A SYSTEM OF LOCATION MODELS FOR THE EVALUATION OF LONG TERM IMPACTS OF TRANSPORT POLICIES

GENNARO NICOLA BIFULCO
Università degli Studi di Roma "Tor Vergata"
Dipartimento d'Ingegneria Civile
Via di Tor Vergata s.n.c., 00133 ROMA, ITALY

Abstract

In this article an analytical model able to deal with the complex interactions among the activities of an urban system is presented. The model is based on a consistent and rigorous analytical framework, where the urban pattern results from the location choices of various decision-makers. The model can be stated in a disaggregate form where choices are simulated within a behavioural framework based on the random utility theory. Aggregation issues can also be consistently treated. An operative specification of the model is also presented, as well as its application to the urban area of Naples.

INTRODUCTION

The main aim of this paper is to present an analytical model able to deal with the complex set of activities' interactions of an urban system. Urban systems are complex, since they are described by a large number of variables, are characterised by non-linear relationships among variables, dynamic dependencies and a system evolution that is dependent on behavioural factors.

Three main sub-systems can be identified within an urban system: transportation, land market and urban activities. Within urban activities, in turn, three further subsystems can be recognised: housing, employment and workplace. Housing depends on where households locate their residences, employment depends on where firms locate their activities and workplace depends on where people choose to locate their job (choosing among the available employment). Transportation system directly influences (through active and passive accessibilities) housing, employment and workplace locations. Minor firms locate "diffuse" services (say retails) and contribute to the whole employment supply together with major public and private firms that locate production and "non-diffuse" services. The whole employment supply and transportation system performances influence workplace choices. Moreover, given workplaces, transportation system performances and land market, households choose housing locations. Workplace, housing and firm locations induce transportation demand. Service and retail locations are influenced by housing and workplace locations and by transportation demand for service purposes. Land market subsystem is directly related with employment (where firms locate their activities) and housing subsystem. It influences, through rents and prices, the location convenience both for households (housing) and firms (service and retail employment). On the other hand, housing choices and firm locations influence floor-space demand, thus influencing rents and prices. Two main decision-makers can be considered in order to describe the mechanisms related to land market: public (mainly defining rules and constraints) and private investors. Sometimes public investors can act as building investors. This happens when residential demand and prices suggest intervening in defence of weaker socio-economic classes. This results in overall price-calming effects because part of the residential demand is "subtracted from the market". In any case, rents and prices also influence private building investment choices, which in turn influence supply and, finally, rents and prices again.

The complex interactions previously described have been differently approached in literature. Excellent reviews of the state of art can be found in Putman (1975) Kim (1983) and De la Barra (1989). A first approach in location models is due to the microeconomic theory, from earlier works of Von Thunen and Wingo to the model proposed by Alonso (1964). These models are characterised, on one hand, by simplistic fundamental hypotheses but, on the other hand, by a rigorous analytical approach. Such kind of models can be defined "internally consistent" but "externally inconsistent", in the sense that simplistic fundamental hypotheses lead to unrealistic representations of the real urban system mechanisms. Other major approaches derive from the early study of Lowry (1964), where spatialisation is represented through an aggregate interaction theory, based on area discretisation. Locations are mainly expressed as the result of the joint effect of transportation costs and location benefits. The location model can be interpreted both from an interaction (mainly gravity models) point of view and from the maximum entropy analogy (Wilson, 1973). Input/output-like models can also be used, in their spatialised-form (multiregional I/O), in order to describe spatial interactions due to "basic" activities and induced ("non-basic") activities. This still is an aggregated modelling approach, mainly based on activity interactions. Lowry-type models can also be treated as multiregional I/O models (see Macgill, 1977). The work of Kim (1983) integrates input/output approaches

together with a simplified transportation model dealing with generation, destination and route choices, treated at an aggregate level. The resulting model is characterised by land-use substitution factors for housing vs. good production, depending on land-use densities. The model, which leads to a non-linear programming problem, is still based on spatial interaction between activities, decision-maker choices are not explicitly taken into account. Perhaps the most widely used urban model is embedded within the MEPLAN software package (for an application see Echenique, 1995). Transportation demand is generated from interaction between activity sizes and activity locations and is also influenced by accessibilities. MEPLAN integrates into a powerful software package elements from I/O theory and transportation models. Interactions are treated within a quasi-dynamic approach where successive steady states are investigated. The hypothesis is made that two different dynamic rates influence transportation system and activity system. Transportation system dynamics is hypothesised to be much faster than urban dynamics, so it can be neglected, while the urban system reacts to subsequent changes of accessibilities in a slower way. Spatial interactions are still treated in an aggregate form and activity relationships are based on a descriptive theory rather than on a theory based on decision makers.
The general model presented in the following is based on a consistent and rigorous analytical framework, in which the urban pattern results from the location choices of several decision-makers. The model can be stated in a disaggregate way, where choices are simulated within a behavioural framework based on random utility theory. Aggregation issues can also be consistently treated. Dynamic and equilibrium approaches can be applied to the resulting problem and are here formalised.

