Combining Spatial and Temporal Dimensions in Destination Choice Models.

Cinzia Cirillo, Eric Cornélis, Laetitia Legrain and Philippe L. Toint

FUNDP – GRT
University of Namur
Transportation Research Group
8, Rempart de la Vierge
5000 Namur BELGIUM
tél : ++32.81.72.49.23, fax : ++32.81.72.49.14,
email: grt@math.fundp.ac.be

ABSTRACT

The context. This paper deals with the question of including travel constraints in a modelled representation of spatial choices.
The basic idea is that the individual's daily activity pattern is inserted within a space-time bounded region and that the number of possible alternatives processed by a single individual is limited (Hägerstrand, 1970). We also want to point out that work trips and non-work trips are chained into complex travel patterns. Thus activity-based framework (Cirillo and Toint, 2001) accounting for complex interactions across trip chains can better represent travel behaviour under spatial and temporal constraints. The use of advanced discrete choice models will give good estimation of the parameters of the utility function and correct prediction of choices. **The data sources.** The data source used in this context is Mobel, the Belgian National Mobility Survey held in 1999. Information on socio-demographic characteristics of the household and each individual in the household are available, as well as a one-day travel diary to be filled out by all members of the household. For this paper, only the travel data related to households from the Antwerpen province were retained. To this core data set, level of service data and land-use data variables have been added using a GIS. **The alternative generation process.** Usually the choice set is constructed by adding to the chosen alternatives a number of alternatives randomly taken from the entire set of possible alternatives. In the literature there are several approaches to drawing the subset of alternatives from the universal choice set (Ben-Akiva et al., 1984). In our *spatial–temporal* context, the number of possible destinations is restricted by keeping only those alternatives that are included in the household individual *action-space* (Dijst and Vidakovic (1997)), i.e. the area containing all activity places which are reachable, subject to a set of temporal and spatial constraints. **The types of variables.** A large number of explanatory variables found significant by other authors will be estimated: level of service variables (in-vehicle travel time, out-of-vehicle travel time and travel cost), level of employment (retail and service employment), and area by land use type (industrial, parking …) (Kockelman (1996); Simma, Schlich and Axhausen (2002) and Bhat, Govindarajan and Pulugurta, (1998)). **The modelling issues.** The assumption that the error terms are IIA distributed is no longer satisfactory for destination choice models. The use of mixed logit allows us to take into account heterogeneity, state dependency and heteroscedasticity across individuals (Revelt and Train, 1998) and across spatial-analytic issues (Bhat and Zhao, 2002).

1. **Introduction**

In traditional transport planning model systems, destination choice models are used to forecast trips/tours made by individuals/households from origins to destinations. Aggregate models have been replaced by more powerful disaggregate choice models (BenAkiva et al., 1984). There are various reasons for this shift: (a) disaggregate models can be estimated with smaller travel surveys, (b) a large number of explicative variables can be estimated, (c) attraction-end models should be casual (Bhat et al.,1998). Recent progresses in travel behaviour studies and discrete choice analysis push future work in this direction. In activity-based models, transport demand is seen as derived demand to pursue activities distributed in space and time (Jones et al, 1990 or Axhausen and Garling, 1992). Temporal aspects have been modelled in departure time choice models (Bradley and al., 2001) and in mode choice models (Cirillo et al.,2002). Spatial effects have yet to influence the fundamental specification of destination choice models. In this paper we attempt to attack three fundamental problems arising in such models: the alternative generation process, the type of variables to be used in the model specification, and the spatial dependency as well as the spatial heteroscedasticity in model estimation.

The destination choice model is characterised by a large number of alternatives. Most of the models are estimated on a sub-set of alternatives in the choice set; the simplest way to extract the alternatives is to use a random sampling approach. The total number of alternatives is fixed (usually from seven to ten) and includes the chosen alternatives and the random sample of non-chosen alternatives. (Bhat et al., 1998; Srour et al, 2000). However, a simple random sample is not an efficient scheme since for any given traveller the vast majority of traffic zones would have a very small choice probability. Ben-Akiva and Lerman (1985) propose to design a sample of alternatives in which the high probability zones have a higher probability of being selected. Two factors are usually considered: distance between origin and destination and size of destination. In this paper we choose to express these factors by using Dijst and Vidakovic's (1997) definition of action space for identifying the sample set of random alternatives.

