ is a work/school tour with at least one additional stop to perform another activity. We label it (HWHs).

$j = 4$ is a simple tour from home to any other activities other than work or school and back home. We label it (HOH).

$j = 5$ is a tour with main purpose any activities other than work or school and at least one additional stop for another activity. We label it (HOHs).

The choices are mutually exclusive; therefore, the model involves five alternatives mutually independent. Availability of alternatives was based on age and occupation. Home-Work-Home tours with (HWHs) and without (HWH) intermediate stops are set not available for people older than 65 years old, because this is the age of retirement. And similarly, Home-Other-Home tours with (HOHs) and without (HOH) intermediate stops are set not available for students and workers during working time.

The measurement equation is defined, as in the typical discrete choice models, by an indicator (d_{jn}) that takes value 1 if the alternative chosen has the highest utility among all the alternatives available in the choice set of each individual:

$$d_{jn} = \begin{cases} 1 & \text{if } U_{jn} > U_{in}, \forall i \neq j; (i, j) \in C_n \\ 0 & \text{in other case} \end{cases} \tag{10.6}$$

where C_n is the choice set of the individual n .

Hence the probability of individual n choosing the alternative j is the probability of choosing the alternative conditional on the observed and unobserved variables, given by:

$$P(j_n | S_n, Z_n, LOS_{nj}, PT_n; \beta, \sigma_\varepsilon) = Prob[U_{jn} \geq U_{in}, j \forall \in C_n] \tag{10.7}$$

Since we assumed the error term is distributed extreme value type 1, the probability of an individual n choosing alternative j , conditional on the unknown propensity to travel (PT), is given by:

$$P_{jn} = \frac{e^{V_{nj}(PT_n)}}{\sum_i e^{V_{ni}(PT_n)}} \tag{10.8}$$

2.3 Estimation of the hybrid choice model

In the simultaneous estimation, the latent variable and the discrete choice are jointly estimated, where the joint probability is given by the following expression:

$$\begin{aligned} &P(j, I_n | S_n, Z_n, LOS_{jn}; \beta, \alpha, \lambda, \sigma_\varepsilon, \sigma_v, \sigma_\omega) \\ &= \int_{PT}^\alpha P(j | S_n, Z_n, LOS_{nj}, PT_n; \beta, \sigma_\varepsilon) f_{PT}(PT_n | S_n, Z_n; \lambda, \sigma_\omega) f_I(I_n | PT_n; \alpha, \sigma_v) dPT \end{aligned} \tag{10.9}$$

Note that the distribution of the indicator is included to improve the accuracy of estimates of the structural parameters as well as to allow for their identification purposes. The maximum likelihood is obtained, as always, from maximizing the logarithm of the likelihood function (\mathcal{L}) over the unknown parameters:

$$\mathcal{L} = \sum_n \sum_{i \in C_n} d_{jn} \log P(j, I_n | S_n, Z_n, LOS_{nj}; \beta, \alpha, \lambda, \sigma_\varepsilon, \sigma_v, \sigma_\omega) \tag{10.10}$$

3. EMPIRICAL APPLICATION

3.1 Data characteristics

The data used in this work comes from a survey conducted originally with the aim of analyzing the influence of the type of questionnaire (activity-based against travel-based) on the mobility patterns (Madrigal and Monzón 2007). The survey was conducted first in 2006 and then repeated (although not with the same individuals) in 2007. One of the reasons for repeating the survey was the need to enlarge the sample and to include a new type of neighbourhood. In 2006, the survey included two zones of Madrid: the CBD and an urban area; in 2007 a suburban area was added. The three neighbourhoods included in the survey have the following characteristics:

- **CBD:** this area, called Chamberí, corresponds to one of the 22 neighbourhoods of the Central Business District of Madrid. It is a traditional neighbourhood where several historic buildings are located and where people live mainly in apartments. It is characterised by good transit (bus and metro) and rail services and by a gross income level that ranks fourth amongst the neighbourhoods of Madrid City. In 2004 the income of Chamberí was also 40 per cent higher than the mean of the Region of Madrid. The information is not available for 2007 and 2008, but the income distribution has not changed recently.
- **Urban:** this area, called Pozuelo de Alarcón, is located 15 km west of Madrid CBD but it is inside Madrid City. This is a car-oriented neighbourhood, where the supply of public transport services is limited. Urban dwellers tend to live in single family houses or detached houses. Average income level of Pozuelo ranks the highest amongst the municipalities of the Region of Madrid. It was 66 per cent higher than the mean of the Region of Madrid in 2004.
- **Suburban:** this area, called Algete, is located 30 km north-east of the Madrid CBD, in the Metropolitan Ring. This district has a lower available gross income and fewer transit services than the other two selected neighbourhoods. The average income of Algete is lower than that of Pozuelo and 17 per cent higher than the mean of the Region of Madrid in 2004.

Data were collected during a working day using a travel diary. The sample includes 345 households and 943 individuals. All individuals older than 4 years were interviewed. The complete sequences of trips made in a

Table 10.1 Definition of explanatory variables

Variable	Parameters in the latent variable	Definition	Parameters in the discrete choice model
Socio-economic characteristics (S_n)			
4–13 years old		Reference	
14–21 years old	$\lambda_{Age1421}$	Yes = 1; otherwise = 0	$\beta_{Age14-21}$
22–39 years old	$\lambda_{Age2239}$	Yes = 1; otherwise = 0	$\beta_{Age22-39}$
40–49 years old		Yes = 1; otherwise = 0	$\beta_{Age40-49}$
40–64 years old	$\lambda_{Age40-64}$	Yes = 1; otherwise = 0	
50–64 years old		Yes = 1; otherwise = 0	$\beta_{Age50-64}$
More than 64 years		Reference	
Female married	$\lambda_{FemaleMarried}$	Yes = 1; otherwise = 0	
Worker	λ_{Worker}	Yes = 1; otherwise = 0	β_{Worker}
Own car	$\lambda_{Own\ car}$	At least 1 car: Yes = 1 ; No car = 0	
Adult \times Child presence		Yes = 1; otherwise = 0	$\beta_{AdultChild}$
Neighbourhood Characteristics (Z_n)			
Condominium	$\lambda_{Condominium}$	Yes = 1; otherwise = 0	
Single family	$\lambda_{Single\ family}$	Yes = 1; otherwise = 0	
Other dwelling		Yes = 1; otherwise = 0	$\beta_{Other\ dwelling}$
Worker density		Ratio of worker density at origin/destination	$\beta_{RatioWorkers_{ob}}$
Commercial land-use		Ratio of urban retails origin/destination	β_{Comm}
PT-600		Number of rail and metro stations in the residence area. Calculated for a 600 meters radius.	β_{PT-600}

day were then collected, with all their characteristics. Socioeconomic characteristics were also gathered from each individual of the interviewed families. Then, household location, origin and destinations from the survey were geocoded and integrated into a geographic information system that included public transport network and street data.²

3.2 Variable description

Table 10.1 reports the list of explanatory variables used in each model (i.e. in the tour choice and propensity to travel models) along with an explanation of how they have been coded. The column “definition” includes the name of the associated parameters as shown in the results.

It is important to mention that many attributes were tested but only those listed above were significant in our models. A detailed description of all the variables built and tested (also in other models) can be found in La Paix et al. (2010, 2012), and La Paix (2010). In particular, to describe the characteristics of the neighbourhoods, several variables were tested, such as dwelling type, population density in the municipality, mixed land use (percentage of land use in the zone of origin and/or destination for residential, commercial or industrial purposes), balance between residence and employment and measures of accessibility. All the available variables were tested in our specifications, as well as several combinations of variables. The final selection of the neighbourhood characteristic was based on statistical grounds. It is also important to mention that correlation among the neighbourhood characteristics used in our model is low. Socioeconomic and neighbourhood characteristics are included in both latent and discrete choice.

4. RESULTS

Table 10.2 shows the results obtained using the specifications described in section 2. In particular, the first two columns include the parameters estimated with the hybrid model, while the last two columns have, for comparison, the parameters estimated with a simple multinomial logit (MNL) without the latent variable. This is a restricted version of the hybrid choice model (HCM).