PROBLEM FORMALISATION

The analytical formalisation introduced in the following section is referred to a general urban system. The model is described in a disaggregate way, then aggregation issues will be treated. Both dynamic and equilibrium approaches for the solution of the location model will be presented, as well as a first analysis on the equilibrium theoretical properties.

Notation and definitions

It is possible to distinguish two main phenomena within an urban system. On one hand the location of activities depends on endogenous mechanisms. On the other hand, activities can also be located exogenously (in an anelastic and fixed way), at least given the current spatial extension of the study area and/or the considered "simulation horizon". Obviously, the location of an activity could be not totally endogenous or exogenous.
Let group numerical indexes identifying all the considered activities in set Y, while X is the set of activity indexes that are, at least in part, endogenously located ($X \subseteq Y$).
The endogenous location mechanism is influenced by "urban system cost functions" (e.g.: transportation costs, housing prices, ...). Be H the set of all cost functions of the urban system and $h_p(\ldots)$ a generic cost function $h_p(\ldots) \in H$.
Moreover, let us indicate the set of all possible activity locations by L.
The location of the activities is subjected to "land-use constraints" (e.g.: maximum number of residential units that can be built in a given area). If a constraint is set up to a null value the corresponding activity is not allowed at all to be localised; on the other extreme, if the activity location can be treated as unconstrained the corresponding constraint value can be set up to infinity.
Let us indicate as u^i_a the maximum amount of activity $a (\in X)$ allowed to be located in $i (\in L)$ ($0 \le u^i_a \le \infty$). The constraint array is $\mathbf{u}_a = [\ldots, u^i_a, \ldots]^T$. The amount of activity located by effect

of endogenous mechanism be expressed by array $\mathbf{ee}_a=[\ldots, ee^i_a, \ldots]^T$ ($\forall$ a$\in$X, $\forall$ i$\in$L). Where ee^i_a is the amount of activity a($\in$X) endogenously located in i($\in$L).
Similarly $\mathbf{ex}_a=[\ldots, ex^i_a, \ldots]^T$ ($\forall$ a$\in$Y, $\forall$ i$\in$L) be the array of exogenous locations.
Consistency with constraint array leads to:

$$\mathbf{ee}_a + \mathbf{ex}_a \le \mathbf{u}_a \quad \forall\ a\in X \tag{1}$$

Since exogenously located activity should respect constraints ($\mathbf{ex}_a \le \mathbf{u}_a$):

$$0\le \mathbf{ee}_a \le \mathbf{u}_a-\mathbf{ex}_a \quad \forall\ a\in X \tag{2}$$

Assume $\mathbf{E}_X$ and $\mathbf{E}_Y$ be the arrays of endogenously and exogenously located amounts:

$$\mathbf{E}_X=[\ldots, \mathbf{ee}_a^T, \ldots]^T,\ \forall\ a\in X;\ \mathbf{E}_Y=[\ldots, \mathbf{ex}_a^T, \ldots]^T,\ \forall\ a\in Y. \tag{3}$$

Endogenous locations of a generic activity (a$\in$X) are chosen by decision-makers. They are "decisional units", not necessarily coincident with a single decision-maker. The decisional unit is defined as a group of individuals choosing locations in a unitary way (for example: household, firm, ...). Different individuals can belong to different decisional units, choosing location of different activities.
Be Φ_a the set of all decisional units that choose locations of activity a($\in$X), d a generic decisional unit and q_{ad} the amount of activity a($\in$X) to be located by decisional unit d.
Note that, generally speaking, the amount of activity q_{ad} could be location-dependent and/or dependent on the cost functions of the urban system. For example, households could choose bigger or smaller houses depending on the housing price in the zone they are choosing in. For instance, if the amount of located activity ($\mathbf{ee}_a$) is "squared meters of flat", q_{ad} is zoning dependent; of course this is not the case if $\mathbf{ee}_a$ simply means "houses". In the following the hypothesis is made that q_{ad}'s can be in any case expressed in a zone-independent way and are not dependent on cost functions.
Within each decisional unit it is possible to distinguishing different elements (for instance the components of a family). They are "forced" to choose in an unitary way, however each of them could have his own different vantages or disadvantages for each current choice.
In any case, define D_d as the set of all the elements of decisional unit d and δ one of those element ($\delta\in D_d$). Each element δ could be influenced in the location choice by a "spatial root". Assume, for instance, that one element of a family would influence the choice of the residential location of the household, say because he/she likes to be as close as possible to his/her workplace. In such a case, workplace is the "spatial root". Of course, within the same family, each element could have a different spatial root. Assume that $r^j_{\delta a}$ is the probability an element δ is influenced by spatial root j in choosing location of activity a$\in$X.
The probability that a generic element of a given decisional unit locates an activity in a given location, provided that a given root influences him, can be computed according to the random utility theory.
In any case, it can be formally written as:

$$\forall\ a\in X,\ \forall\ d\in\Phi_a,\ \forall\ \delta\in D_d,\ \forall\ i\in L,\ \forall\ j\in L;\quad p^{ji}_{\delta a}=p^{ji}_{\delta a}(\mathbf{V}^{j\cdot}_{\delta a}) \tag{4}$$