The choice of explanatory variables in the attraction-end models is another important point of this research work. We have included in the model four sets of variables: impedance variables, zone size measures, land use variables and interaction of spatial variables with impedance. In particular we use land use explicative variables that are derived from a GIS transformation of a satellite image of the city of Antwerpen (Belgium) and its surroundings. We believe that the explicit consideration of those variables in destination choice model can better sustain the analysis of land use allocation and travel demand in an integrated land use-transport model system. The behavioural data used for this model is obtained from Mobel, the first Belgian household and travel survey held in 1999 (Hubert and Toint, 2002).

Following Bhat and Zhao (2002), who studied spatial dependency, spatial heterogeneity and spatial heteroscedasticity in activity stop generation models, we also investigate these spatial effects. Those aspects should play an important role also in modelling individual destination choices.
Individual and spatial dependencies are defined as the correlation existing across observations collected from the same individuals/households, and the correlation among destination choices belonging to zones with similar characteristics. For example individuals can chain trips into complex patterns by choosing destinations close to their work or home locations; they can be influenced by unobserved socio-demographic or spatial variables. The explicit assumption of independence of individual and spatial observations can induce an under-

estimation of the standard error and minor changes in coefficients estimation (Cirillo et al, 2000).
Heterogeneity effects will cause diversity of preferences and sensitivity in spatial choices. Individuals can have different perception of distances, in-vehicle travel time and travel cost; we expect the impedance variable to vary a lot across the population. Mixed logit formulation allows the estimation of the mean variable and the standard deviation for an assumed distribution.
Heteroscedasticity is caused by the variance of the error terms across spatial units. The unobserved component accounts for both errors made by the choice maker and by the analyst during model estimation. For example the former may not be exactly aware of the number of shopping facilities in the chosen or competitive areas, or miss-estimate the accessibility to those zones. The latter may not be exactly sure about the measure of the land use and size variables.
The modelling of these effects is carried out using the mixed logit model and results are compared to a traditional multinomial logit model.

The paper is organised as follows. Section 2 is devoted to the explanation of the model specification and estimation. We then explain the procedure to calculate the action space and to generate the alternatives in Section 3. Section 4 is dedicated to the description of all variables used in model estimation and section 5 presents the numerical results. Conclusions and research perspectives are given in Section 6.

2. Model structure

Discrete choice models are used below to describe how individuals perform a selection among a finite number of destination alternatives. Let I be the population size and $A(i)$ the set of available destinations for individual i, $i = 1,..., I$. For each individual i, each alternative A_z, $z = 1, ..., |A(i)|$, has an associated utility, depending on the individual characteristics and the relative attractiveness of the alternative. In our case the destination choice model can be specified as:

$$U_{iz} = \lambda^T \ln y_z + \mu \ln M_z + \sum_j \gamma x_{izj} + \varepsilon_{iz} \qquad (1)$$

where:

y_z is a vector of observed land use zone-specific variables and λ is a column vector of coefficients fixed across all zones;

$M_z = \sum_{k=1}^{s} \beta_k M_{zk}$ is the measure of the size of the alternative z, where M_{zk} is the kth size variable for zone z, β_k is the corresponding coefficient and μ is a positive scale parameter, with $0 \le \mu \le 1$. In some case, when the average utility is supposed to be well specified, we may impose $\mu = 1$. In this case the parameters of the choice model are not dependent on the definition of aggregate alternatives. (Ben Akiva and Lerman, 1985);

x_{izj} are the components of a column vector of exogenous accessibility variables for individual i in zone z and γ is a vector of random parameters;

ε_{iz} is the error term, which is assumed to be independently and identically Gumbel distributed across zones alternatives and individuals.

The theory then assumes that individual *i* selects the alternative that maximizes his/her utility.