Following the description in section 2, the parameters estimated in the HCM include: 17 parameters (β) for the SE, LU attributes and the PT variables in the discrete choice model; 9 parameters (λ) for SE and LU characteristics in the latent variable; 4 alternative specific constants (*ASC*) in the discrete choice model and 1 constant (*Mean PT*) for the latent variable; 1 parameter (α) associated with the latent variable in the measurement equation (this parameter was constrained to one, for identification purposes) and 2 standard deviations respectively for the latent variable (σ_w) and for the indicator (σ_v). Parameters are estimated using an extended version of the software package Biogeme (Bierlaire 2003). For further details on the software see Bierlaire and Fetiariison (2009).

First of all we note that both socioeconomic and neighbourhood characteristics are significant in explaining the propensity to travel as well as the discrete choice among tours. Also the propensity to travel, included in all the tours except work/study based tours without additional stops and home, is highly significant, positive and significantly different among the three types of tours. This result confirms our a priori hypothesis about individual behaviour, that higher propensity to travel leads to more

Table 10.2 Model estimation results

Name	Affected alternative	Hybrid choice model		Multinomial logit	
		Value	Robust T-test	Value	Robust T-test
Discrete choice					
ASC_{HOME}	HOME	-6.21	-3.18**	-5.80	-2.71**
ASC_{HWHs}	HWHs	-26.6	-4.14**	-3.99	-3.16**
ASC_{HOH}	HOH	-6.43	-3.17**	12.2	1.74
ASC_{HOHs}	HOHs	-22.5	-8.50**	-5.98	-2.80**
$\beta_{Age1421}$	HWH, HWHs, HOH, HOHs	0.80	1.36	1.94	4.23**
$\beta_{Age2239}$	HWH, HWHs	-0.77	-1.58	0.22	0.85
$\beta_{Age2239}$	HOHs	-0.72	-0.88	-0.42	-0.77
$\beta_{Age4049}$	HWH, HWHs	-1.51	-3.09**	-0.58	-2.17**
$\beta_{Age5064}$	HWH	-1.79	-3.78**	-0.69	-3.31**
$\beta_{Age5064}$	HOH, HOHs	0.56	1.26	0.76	2.58**
β_{Worker}	HWH, HWHs	1.01	4.33**	1.07	4.82**
$\beta_{Adult Child}$	HWHs, HOHs	1.53	2.69**	0.65	3.23**
$\beta_{Other dwelling}$	HOHs	1.96	2.25**	-0.22	-0.20
β_{PT600}	HOHs	-1.94	-1.50	-1.73	-2.62**
$\beta_{RatioWorkers_{od}}$	HWH	-4.80	-2.51**	-5.59	-2.58**
$\beta_{RatioWorkers_{od}}$	HWHs	1.77	0.31	-2.95	-1.40
$\beta_{RatioWorkers_{od}}$	HOH	-6.16	-2.76**	-18.4	-2.76**
$\beta_{Comm_{od}}$	HWHs, HOHs	-0.13	-1.10	-0.02	-0.40
$\beta_{PT_{inw}}$	HWHs	7.52	12.14**		
$\beta_{PT_{non}}$	HOH	3.93	6.74**		
$\beta_{PT_{non}}$	HOHs	7.31	11.06**		
Latent Variable					
Mean PT		1.83	262.72**		
$\lambda_{Age1421}$		0.34	43.77**		
$\lambda_{Age2239}$		0.01	1.67		
$\lambda_{Age4064}$		0.33	48.43**		
$\lambda_{FemaleMarried}$		0.34	73.37**		
λ_{Worker}		0.34	89.07**		
$\lambda_{Own car}$		-0.002	-0.99		
$\lambda_{Single Family}$		-0.34	-126.49**		
$\lambda_{Condominium}$		0.32	53.56**		
$\lambda_{Metro600}$		-0.03	-98.60**		
$\sigma_{\omega_{pr}}$		0.41	284.85**		
σ_{ν_n}		-3.08	-78.75**		
Sample size		890		890	
Final log-likelihood:		-1537.44		-921.299	
Rho bar for the initial model		0.507		0.129	

Note: (**) significant at 95%.