(subjected to $p^{ji}_{\delta a}\ge 0$, $\Sigma_{i\in L}\, p^{ji}_{\delta a}=1$; $u^i_a-ex^i_a=0 \Rightarrow p^{ji}_{\delta a}=0$)
where:

$p^{ji}_{\delta a}$ = probability that component $\delta(\in D_d)$ of the decisional unit $d(\in\Phi_a)$ chooses to locate activity a($\in$X) in location i($\in$L), given that he/she is influenced b root j($\in$L);

$V^{ji}_{\delta a}$ = systematic utility related, within the random utility theory, to location alternative i($\in$L) for activity a($\in$X) in the choice of component $\delta(\in D_d)$ of decisional unit $d(\in\Phi_a)$, influenced by spatial root j($\in$L);

$\mathbf{V}^{j\cdot}_{\delta a}$ = array of $V^{ji}_{\delta a}$'s $\forall i\in L$, $\mathbf{V}^{j\cdot}_{\delta a} = [\ldots, V^{ji}_{\delta a}, \ldots]^T$

For instance, if a logit-type model is used, the probability is:

$$p^{ji}_{\delta a} = p^{ji}_{\delta a}(\mathbf{V}^{j}_{\delta a})= \frac{exp\left(\alpha V^{ji}_{\delta a}\right)}{\sum_{k\in L} exp\left(\alpha V^{jk}_{\delta a}\right)} \tag{5}$$

Thus the location probability of the whole decisional unit can be written as:

$$\forall\ a\in X,\ \forall\ d\in\Phi_a,\ \forall\ i\in L;\quad \varphi^i_{da}= g_a(\mathbf{p}^{\cdot i}_{da}, \mathbf{W}^{da}_{\cdot\cdot}, \mathbf{r}^{\cdot}_{da}) \tag{6}$$

(subjected to. $\varphi^{i}_{da} \geq 0$, $\Sigma_{i \in L} \varphi^{i}_{da}=1$, $u^{i}_{a}-ex^{i}_{a}=0 \Rightarrow \varphi^{i}_{da}=0$)

where:

$g_a(\ldots)$= function composing probability choices of decisional unit components into one unitary choice;

φ^{i}_{da} = probability that decisional unit $d(\in \Phi_a)$ locate activity $a(\in X)$ in location $i(\in L)$;

$W^{da}_{j\delta}$ = "weight" of element $\delta(\in d)$, influenced by root $j(\in L)$, in the overall choice of decisional unit $d(\in \Phi_a)$ regarding the location of activity $a(\in X)$;

$\mathbf{W}^{da}_{..}$ = array of $W^{da}_{j\delta}$'s $\forall \delta \in Dd$, $\forall j \in L$; $\mathbf{W}^{da}_{..} = [\ldots, W^{da}_{j\delta}, \ldots]^T$

$\mathbf{p}^{\cdot i}_{da}$ = array of $p^{ji}_{\delta a}$ $\forall \delta \in Dd$, $\forall j \in L$; $\mathbf{p}^{\cdot i}_{da} = [\ldots, p^{ji}_{\delta a}, \ldots]^T$

$\mathbf{r}^{\cdot}_{da}$ = array of $r^{j}_{\delta a}$'s $\forall \delta \in Dd$, $\forall j \in L$; $\mathbf{r}^{\cdot}_{da} = [\ldots, r^{j}_{\delta a}, \ldots]^T$

The probability distribution of spatial root ($r^{j}_{\delta a}$) is generally dependent on the spatial distribution of activities. In general analytical terms:

$$r^{j}_{\delta a} = r^{j}_{\delta a}(\mathbf{E}_x, \mathbf{E}_y) \quad \Rightarrow \quad \mathbf{r}^{\cdot}_{da} = \boldsymbol{r}_{da}(\mathbf{E}_x, \mathbf{E}_y); \tag{7}$$

It is well known that, within the random utility theory, the systematic utilities ($V^{ji}_{\delta a}$) are generally made up as linear combinations (through proper parameters) of "attributes". Attributes depend, in general, on both the "endogenous" and "exogenous" location of activities. For instance, one could consider a given zone more attractive than other ones, in residential location choice, because it has a better active-accessibility to leisure facilities. In turn, active-accessibilities to leisure facilities depend on location of leisure facilities, some of these being "exogenous" while other "endogenous". Formally:

$$\forall\, a \in X,\ \forall\, d \in \Phi_a,\ \forall\, \delta \in D_d,\ \forall\, i \in L,\ \forall\, j \in L \qquad V^{ji}_{\delta a} = V^{ji}_{\delta a}(\mathbf{E}_x, \mathbf{E}_y, \mathrm{H}) \tag{8}$$