Mixed logit models relax the assumption that the parameters γ are the same for all individuals, by assuming instead that individual parameters vectors $\gamma(i)$, $i = 1,..., I$, are realisations of a random vector γ'. We assume that γ' is itself derived from a random vector δ and a parameters vector θ, which we express $\gamma' = \gamma'(\delta,\theta)$. For example, if γ' is a *J*-dimensional normally distributed random vector, we may choose $\delta = (\delta_1, \delta_2, ..., \delta_J)$ with $\delta_j = N(0,1)$, and let θ specify the means and standard deviations of the components of γ'. The probability choice is then given by:

$$P_{iz}(\theta) = E_P[L_{iz}(\delta,\theta)] = \int L_{iz}(\delta,\theta)P(d\delta) = \int L_{iz}(\delta,\theta)f(\delta)d\delta \qquad (2)$$

where P is the probability measure associated with δ and $f(\bullet)$ is its distribution function, and L_{iz} is the logit formula given by

$$L_{iz} = \frac{e^{V_{iz}}}{\sum_{l=1}^{|A(i)|} e^{V_{il}}}.$$

The vector of parameters θ is then estimated by maximizing the log-likelihood function, i.e. by solving the program:

$$\max_{\theta} LL(\theta) = \max_{\theta} \frac{1}{I} \sum_{i=1}^{I} ln\, P_{iz_i}(\theta) \qquad (3)$$

where z_i is the alternative choice made by the individual *i*. This involves the computation of $P_{iz_i}(\theta)$ for each individual *i*, *i = 1, ..., I*, which is impractical since it requires the evaluation of one multidimensional integral per individual. The value of $P_{iz_i}(\theta)$ is therefore replaced by a Monte-Carlo estimate obtained by sampling over δ, and given by

$$SP_{iz_i}^R = \frac{1}{R}\sum_{r=1}^{R} L_{iz_i}(\delta_r,\theta) \qquad (4)$$

where R is the number of random draws δ_r, taken from the distribution function of δ. As a result, θ is now computed as the solution of the simulated log-likelihood problem

$$\max_{\theta} SLL^R(\theta) = \max_{\theta} \frac{1}{I} \sum_{i=1}^{I} ln\, SP_{in_i}^R(\theta) \qquad (5)$$

We will denote by θ_R^* a solution of this last approximation (often called Sample Average Approximation, or SAA), while θ^* denote the solution of the true problem (3).

In order to take into account spatial correlation we need to adopt the mixed logit formulation on panel data first provided by Revelt and Train in 1998 and then largely applied in a number of application studies. (Algers et al, 98). The part of the utility given by random variables

should be $U_{izt} = \varsigma x_{izt} + \xi x_{izt} + \varepsilon_{izt}$ where t is the choice occasion registered in the clustered zone z for individual i. The maximisation algorithm is changed by drawing a random zone specific coefficient γ_z from its distribution for each cluster. This coefficient is kept constant for the observations in the zone cluster.

A major drawback of Monte-Carlo methods is their slow convergence rate. Quasi Monte-Carlo methods attempt to speed up this convergence by being more directive in the choice of the sampling points used to evaluate the choice probabilities. This led to suggest the use of Halton (Bhat, 1999) draws instead of Monte-Carlo random draws. He found that when the number of random parameters is small (typically less or equal to five), approximation of the objective function based on Halton sequence usually succeeds to give the same results as pure Monte-Carlo sampling with less random draws.

To estimate (1) we adopt a two steps calibration procedure:

step 1 Estimate the "size variables" and calculate the attraction utility part, lnM_z as follows

$$lnMz = \ln(M_{z0} + \exp(\beta_1) * M_{z1} + \exp(\beta_2) * M_{z2} + ... + \exp(\beta_s) * M_{zs}) \qquad (6)$$

In destination choice models the "size variables" include attraction measures such as population, employment, number of shopping centres, number of stores, leisure activity area and surface. The general case with two or more size variables yields a logit model that is non-linear in its parameters. We adopted the special logit estimation procedure developed by Daly (1982) and implemented into Alogit. Not all size variable coefficients can be identified, and it is necessary to impose a scaling restriction such as $\beta_0 = 1$.

step 2 Estimate the mixed logit model with random parameters on impedance measure.