Table 10.3 *Marginal utilities accounting for the effect of the propensity to travel*

Variable	HWH	HWHs	HOH	HOHs
Age 14–21	0.801	3.394	2.156	3.321
Age 22–39	-0.771	-0.682	0.046	-0.630
Age 40–49	-1.512	0.999	1.312	2.440
Age 50–64	-1.792	2.511	1.876	3.004
Female married		2.527	1.321	2.455
Worker	1.006	3.546	1.327	2.468
Own car		-0.014	-0.007	-0.014
Single family		-2.544	-1.329	-2.472
Condominium		2.409	1.259	2.341
Metro 600		-0.224	-0.117	-0.237

intermediate stops and that the propensity to travel has different effects depending on the type (or complexity) of the tour performed and depending on the primary activity of each tour.

In particular the propensity to travel increases with the age between 14 and 64 years, but not linearly. Two of three age parameters in the latent variable models are in fact significantly different among them (t-test for generic parameters higher than 1.96). Age has a non linear effect also on the choice of the type of tours. In fact, if we do not consider the effect of the propensity to travel (as in the MNL model or if we consider only the effect of the DCM in the HCM) the probability to choose any tour is positive for people between 14 and 21 years old, and negative for those between 22 and 49 years old; while for those in the cohort 50–64 it is negative for HWH tours and positive for HOH and HOHs tours. If we correctly consider the effect of age through the propensity to travel (see values reported in Table 10.3), the effect is still highly non linear among age cohorts and for the same cohort among types of tours, but the effect is slightly different, because all the marginal utilities are positive except those for HWH and those of the age cohort 22–39.

Owning a car has a negative sign in the propensity to travel, meaning that it decreases the propensity to travel. This is somehow a counterintuitive effect but the variable is not significant at 95 per cent. The propensity to travel instead is positive for a worker and married female, of which 48 per cent are housewives. However, because of the specific effect of the latent variable on the type of tours, workers have a high probability to make complex tours around the main working activities, less probability to realize tours for purposes different from work and even less probability to carry out simple working tours; and analogously for married females.

Interestingly, the number of metro stations in a radius of 600 metres from the living place does not increase the propensity to travel and does not even increase the probability of realizing non-working tours. All the marginal utilities are negative either when only considering the effect in the discrete choice or through the latent propensity to travel. Regarding dwelling type, people living in a “single family” unit have lower probability of undertaking tours for work/study with more than one stop and in general tours with other purposes. Single family houses are usually bigger houses, with more facilities, so people are more likely to stay at home than those living in apartments, terraced or detached units. In contrast, people living in “condominium” have higher probability of undertaking multistage tours. Condominiums are dwelling types mostly frequent in the CBD, and more facilities are located around those areas.

Another important neighbourhood characteristic is the opportunity to perform activities (either work or other non work activities) in the neighbourhood where the person lives compared to the place of destination. Our results show that more commercial activities in the zone of origin than at destination decreases the probability to perform complex tours, i.e. with more than one stop (the variable called Commercial Land Use, which is the ratio of commercial land at origin and destination, has a negative parameter for the tours with stops, although the effect is significant at 86 per cent). On the other hand more service employment in the zone of origin than at destination decreases the probability to perform simple tours, i.e. with only one stop. The “RatioWorkers” variable, which is the ratio of the service employment between origin and destination, has a negative significant effect on simple tours.

Finally, it is worth mentioning that some variables have opposite effects in the MNL and HCM, but in these cases they are not significant.

Finally, Table 10.4 reports the aggregate elasticity for some of the attributes. Elasticities are computed for each type of tour, and for both the hybrid and the simple discrete choice model. As expected, the ratio

Table 10.4 Average elasticity for ratio of workers and metro station

Type of tour	Ratio of workers (O/D)		Metro station	
	HCM	DCM only	HCM	DCM only
Home	3.700	5.953	0.077	0.010
HWH	-0.915	0.488	0.069	0.008
HWHs	2.143	2.778	-0.097	0.008
HOH	-0.591	-3.905	-0.020	0.004
HOHs	0.400	2.080	-0.081	-0.031

of workers between origin and destination has the highest elasticity in both HCM and MNL models, with the exception of HWH tours. This is because the attribute “RatioWorkers” is not included in the latent model, so its effect is similar for both the hybrid and the simple multinomial logit model. Note also that, although the sign of the “RatioWorkers” coefficients is negative in all the alternatives except for HWHs in the hybrid model, some direct elasticity is positive (namely the elasticity for HWHs in the HCM and the elasticities for HWH and HWHs in the DCM model). This is because the ratio of workers is a socioeconomic attribute and it does not vary among tours, so a variation of the ratio of works directly affects all the three alternatives.