Urban cost functions could also depend on the spatial distribution of activities, both endogenously and exogenously located, it results that:

$$\forall\, a \in X,\ \forall\, d \in \Phi_a,\ \forall\, \delta \in D_d,\ \forall\, i \in L,\ \forall\, j \in L \qquad V^{ji}_{\delta a} = V^{ji}_{\delta a}(\mathbf{E}_x, \mathbf{E}_y, H(\mathbf{E}_x, \mathbf{E}_y)) \tag{9}$$

Finally, it is possible to define "location maps" ($\boldsymbol{m}_a$'s) formalised as:

$$\mathbf{ee}_a = \boldsymbol{m}_a(\mathbf{V}^{\cdot\cdot}_{.a}, \mathbf{r}^{\cdot}_{.a}) = \left[\ldots, \sum_{d \in \Phi_a} q_{ad} \varphi^{i}_{da}, \ldots\right]^T = \left[\ldots, \sum_{d \in \Phi_a} q_{ad}\, g_a(\boldsymbol{p}^{\cdot i}_{da}(\mathbf{V}^{\cdot\cdot}_{.a}), \mathbf{W}^{da}_{..}, \mathbf{r}^{\cdot}_{da}), \ldots\right]^T \quad \forall a \in X \qquad \text{s.a.}: \mathbf{ee}_a \leq \mathbf{u}_a - \mathbf{ex}_a \tag{10}$$

where:

$\mathbf{V}^{\cdot\cdot}_{.a}$ = array of $\mathbf{V}^{j\cdot}_{\delta a}$'s $\forall j \in L$, $\delta \in D_d$ and $d \in \Phi_a$ $\qquad \mathbf{V}^{\cdot\cdot}_{.a} = [\ldots, \mathbf{V}^{j\cdot}_{\delta a}{}^T, \ldots]^T$

$\mathbf{r}^{\cdot}_{.a}$ = array of $\mathbf{r}^{\cdot}_{da}$'s $\forall d \in \Phi_a$ $\qquad \mathbf{r}^{\cdot}_{.a} = [\ldots, \mathbf{r}^{\cdot}_{da}{}^T, \ldots]^T$

Aggregation issues

In the previous the model has been formalised in a general disaggregate way. For instance households are considered as decisional units. The amount q_{ad} can be viewed as "one residential unit", each different household chooses among all residential units built in the study area. The decisional units (d's) are each household and the set (L) of locations contains all built residential unit. The amount of residential activity located (ee^{i}_{a}) gets a value from 0 to 1 that is the probability it is occupied by effect of endogenous mechanism, analogously for ex^{i}_{a} (exogenous mechanism). Each element of the constraint-array (u^{i}_{a}) gets unitary value (each residential unit can be occupied by only one household). This specification is disaggregated with respect to both "decision makers" and "locations". Such a completely disaggregate approach is not suitable in most practical cases, mainly because of its complexity and the lack of the required data disaggregation. A first possible aggregation is "spatial aggregation". The study area can be divided into "zones". Each zone is one of the possible housing locations and L is the set of the identifying indexes of all zones. Assume "b" is the identifying index of the activity "building residential units" and ex^{i}_{b} is the amount of residential units built in a given zone (i) or, rather, a proper "size-function of the built residential (size functions are a way to aggregately take into account each element of a set, see Ben Akiva and Lerman, 1985). For sake of simplicity assume that activity b is completely exogenous ($b \in Y$, $b \notin X$), that is the choice of building residential units is not simulated and the amount of built residential units is fixed and "a priori" known. Say "a" is the residential activity (like in the previous disaggregated example); ex^i_a is the amount of residential units occupied in zone i by effect of exogenously (non-simulated) choices and ee^i_a is the amount of residential units occupied in zone i by effect of endogenous choices. The set of decision-maker (Φ_a) is each household in the study area. A generic element of the constraint vector (u^i_a) is the (known) amount of residential units built in zone i. In such a case the model specified is "spatially aggregated" but disaggregate in "decision-makers". The modelling approach can also be aggregate in both decision-makers and space. Assume, for instance, to consider as decisional unit a sort of "average household" that is in some way representative of all the families of the study area. In such a case, the amount q_{ad} of residential units to be endogenously located represents all the residential units that all the "real" households of the study area have to locate. Set Φ_a contains only one element: the "average household". The aggregation level of the decision-makers can be, of course, intermediate between "totally aggregated" and "totally disaggregate". For instance, one could consider some "classes of households": say low income, medium income, high income, etc. For each of these classes an average household represents all the households of the class, a different amount of residential units (q_{ad}) can be consistently associated to each class (or "segment") and the set Φ_a contains so many decisional units as many segments. Such an approach could be defined "partially disaggregate" or, more rigorously, "segmented".

The interaction model

In general, the location of each activity depends on the location of all others, see eqns 8 and 10. In this section such an interaction between activities will be formalised.