3. Alternatives generation process

In a destination choice model, the problem of the huge size of the choice set obliges modellers to use some restriction assumptions in order to define a subset of alternatives. The most common method to draw this subset is to consider a random sampling of, say, nine non-chosen alternatives from the entire choice set to which the chosen alternative is added.

Because of our activity based approach, our aim is to eventually identify a restricted area where the individual is supposed to achieve all the trips contained in his/her daily activity chain. This area will define a restricted choice set among which a set of nine alternatives will be randomly selected. We were inspired by the notion of "daily prism" previously developed by Hägerstrand (1970). As defined in Dijst and Vidakovic (1997), "the prism describes a set of positions in space-time for which the probability of being included in the individual path is greater than zero. The projection of the prism onto space gives the potential **action space**. The potential action space is defined by the area containing all activity places which are reachable subject to a set of temporal and spatial conditions". This action space approach represents a way of taking space and time into account in our model.

We suppose that the individuals organise all their activities around spatially fixed *bases* as home and work places. All *non-base stops* made between a base to the same or another base are subject to modelling.

The spatio-temporal constraints consist of the *time-budget*, defined by the total time interval available to make a trip between two bases including the activity duration, the *travel speed*, the *distance between home and work* and the *travel time ratio* which is the ratio between the travel time and the time budget.

Workers and non-workers are considered separately. For workers, the daily activity chain can be divided into four parts: the first one consists of leaving home to commute to the work place (the *morning commute*); the second represents a tour from work at lunchtime (*midday tour*); the third part is the *evening commute* from work to the home place and the last one is constituted by a tour from home after the working day (*after tour*). We suppose that the individual can undertake at most two non-base stops during the midday and after tours while only one non-base stop both in the morning and evening commutes can be considered.

At variance, all non-worker's trips are assumed to be home-based. Hence, the non-worker daily chain can be divided into three parts: a tour during which the individual undertakes his/her daily activity with the longest duration (the *main tour*); the *before tour* represents the tour before the individual begins his/her main tour; and finally the *after tour* consists of the tour made after the return from the main tour. The individual can have at most two non-base stops during the before and after tours. In the main tour, two non-base stops can be undertaken in addition to the activity.

3.1 The action space building

An action space has to be built for each tour and/or commute forming the daily activity chain. Indeed non-base stops made between to different bases determine an action space which consists in an ellipse while stops included in a tour induce a circle.

According to Dijst and Vidakovic (1997), the ellipse is given by the following equation:

$$\frac{(x-\frac{l}{2})^2}{(\frac{\tau TV}{2}+\frac{L}{2}(1-\tau))^2}+\frac{y^2}{(\frac{\tau TV}{2}+\frac{L}{2}(1-\tau))^2-(\frac{L}{2})^2}\leq 1 \qquad (7)$$

where,
T is the time-budget;
V is the travel speed;
L is the distance between bases;
τ is the travel time ratio;
x, y are the coordinates of points belonging to the action-space.

Note that the action space becomes a circle when L vanishes. Moreover, when no non-base stops are undertaken (i.e. when $\tau = 1$) the action space becomes a line. In this case, all the time-budget is spent to travel between home and work.

The travel time is the sum of all travel times from the departure base to arrival at the same or another base (i.e. addition of the travel times for all the stages (from departure base to non-base stop, from non-base stop to other non-base stop if appropriate and from non-base stop to arrival base)), while the activity duration required in computation of time-budget is the one involved in non-base stop activities i.e. the sum of non-base activity stops stay times.

We then compute the longer principal axis from (7) to use the bifocal definition of the ellipse in order to approximate the action space by:

$$dist(B_1, S) + dist(B_2, S) \leq LA \tag{8}$$

where B_1 is the first base, B_2 is the second base, S is a non-base stop and LA is the measure of the longer principal axis for the action-space ellipse.

In the case of non-base stops included in a home-based or work-based tour, (8) reduces to:

$$dist(B, S) \leq R \tag{9}$$

where B is a the base, S is a non-base stop and R is the radius of the circle forming the action space.