The number of metro stations instead is included in both the latent variable (alone) and in the discrete choice among tours (summed to the number of bus stops), so the elasticity is of course different in the two model formulations. In the MNL, the attribute is included only in the HOHs tour, so the elasticity is negative for the HOHs tours (because the attribute has a negative coefficient) and it is positive in all the other alternatives. In the HCM instead, the elasticity is affected by the effect of the number of stations in the latent propensity to travel, and it is in fact negative for the three alternatives where the latent variable is included. Moreover, because of the strong effect of the latent propensity to travel, the effect of the number of metro stations in the HCM is also higher (in absolute value) in the HCM than in the MNL model. This reinforces the importance of properly accounting for the latent propensity to travel into the discrete choice among tours.

5. SUMMARY AND FUTURE DIRECTIONS

In this work we studied the effect of land use and socioeconomic characteristics in the choice of tour complexity according to three primary activities: home, work/school, shopping/others; and five types of tour: HOME, HWH, HWHs, HOH and HOHs. A hybrid model was estimated where the latent variable is the propensity to travel and the choice model is the discrete choice among tour structures in terms of type of main activities of the tour and number of stops during the tour for purposes other than the main activity.

Many authors have studied the relationship between urban environment and travel behaviour with different structures of discrete choice models. However, the study of this relationship with hybrid models is new. The empirical results obtained from our case study are also new in this field. This approach has two key advantages. The first is that tour-based

analysis provides a better explanation than trip-based because it takes into account temporal and spatial constraints; and the second is that the integrated choice and latent variable model allows for a more detailed analysis of the behavioural process.

The main motivation to select the propensity to travel as the latent variable is because we believe there could be an unobservable attitude that affects the discrete choice model, but it is not reflected in the explanatory variables. The PT for a single individual measures in fact how frequently s/he travels depending on her/his socioeconomic characteristics and neighbourhood characteristics of the location where s/he lives and performs other activities during the tour. This latent variable also allows accounting for self-selection effects due to the spurious associations attributable to the simultaneous effect of preferences on both the choice of residential location (and thus built environment) and travel behaviour (Handy et al. 2005; Cao et al. 2009).

The results support our hypotheses that PT as a latent variable exists, and it is important to explain the choice among tour complexity. Our results also show that both the propensity to travel and the discrete choice of the type of tours can be explained by observed socioeconomic and neighbourhood characteristics. There are specific effects tested in this empirical study between urban structure variables, such as commercial ratio, neighbourhood type, presence of bus/metro at residence area and so on, and tour complexity according to different primary activities (working, school, shopping and others). It is worth emphasizing also the contribution of the latent propensity to travel in the discrete choice model to better understanding individual decision process. Additionally, the extent to which the propensity to travel influences the decision for tour complexity strongly depends on the flexibility of the individual commitments. For example, low propensity to travel represents the behaviour of people who do not have work/study commitments, or aging people or those who are living in a low density area with no commercial retail outlets. High propensity to travel captures the behaviour of people who have to travel daily for work or study, people who need to travel for household commitments or people who undertake multi-stage tours during a working day.

ACKNOWLEDGMENTS

Research was mainly developed during the stay of the first author at the Transport and Mobility Laboratory at EPFL (Transp-OR). The first author would like to thank Transp-OR staff, for their feedback and useful comments; special thanks go to Bilge Atasoy, for her help in the estimation

process. The authors would also like to thank the Department of Human Geography at the Complutense University of Madrid for providing the transport network and Concepción García and Juan de Dios Ortúzar for their comments on the availability of the tours.

NOTES

1. In the psychological literature “propensity” is often used with specific meanings: as “propensity to trust” others (Mayer et al. 1995), as “risk propensity, i.e. as the tendency of a decision maker to either take or avoid risks” (Sitkin 1992), or as “propensity effect” i.e. as a reversal of the traditional hindsight bias (Roese 2006).
2. The public transport network was facilitated by the Department of Human Geography of Complutense University of Madrid.