In real world activity relationships are dynamic: changes in location of a generic activity induce some "reacting" mechanisms in location of all "dependent" activities. The "reacting" activities induce, in turn, more reactions that are progressively and dynamically propagated, eventually leading to circular dynamic dependencies. Different interaction mechanisms, related to the different activities, can have different dynamics. The time-dependence of the "reaction" of a given activity could be different than for other activities; dynamics is not synchronous and the whole system dynamics results from the different dynamic interactions between activities. After some time the whole system could reach again a steady state, generally different from the starting one. Formally, it can be stated that:

$$\forall\ a \in X, \quad \mathbf{ee}_a^{\mathbf{t}} = m_a(\widetilde{\mathbf{V}}^{\cdot\cdot}{}_{.a}[t],\ \hat{\mathbf{r}}^{\cdot}{}_{.a}[t]) \tag{11}$$

Where t is a generic time period (for instance an year) and $\widetilde{V}^{ji}{}_{\delta a}[t]$ and $\hat{r}^{j}{}_{\delta a}[t]$ are respectly the value of the systematic utility and the probability distribution of spatial roots determining the location map at time t.

To define $\widetilde{V}^{ji}{}_{\delta a}[t]$'s and $\hat{r}^{j}{}_{\delta a}[t]$'s proper "smoothing functions" are used:

$$\widetilde{V}^{ji}{}_{\delta a}[t] = fv_a\,(V^{ji}_{\delta a}(E_x^{t0}, E_y^{t0}, H(E_x^{t0}, E_y^{t0})),\ \dots,\ V^{ji}_{\delta a}(E_x^{t}, E_y^{t}, H(E_x^{t}, E_y^{t}))) \tag{12}$$

$$\hat{r}^{j}{}_{\delta a}[t] = fr_a\,(\ r^{j}_{\delta a}(E_x^{t0}, E_y^{t0}),\ r^{j}_{\delta a}(E_x^{t1}, E_y^{t1}),\ \dots,\ r^{j}_{\delta a}(E_x^{t}, E_y^{t})\ ) \tag{13}$$

For example:

$$\widetilde{V}^{ji}{}_{\delta a}[t] = \beta\, V^{ji}{}_{\delta a}(E_x^{t}, E_y^{t}, H(E_x^{t}, E_y^{t})) + (1-\beta)\ \widetilde{V}^{ji}{}_{\delta a}[t]; \qquad \widetilde{V}^{ji}{}_{\delta a}[t=0] = V^{ji}{}_{\delta a}(E_x^{0}, E_y^{0}, H(E_x^{0}, E_y^{0}))$$

with $\beta \in R^+$

Smoothing functions used for calculating $\widetilde{V}^{ji}{}_{\delta a}[t]$'s can be different than those for $\hat{r}^{j}{}_{\delta a}[t]$'s, as well as different smoothing functions can be associated to different activities and/or decision-makers. Different smoothing functions allow different dynamics to be simulated.

It is worth noting that in previous eqns (12) and (13) the spatial root probability distribution has been assumed to be independent on system performance (H) and the cost functions have been considered only dependent on the current exogenous and endogenous activity locations. Even if in general it could be hypothesised that also cost functions are dependent on previous states of system configuration, in practice a common assumption is widely made that cost functions dynamics is much faster than other interactions, so that it is considered to be instantaneous. It is also usual, in the hypothesis that a discrete time approach is used, to assume that from a generic time t-1 to time t not all the activities are relocated:

$$\forall\, a \in X, \quad \mathbf{ee}^t_a = \lambda_a\, \boldsymbol{m}_a(\, \widetilde{\mathbf{V}}^{\cdot\cdot}{}_{.a}[t],\ \hat{\mathbf{r}}^{\cdot}{}_{.a}[t]\,) + (1-\lambda_a)\, \mathbf{ee}^{t-1}{}_a \tag{14}$$

The interaction mechanism between to or more activities or, rather, the whole system interaction mechanism can be investigated also with reference to steady state condition(s). Often only the final re-equilibrated state(s) (if any) keeps the interest of the analysts. An equilibrium approach can be used to this aim. In our case the fixed point representative of the equilibrium can be stated through the following system of equations, where the role of the "choice model" (location maps m_a's) is distinguished from the role of the "supply model" (V's and r's):

$$\forall a \in X \begin{cases} \mathbf{ee}^*_a = \boldsymbol{m}_a(\mathbf{V}^*_a, \mathbf{r}^*_a) = \left[\ldots, \sum_{d \in \Phi_a} q_{ad}\, g_a(\, \boldsymbol{p}^{\cdot l}_{da}(\mathbf{V}^{\cdot\cdot\,*}_{.a}), \mathbf{W}^{da}_{..}, \mathbf{r}^{\cdot\,*}_{da}), \ldots \right]^T \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{s.a.}: \mathbf{ee}^*_a \le \mathbf{u}_a - \mathbf{ex}_a \\ \forall d \in \Phi_a; \forall \delta \in D_d; \forall i,j \in L; \begin{cases} V^{*ji}_{\delta a} = V^{ji}_{\delta a}(\mathbf{E}^*_X, \mathbf{E}^*_Y, H(\mathbf{E}^*_X, \mathbf{E}^*_Y)); & \mathbf{E}^*_X = [\ldots, \mathbf{ee}^*_a, \ldots]^T \\ r^{*j}_{\delta a} = r^{*j}_{\delta a}(\mathbf{E}^*_X, \mathbf{E}^*_Y); & \mathbf{E}^*_Y = [\ldots, \mathbf{ee}^*_a, \ldots]^T \end{cases} \end{cases} \tag{15}$$