All trips information collected in the MOBEL survey data has been previously geocoded at the street level with help of ArcGIS 8.1. For this study, we restrict the MOBEL data set to the trips located in the Antwerpen province. This leads us to a set of 351 individuals for which the entire daily activity chain has been successfully geocoded. Each geo-coded activity chain is separated in different tours and/or commutes leading to three or four, respectively for non-workers and workers, types of action space to determine. In this way, 456 non-base stops are observed which represent a set of 369 action spaces to build.

For this model, the elemental alternatives are aggregated to statistical sectors (a unit used by the National Statistical Office and the smallest geographical unit for which land use or commercial information is available). Hence all geocoded information was reported to a statistical sectors map.
The province of Antwerpen is divided into 13 914 statistical sectors. Given a file containing the coordinates of the center of each statistical sector and given the coordinates of home and work places we keep, for each pair of bases or base (for non-workers), all statistical sectors satisfying the equation (8) (or (9)). In this way, we get a representation of our restricted choice set, consisting of this action space. Then for each observation (i.e. non-base stop), the actual subset of alternatives is determined by randomly selecting nine alternatives from the set of statistical sectors in the action-space to which the real chosen destination is added.

(Examples of action space are available in appendix.)

4. Variables description

In our model, we make use of the following aggregated **land use variables**:

- Statistical sector area (total geographical area of the sector [m²]) ;

- densely-built housing ;
- built-up housing ;
- housing and other developments ;
- industrial / commercial / port area ;
- infrastructure (highways, district roads, airport and/or railway infrastructure and so on);
- green / nature area (broad-leaved, coniferous and/or mixed forests, municipal parks, heath land and moors, dunes and beaches, water);
- agriculture and meadowland (agriculture and open space, meadowland and orchards).

These variables were provided by the team *"Vakgroep transport en ruimte"* from the University of Antwerpen (UFSIA). The data are expressed as percentages of the geographical area of the statistical sector by land use type. They assembled these variables using maps[1] from the Flemish Land Agency/Support Centre GIS-Flanders and geographical information system technology. Moreover, they provided population and employment related variables which they assembled respectively from National Institute of Statistics and Regional Development Agency data.

From a set of 29 **commercial variables**, we consider the following aggregated variables:

- shopping variable;
- financial variable (banks);
- hotel / restaurant / café;
- cinemas;
- sport activities;
- cultural, recreational and leisure activities (museum, library, school of music, zoo, nature reserve, theatres, casino and so on);
- car retail;
- personal service (beauty center and so on).

These variables were provided by the Statistical Services (SES) from the Ministry of the Walloon Region and calculated using data from EURO DB (which manages and delivers data from the commercial register) and the National Office for Social Security (N.O.S.S.).

The last variables are **impedance variables,** which include cost and in-vehicle travel time (IVTT). For each origin-destination pair, we used the distance to calculate the COST and IVTT variables for the chosen travel mode. Four categories of travel mode are considered: *walk and bike*, *train*, *bus* and finally *car*. For the first one, we assign a cost equal to zero. The cost of train is established by means of SNCB (the Belgian railways) tariff distances. A fixed cost chosen to be 1.1 euros is assigned to the bus mode. When the car is used, we assumed that it costs 0.17 euros per kilometer.
IVTT is obtained by the ratio of the travel distance (between base and non-base stop) and the mean travel speed (determined for each tour and commute).

By definition, these variables are strongly correlated. The correlation between COST and IVTT is overcome by converting COST into equivalent IVTT. The accessibility variable is then expressed as

[1] Rasterised version of the land cover/use data set, survey of 2001, Flemish Land Agency/Support Centre GIS-Flanders

$$Access = IVTT + \alpha\, COST\ ,$$

where α is a parameter allowing the conversion of COST into time value. The value of time was chosen to be 5 euros/hour in a European study comparing the different values of time across the European Union (Trace, 1999); it means that each euro of travel cost is equivalent to 12 minutes of travel time which is the value employed for α.