REFERENCES

- Agyemang-Duah, K., Anderson, W. and Hall, F. (1995), ‘Trip generation for shopping travel’, *Transportation Research Record*, **1493**, 12–20.
- Atasoy, B., Glerum, A., Hurtubia, R., and Bierlaire, M. (2010), ‘Demand for public transport services: Integrating qualitative and quantitative methods’, *Proceedings of the 10th Swiss Transport Research Conference (STRC)*, September 1–3, 2010.
- Ben-Akiva, M., Mcfadden, D., Gärling, T., Gopinath, D., Walker, J., Bolduc, D., Börsch-Supan, A., Delquié, P., Larichev, O. and Morikawa, T. (1999), ‘Extended framework for modeling choice behaviour’, *Marketing Letters*, **10**(3), 187–203.
- Ben-Akiva, M., Mcfadden, D., Train, K., Walker, J.L., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D., Daly, A., De Palma, A., Gopinath, D., Karlstrom, D. and Munizaga, M. (2002a), ‘Hybrid choice models: Progress and challenges’, *Marketing Letters*, **13**(3), 163.
- Ben-Akiva, M.E., Walker, J.L., Bernardino, A.T., Gopinath, D.A., Morikawa, T. and Polydoropoulou, A. (2002b), ‘Integration of choice and latent variable models’, in H.S. Mahmassani, (ed.), *In Perpetual Motion: Travel Behaviour Research Opportunities and Challenges*, Amsterdam: Pergamon, pp. 431–470.
- Bhat, C. (1999), ‘An analysis of evening commute stop-making behavior using repeated choice observations from a multi-day survey’ *Transportation Research Part B: Methodological*, **33**(7), 495–510.
- Bierlaire, M. (2003), ‘BIOGEME: a free package for the estimation of discrete choice models’, *Proceedings of the 3rd Swiss Transport Research Conference*, Monte Verità, Ascona, Switzerland.
- Bierlaire, M., and Fetiariison, M. (2009), ‘Estimation of discrete choice models: Extending BIOGEME’, *9th Swiss Transport Research Conference*, September, Monte Verità, Ascona, Switzerland.
- Bolduc, D. and Álvarez-Daziano, R. (2009), ‘On estimation of hybrid choice models’, in S. Hess and A. Daly (eds), *Choice Modelling: The state-of-the-Art and The State-of-Practice*, Bingley, UK: Emerald.

- Bolduc, D., N. Boucher and Álvarez-Daziano, R. (2008), 'Hybrid choice modeling of new technologies for car choice in Canada', *Transportation Research Record: Journal of the Transportation Research Board*, **2082**, 63–71.
- Cao, X., Mokhtarian, P. and Handy, S. (2009), 'Examining the impacts of residential self-selection on travel behaviour: a focus on empirical findings', *Transport Reviews*, **29**(3), 359–395.
- Handy, S., Cao, X. and Mokhtarian, P. (2005), 'Correlation or causality between the built environment and travel behavior? Evidence from Northern California', *Transportation Research Part D*, **10**(6), 427–444.
- Hanson, S. (1982), 'The determinants of daily travel-activity patterns: relative location and sociodemographic factors', *Urban Geography*, **3**(3), 179–202.
- Hurtubia, R., Atasoy, B., Glerum, A., Curchod, A. and Bierlaire, M. (2010), 'Considering latent attitudes in mode choice: The case of Switzerland', *Proceedings of the World Conference on Transport Research (WCTR 2010)*, July 11–15.
- Kitamura, R., Mokhtarian, P.L. and Daidet, L. (1997), 'A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area', *Transportation*, **24**(2), 125–158.
- Kitrinou, E., Polydoropoulou, A. and Bolduc, D. (2010), 'Development of integrated choice and latent variable models for the residential relocations decision in island areas', in S. Hess and A. Daly (eds), *Choice Modelling: The State-of-the-Art and the State-of-Practice. Proceedings from the Inaugural International Choice Modelling Conference*, Bingley, UK: Emerald.
- La Paix, L. (2010), 'The relationship between trip generation and urban density: mobility patterns based on ordered probit models' (in Spanish: 'Relación entre la generación de viajes y la densidad urbana: caracterización de la movilidad con modelos probit ordenados', *IX Conference of Transport Engineering CIT*, 5–7 July, Madrid, Spain.
- La Paix, L., Monzón, A. and Cherchi, E. (2010), 'Location effects and trip generation: evidence from Madrid Metropolitan Area', *12th World Conference on Transport Research*, 11–15 July, Lisbon, Portugal.
- La Paix, L., Monzón, A. and Cherchi, E. (2012), 'Modelling the relationship between urban environment and travel behaviour: policy and indicators', *Conference Proceedings of MOBIL.TUM 2012 International Scientific Conference on Mobility and Transport*, 19–20 March, Munich, Germany.
- Limanond, T. and Niemeier, D.A. (2004), 'Effect of land use on decisions of shopping tour generation: a case study of three traditional neighborhoods in Wa', *Transportation*, **31**(2), 153–181.
- Litman, T. (2005), 'Land use impacts on transport: how land use factors affect travel behavior', *working paper*, Victoria Transport Institute.
- Madrigal, E. and Monzón, A. (2007), 'Applying an activity-based travel diary compared to a trip-based travel diary in a central and an outlying zone in Madrid', *86th Annual Meeting of Transportation Research Board*, Washington D.C.
- Mayer, R.C. (1995), 'An intergrative model of organizational trust', *The Academy of Management Review*, **20** (3), 709–734.
- Mitchell, R.B., and Rapkin, C. (1954), *Urban Traffic: A Function of Land Use*, Columbia University Press.
- Morency, C., Paez, A., Roorda, M.J., Mercado, R. and Farber, S. (2011), 'Distance