Which is equivalent to the following fixed-point formulation:

$$\forall\, a \in X, \quad \mathbf{ee}^*{}_a = \boldsymbol{m}_a(\ \boldsymbol{V}^{\cdot}{}_{.a}(\mathbf{E}^*{}_x, \mathbf{E}^*{}_y, \boldsymbol{H}(\mathbf{E}^*{}_x, \mathbf{E}^*{}_y), \boldsymbol{r}_{.a}(\mathbf{E}^*{}_x, \mathbf{E}^*{}_y)\ ) \tag{16}$$

Existence and uniqueness of the equilibrium problem can be studied within the general framework of the fixed-point theory stated by Cantarella (1997) for transportation systems. Given a fixed point problem, described in the form:

$$\left.\begin{matrix} \mathbf{y}^* = y(\mathbf{x}^*) \\ \mathbf{x}^* = x(\mathbf{y}^*) \end{matrix}\right\} \Rightarrow \quad \mathbf{y}^* = y(x(\mathbf{y}^*)) = f(\mathbf{y}^*) \qquad \mathbf{y}^* \in S_y;$$

existence is ensured if:

- function $f(\ldots)$ is defined in a non empty, compact and convex domain S_y;
- values of function $f(\ldots)$ belongs to a set $F \subseteq S_y$;
- function $f(\ldots)$ is continuous in set S_y.

Uniqueness is ensured if (but not necessarily if) contemporarily hold:

$$[x(y') - x(y'')]^T (y' - y'') \ge 0 \ \ \forall\, y', y'' \quad \text{or} \quad [x(y') - x(y'')]^T (y' - y'') \le 0 \ \ \forall\, y', y''$$

$$[y(x') - y(x'')]^T (x' - x'') < 0 \ \ \forall\, x', x'' \quad \text{or} \quad [y(x') - y(x'')]^T (x' - x'') > 0 \ \ \forall\, x', x''$$

In our case:

the role of $\mathbf{y} = \boldsymbol{y}(\mathbf{x})$ is played by $\mathbf{E}_X = [\ldots,\ \boldsymbol{m}_a(\ \boldsymbol{V}^{\cdot}{}_{.a}, \boldsymbol{r}_{.a}\ )\ , \ldots]^T$;

the role of $\mathbf{x} = \boldsymbol{x}(\mathbf{y})$ is played by $\mathbf{V}^{\cdot\cdot}{}_{.a} = \boldsymbol{V}^{\cdot}{}_{.a}(\mathbf{E}_X, \ldots)$ and $\mathbf{r}^{\cdot}{}_{.a} = \boldsymbol{r}_{.a}(\mathbf{E}_X, \ldots)$; $\forall a \in X$

the role of S_y is played by $S_{ee} = \left\{ \mathbf{E}_x : \forall a \ \left[\begin{matrix} \mathbf{0} \le \mathbf{ee}_a \le \mathbf{u}_a - \mathbf{ex}_a \\ \mathbf{1}^T \mathbf{ee}_a = \sum_{d \in \Phi_a} q_{ad} \end{matrix} \right] \right\}$

With respect to existence conditions it should be noted that:

- S_{ee} derives from equalities and inequalities such that it is convex, closed and bounded;
- function $[\ldots,\ \boldsymbol{m}_a(\boldsymbol{V}^{\cdot}{}_{.a}(\mathbf{E}_X, \ldots), \boldsymbol{r}_{.a}(\mathbf{E}_Y, \ldots))\ , \ldots]^T$ is composed by functions:
  - $p^{ji}_{\delta a}$, which are usually continuous, because of use of probabilistic models;
  - $V^{ji}_{\delta a}$, which are usually continuous;
  - $r^j_{\delta a}$, usually continuous
  - g_a;

- if constraints ($0 \leq ee_a \leq u_a - ex_a$) are in some way embedded within functions $m_a(\ldots)$, the values of function $[\ldots,\ m_a(\ V^{\cdot}{}_{.a}(E_x, \ldots),\ r_{.a}(E_y, \ldots)\ )\ ,\ \ldots]^T$ are ensured to belong to a set $F \subseteq S_Y$, otherwise this property is not automatically ensured.

Finally, existence critically depends on continuity of functions g_a's.

Uniqueness conditions are still an open issue. Author's feeling is that it is not difficult to prove that, with reference to our case, $[x(y') - x(y'')]^T (y' - y'') \geq 0$. However it is much more difficult to prove, in general cases, that $[y(x') - y(x'')]^T (x' - x'') < 0$, since it is strictly related to the use of urban cost functions, both related to space availability for location of activities and on transport system congestion. In fact these are the only aspects of the urban system that play a role against activity concentration.