Because the mean and standard deviation of the accessibility can only be estimated in the mixed logit model, a spatial related variable called URBAVAR, which allows distinguishing perception of impedance between three levels, is introduced in the multinomial logit model. This zone related variable can take the three following values: 0 for a zone which lies in the country (outside built-up area or in a built-up area including less than 7 000 inhabitants), 1 for a built-up area with more than 7 000 inhabitants included in the set of the less dense sectors and finally, 2 for a built-up area with more than 7 000 inhabitants and included in the set of the most dense sectors.

5. Model results

Using two steps calibration procedure of Section 2, we estimated two models (multinomial logit (MNL) and mixed logit model (MXL)) using three groups of variables: accessibility measures, land use and size variables. The estimated explanatory variables are presented in Table 1.
In the MNL the accessibility on the three spatial levels identified by the variable URBAVAR was found to be negative and significant. Four land use variables affect the destination choice behaviour; all except green area were found to be positive and significant. Four size variables fit into the model, of which three could be estimated. The density of employment was found to be strongly negative: areas attracting work trips do not offer great opportunity for shopping and leisure activities.
Surprisingly cinema and restaurants were found not significant and are not shown in the table. The size variables that affected the destination choice model are the number of shops and the number of recreation activities defined as the sum of the number of recreation, sport and cultural activities reported in the statistical sectors.

In the second columns of Table 1 we report results of the mixed logit models; ten variables are reported to be significant. The MXL model allows the accessibility of destination locations to vary with normal distribution across the observations. Both mean and standard deviation of accessibility were found significant; the mean is negative and after looking at the distribution we do not report positive values. Three land use variables (all positive) have been included into the model. Six size variables are reported to be significant. The area of industry and shopping activities is the reference, the density of employment is the only variable found to be negative (as in the MNL), four variables (recreation locations, cinema, restaurants, and shops) were found positive and significant.

We also attempted to use mixed logit models to account for individual and spatial correlation as described in Section 2. At individual levels we calculated the number of tours and we generate for each of them an identical error component. The value of the log-likelihood is lower than the one registered for the global mixed logit model, but, unfortunately the standard deviation (S.D.) of the accessibility and of two of the three land use variables are no longer

statistically significant. To take spatial correlation into account we calculated the number of observations for each cluster and then we applied the same error component technique. The results again show improvements in the log-likelihood function, but we are not able to obtain good estimates of the parameters. We believe that increasing the number of observations and applying different strategies of spatial aggregation should bring further improvements in the global understanding of the destination choice behaviour.

Variable	Multinomial logit (MNL)		Mixed logit (MXL)	
	Parameter	*t*-Stat.	Parameter	*t*-Stat.
Accessibility variables				
Accessibility (Mean)			-0.4067	2.6
Accessibility (S.D.)			0.1353	2.9
Accessibility (URBAVAR = 0)	-0.0079	5.0	-	-
Accessibility (URBAVAR = 1)	-0.0089	8.2	-	-
Accessibility (URBAVAR = 2)	-0.0068	6.6	-	-
Land use variables				
Ln (agricultural area)	0.1139	1.7	0.2569	2.3
Ln (green area)	-0.1264	3.2	-	-
Ln (built area)	0.4204	11.7	0.1176	5.2
Ln (densely built area)	0.2156	3.7	-	-
Ln (housing area)	-	-	0.458	1.1
Size variables				
Industry/Shopping area	1	-	1	-
Employment	-8.075	3.5	-3.084	4.0
Recreation	1.137	1.1	2.939	2.8
Cinema	-	-	7.927	8.3
Restaurants	-	-	3.375	3.4
Shopping	3.529	4.4	2.469	2.2
Number of observations	448		448	
Log-likelihood at convergence	865.58		852.92	
Degrees of Freedom	11		10	
Adjusted ρ	0.3766		0.3858	

Table 1: Model estimation results

6 Conclusions

In this paper we propose a destination choice model which accounts for both temporal and spatial dimensions in travel demand. We exploit an activity travel framework to describe daily patterns and to calculate time budget. We propose to sample destination alternatives by using the notion of the action space as formulated by Dijst and Vidakovic.
We use accessibility measures on spatial levels, land use and size variables to estimate destination choice model for shopping and leisure activities.
We estimate the taste variation on impedance variable and we account for heteroscedasticity by the mean of a mixed logit formulation. We tested individual and spatial correlation by way of error components techniques.