- traveled in three Canadian cities: Spatial analysis from the perspective of vulnerable population segments', *Journal of Transport Geography*, **19**(1), 39–50.
- Naess, P. (2006), 'Accessibility, activity participation and location of activities: exploring the links between residential location and travel behaviour', *Urban Studies*, **43**(3), 627.
- Paez, A., Scott, D., Potoglou, D., Kanaroglou, P. and Newbold, K.B. (2007), 'Elderly mobility: Demographic and spatial analysis of trip making in the Hamilton CMA, Canada', *Urban Studies*, **44**(1), 123–146.
- Raveau, S., Alvarez-Daziano, R., Yáñez, M.F., Bolduc, D. and Ortúzar, J. de D. (2010), 'Sequential and simultaneous estimation of hybrid discrete choice models', *Transportation Research Record: Journal of the Transportation Research Board*, **2156**, 131–139.
- Roese, N., Fessel, F., Summerville, A., Krnger, J. and Dilich, M. (2006), 'The propensity effect when foresight trumps hindsight', *Psychological Science*, **17**, 305–310.
- Roorda, M.J., Páez, A., Morency, C., Mercado, R. and Farber, S. (2009), 'Trip generation of vulnerable populations in three Canadian cities: A spatial ordered probit approach'. *88th Annual Meeting of Transportation Research Board*, Washington D.C.
- Schmöcker, J.D., Quddus, M.A., Noland, R.B. and Bell, M.G.H. (2005), 'Estimating trip generation of elderly and disabled people: analysis of London data', *Transportation Research Record: Journal of The Transportation Research Board*, **1924**, 9–18.
- Shiftan, Y. (2000), 'The advantage of activity-based modelling for air-quality purposes: theory vs practice and future needs', *Innovation*, **13**(1), 95.
- Shiftan, Y. (2008), 'The use of activity-based modeling to analyze the effect of land-use policies on travel behavior', *The Annals of Regional Science*, **42**(1), 79.
- Sitkin, S.B. (1992), 'Reconceptualising the determinants of risk behaviour', *The Academy of Management Review*, **17** (1), 9–38.
- Walker, J. and Ben-Akiva, M. (2002), 'Generalized random utility model', *Mathematical Social Sciences*, **43**(3), 303–343.
- Walker, J.L. and Li, J. (2007), 'Latent lifestyle preferences and household location decisions'. *Journal of Geographical Systems*, **9**(1), 77–101.
- Yáñez, M.F., Raveau, S. and Ortuzar, J.de D. (2010), 'Inclusion of latent variables in mixed logit models: modelling and forecasting', *Transportation Research Part A: Policy and Practice*, **44**(9), 744–753.