A SIMPLIFIED SPECIFICATION AND APPLICATION

In this section a simplified operative version of the general location model described in previous section is presented. It will be applied to the city of Naples and calibration and validation issues will be also addressed.

Model specification

The study area has been split in 39 zones, 27 for the city of Naples and 12 for neighbouring. Activities are as follows (Y= {1, 2, 3, 4, 5, 6, 7} and X= {4, 5, 6, 7}):

Table 1 - Considered urban activities

Index	Activity Description	Endog.	Units of measure
1	Industries	N	Employment amount
2	Manufacture firms	N	Employment amount
3	Educational (and other non-diffuse services)	N	Employment amount
4	Wholesales	Y	Employment amount
5	Retails	Y	Employment amount
6	Services	Y	Employment amount
7	Housing	Y	Number of households

As "urban cost functions" only transportation costs are considered, H= $\{h_1(\ldots)\}$.

The set of possible activity locations is L= {1, 2, ..., 39}.

A "market segmented" approach is used. Decisional unit of a given segment represents (in average) all "real" decisional units with similar characteristics:

Table 2 - Decision-maker definitions

Activity	Activity definition	Dec. unit	Dec. units definition
1, 2, 3	Production, educational and other non-diffuse	NONE	NONE - exogenous -
4	Wholesales	1	All wholesale firms of the study area
5	Retails	2	All retail firms of the study area
6	Services	3	All service firms of the study area
7	Housing	4	All households with low-income
7	Housing	5	All households with medium-income
7	Housing	6	All households with high-income

It results that sets Φ_a's are:

$\Phi_1 = \Phi_2 = \Phi_3 = \varnothing$, since 1,2 and $3 \notin X$

$\Phi_4 = \{1\}$; $\Phi_5 = \{2\}$; $\Phi_6 = \{3\}$; $\Phi_7 = \{4, 5, 6\}$

In this simplified approach, the hypothesis is made that household can be represented only by an element, which is the "best-job employed" among the components of the family. Firms' management is considered to be not relevant for location choices. It results that each decisional unit contains a single element. In the following, for sake of simplicity, decisional units and elements will be used as synonymous. It is worth nothing that, in this case, it makes not sense to define $W^{da}_{j\delta}$'s and $g_a(...)$'s and that:

$\varphi^i_{da} = \sum_j p^{ji}_{\delta a} \times r^j_{\delta a}\ \forall a \in X,\ \forall \delta = d \in \Phi a,\ \forall i \in L,\ \forall j \in L$ (17)

The amount of activity that each decisional unit can locate in each zone is shown in the following table:

Table 3 - Amounts of activities to be located

Dec. unit	Dec. unit definition	Amount of activity to be located (q_{ad}'s)	*(ACTIVITY)*
1	Wholesale firms	16627 of employment (*3265* low+*12064*medium+*1297*high)	4
2	Retail firms	77745 of employment (*15269*low+*56411*medium+*6066* high)	5
3	Services	238297 of employm. (*46801*low+*172908*medium+*18594*high)	6
4	Households. (low-income)	184918 houses	7
5	Househ. (medium-income)	610757 houses	7
6	Households. (high-income)	82171 houses	7

The amount of exogenously located activities (ex^i_a's) is shown in the following table:

Table 4 - Amount of exogenously located activities (ex^i_a's

	Zone	Industry ex^i_1	Manuf. ex^i_2	Educ. ex^i_3		Zone	Industry ex^i_1	Manuf ex^i_2	Educ. ex^i_3
1	Posillipo	78	195	431	21	S. Pietro	1736	780	423
2	Bagnoli	1925	1555	1173	22	Doganel.	370	558	486
3	Fuorigrot	741	1600	2649	23	Poggior.	453	1064	457
4	Pianura	160	993	977	24	Industr.	2631	1816	203
5	Soccavo	159	1173	1549	25	Barra	2247	1070	165
6	Camald	68	125	332	26	Ponticelli	292	985	1058
7	M. d'oro	349	701	1471	27	S. Giov.	1890	1006	1824
8	Vomero	148	431	1581	28	*Pozzuoli*	*5828*	*2341*	*3488*
9	Chiaia	261	2158	1649	29	*Giugliano*	*2086*	*2937*	*3155*
10	Munic.	2178	2526	1872	30	*Aversa*	*3852*	*1984*	*2744*
11	Spagnoli	104	240	880	31	*Casoria*	*5803*	*6899*	*2748*
12	Duomo	531	610	4209	32	*Afragola*	*3018*	*8737*	*5178*
13	Sanità	80	387	1124	33	*Pomigl.*	*17320*	*4410*	*3857*
14	Carlo III	242	809	1872	34	*Somma*	*1734*	*1903*	*2093*
15	Mercato	259	723	1736	35	*Nola*	*1611*	*3831*	*3157*
16	Aminei	98	409	1047	36	*S. Gius..*	*786*	*2293*	*2027*
17	Miano	296	469	854	37	*Pompei*	*2842*	*2060*	*2988*
18	Chiaiano	101	329	1177	38	*Tor del G*	*2341*	*3561*	*7108*
19	Scampia	317	49	597	39	*Sorrento*	*3572*	*5407*	*5530*
20	Second.	226	547	847		TOT	68733	69671	76716

Amounts of endogenously located activities are defined as follows:

- ee^i_4 , ee^i_5 and ee^i_6 = employment amount in wholesales, retails and services that the model endogenously locates in zone i;
- ee^i_7 = number of households that the model endogenously locates in zone i (this is also the number of occupied houses in zone i).