Looking at the significativity of the estimated variables, we conclude that combining temporal and spatial dimensions of travel demand bring improved realism for the constitution of the available set of choice alternatives.
The improvements in the value of the log-likelihood at convergence and in the number of estimated variables, comfort us in the use of a mixed logit formulation, moreover, it allows a number of analyses not possible under classical multinomial logit hypotheses.

This study is still ongoing and we would like to improve several aspects of our methodology.
The number of observations can be considerably increased, by adding the OVG (travel behaviour survey (1999)) for the city of Antwerp. This will allow the estimation of a larger number of land use and size variables already available in the database created for this paper. Work and school trips should be included into the model. Major attention should be given to spatial correlation; demographic and social aspects are not yet considered but are supposed to play an important role in destination choice model.

There are a number of theoretical issues to be explored. Destination choice models with size variables are a typical example of non-linear utility models; the implication on mixed logit function maximisation is not yet explored. The IIA property (Independence from Irrelevant Alternatives), which holds in the multinomial logit model for the estimation of a sample of alternatives, is not longer applicable in mixed logit estimation. This requires further investigation.

This investigation is part of the SAMBA project in which results from model will be used for estimating origin/destination matrices for a synthetic population.

Acknowledgements

We would like to thank Andrew Daly for quickly replying to all our questions concerning destination choice models and size variable estimation. The authors are also grateful to all those who provided help throughout the development of this project. These include Dries Van Hofstraeten at University of Antwerpen (UFSIA) for creating land use variables, Jean Paul Duprez at “Service des études et de la statistique” (Ministère de la Région Wallonne) for providing shopping and leisure activity indices, Pierre Yves Bernard and Joëlle Vandervoorde at FUNDP-GRT for support in program developments and in the use of GIS.
This research was funded by the SSTC (Services Fédéraux des Affaires Scientifiques Techniques et Culturelles) project **SAMBA n° CP/13/411.**

Appendix

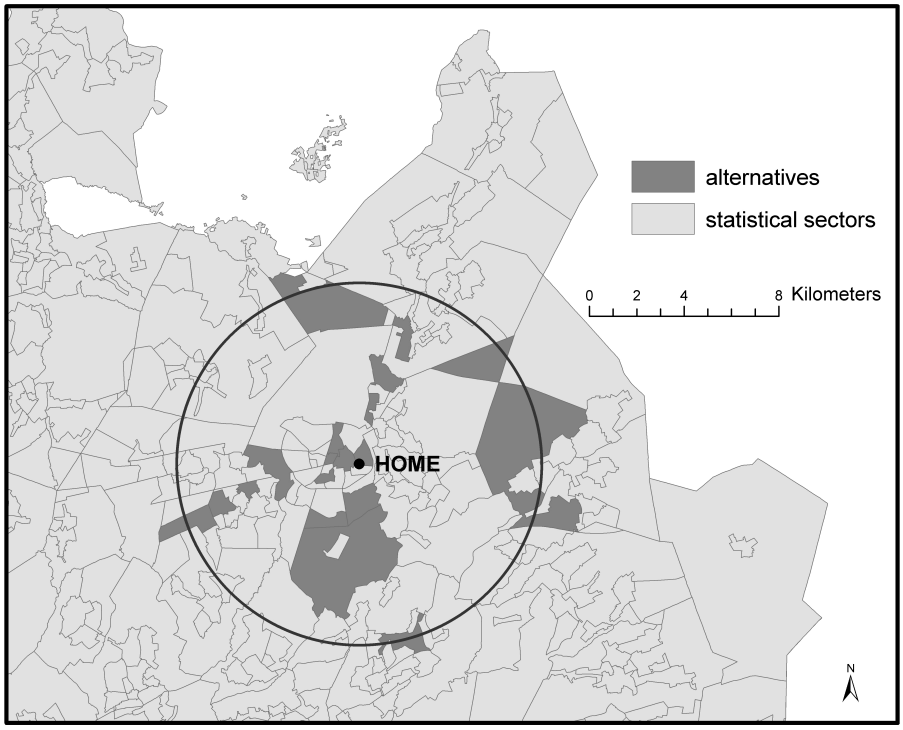

Figure 1. An example of action space for non-worker (before principal tour)

Figure 2. An example of action space (morning commute) for worker

References

Algers S, Bergstrom P., Dahlberg M and Dillen J. (1998) Mixed logit estimation of travel time. Working paper. dept of infrastructure and planning, Royal Institute of technology. Sweden

Aptech Systems Inc (2002) GAUSS 5.0 for Windows.

Axhausen, K. and T. Garling (1992) *Activity-Based approaches to travel analysis: conceptual frameworks, models and research problems*, Transport Reviews, 12, 324-341.

Barette P., Cirillo C., Cornelis E., Dessy A. and Toint P.L *Enquete Nationale sur la mobilite des ménages*, Rapport final Publications of Mathematics Department N.2001/11 at FUNDP – GRT.

Bastin F., Cirillo C. and Toint P.L. (2003) *Numerical experiments with AMLET, a new Monte-Carlo algorithm for estimating mixed logit models* in Proceedings of IATBR International Association for Travel Behaviour Research 2003 Lucerne Switzerland.

Ben Akiva M. Gunn H. and Silman L. (1984) *Disaggregate trip distribution models*, Proceedings japanese society of civil engineers.

Ben Akiva M., Lerman S.R., (1985) *Discrete choice analysis: theory and application to travel demand*, Cambridge, MA: MIT Press.

Bhat, C. R. (2001) *Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model*, Transportation Research, **35B** (7) 677-693.

Bhat, C. R. & Singh S. K., (2000). *A comprehensive daily activity-travel generation model system for workers*, Transportation Research 34A, 1-22.

Bhat, C. R., Govindarajan A. and Pulugurta V. (1998) *Disaggregate attraction-end choice modeling: Formulation and empirical analysis*, Transportation research record 1645 60-67

Bhat C.R. and Zhao H. (2002) *The spatial analysis of activity stop generation*, Transportation Research, **36B** 557-575.

Bradley M., Daly A., Hofman, Jong, Vellay C. (2001) *A model for time of day and mode choice using error components logit.* 2001 ETC proceeding

Cirillo C. and Axhausen K.W. (2002) *Modal choice of complex tours* in Proceedings of European Transport Conference Cambridge UK

Cirillo C. and Toint P.L.(2001) *An activity based approach to the Belgian national travel survey,* Technical Report – Transportation Research Group, FUNDP, 24p.

Daly A. (1982) *Estimating choice models containing attraction variables*, Transportation Research, 16**B** (7) 5-15.

Dijst M. and Vidakovic (1997) *Individual Action-Space in the city* in Activity-based approaches to travel analysis, D.F. Ettema and H.J.P. Timmermans, Pergamon, p.117-134.

Hägerstrand, T. (1970) *What about people in regional science?* Papers of the Regional Association 24, 7-21.

Hubert J.-P. and Toint P. (2002) *La mobilité quotidienne des Belges.* Presses Universitaires de Namur, 2002

Jones, P.M., Koppelman F. and Orfeuil J.P. (1990) *Activity analysis: state-of-the-art and future decision*. In Developments in Dynamic and Activity-based approaches to travel Analysis, Gower, Aldershot, UK

Kockelman, K. (1996) *Travel behaviour as function of accessibility, land use mixing, and land use balance: evidence from San Fransico bay area*, Transportation Research Record 1607.

Revelt, D. and Train K. (1998) *Mixed logit with repeated choices: household's choices of appliance efficiency level*. The review of Economics and Statistics. Vol. LXXX, No.2

Srour, I.S., Kockelman, K and Dunn T.P. (2000) *Accessibility indices: A connection to residential land prices and locations choices*. Paper presented at the 81st Annual Meeting of the transportation research board. Forthcoming in Transportation Research Record.

Train K. (1999) Mixed Logit Estimation Routine for Cross-Sectional Data, Mixed Logit Estimation for Panel Data (http://elsa.berkeley.edu/~train/software.html).