The only "spatially rooted" decisional units are considered to be households. It is assumed that they are influenced (in housing choices) by their workplaces. So, j is the "root" that influences decision-makers 4, 5 and 6 (households) in location of activity 7 (housing). The hypothesis is made that probability distribution of spatial roots is equal to probability distribution of households' workplaces. In other term, the probability that a given root (j) influences household component $\delta \in \{4, 5, 6\}$ in housing choices (a=7) is given by the probability that the household element works in zone j:

$$r^j_{\delta a} = \text{Prob}[\delta \text{ employed in } j] = \frac{employed(j)}{\sum_w employed(w)} = \frac{\sum_{b=1}^{3} ex^j_b + \sum_{b=4}^{6} ee^j_b}{\sum_{w \in L}\left(\sum_{b=1}^{3} ex^w_b + \sum_{b=4}^{6} ee^w_b\right)} \tag{18}$$

So it results that $r_{4,7} = r_{5,7} = r_{6,7} = r_{4,7}(ex_1, ex_2, ex_3, ee_4, ee_5, ee_6)$.

Such a $r^{j}_{\delta a}$'s probability distribution is hypothesised to be equal for each $\delta \in \{4, 5, 6\}$ (income-independent), so the implicit assumption made is that the proportions between "low-profile", "medium-profile" and "high-profile" employees are the same in wholesales, retail and service firms; moreover these are independent on locations. It is also worth noting that the overall operative modelling specification here presented does not consider any type of workplace choice. In general, firms locate employment but the corresponding employees depend on people's workplace choices. The latest have been not considered: firms locate employment and so many workers result as many needed by firms. Moreover, for each wholesale, retail and service unit the average number of required employees is, respectively 8.46 (1.67 low+6.14 medium+0.65 high), 1.93 (0.38 low+1.4 medium+0.15 high) and 5.04 (0.99 low+3.66 medium+0.39 high).

For each decisional unit a logit-type choice model is used. For each considered choice model (that is for each considered decisional unit) and for each choice alternative (and possibly for each spatial root) a systematic utility function has to be identified. It is common use in random utility theory to consider systematic utilities as a linear combination of "attributes" ($T^{ji}_{\delta a,k}$) through "parameters" ($\beta_{\delta a,k}$):

$$V^{ji}_{\delta a} = \sum_k (\beta_{\delta a,k} \times T^{ji}_{\delta a,k}) \qquad (19)$$

Attributes depend on endogenous and exogenous amounts of located activities, so utilities also depend on. The attributes related to each choice model are listed in the following table:

Table 5 - Attributes of location choices

Decision Maker	**Activity Index**	**Rooted choice**	**Systematic Utility Attributes**					
			1st	2nd	3rd	4th	5th	6th
1	4	N	Passive access	Avail. locations	Aver. floor price	Dummy	---	---
2	5	N	Passive access	Avail. locations	Aver. floor price	Dummy	---	---
3	6	N	Passive access	Avail. locations	Aver. floor price	Dummy	---	---
4	7	Y	Active access.	Avail houses	Aver. house prices	Ambient	Work-trip costs	Saturation
5	7	Y	Active access.	Avail houses	Aver. house prices	Ambient	Work-trip costs	Saturation
6	7	Y	Active access.	Avail houses	Aver. house prices	Ambient	Work-trip costs	Saturation

Note that passive accessibilities of a given zone from retails of all zones (location of activity 4) depend on both transportation costs and amount of located retails in each zone. Similarly, passive accessibilities from households (location of activities 5 and 6) depend on transportation costs and household locations. Active accessibilities to retails influence residential choices for low, medium and high income decision makers. Active accessibilities depend on transportation costs and amount of located retails in each zone. Location availability for activities 4, 5 and 6 have been exogenously treated in the operative specification of the model (building mechanism not treated), in particular "size functions" have been used. Similarly, exogenously estimated size-functions for house (decision makers 4, 5 and 6) have been used, in particular for low incom decision makers both public and private housing market is considered while for medium and high income decision makers only private housing market is considered. Average location costs are exogenously estimated for activities 4, 5, 6 and 7 (housing and land-use market are not explicitly treated). Demand/Supply equilibrium in housing market and floor rents is approximately estimated through saturation variables. Environmental aggregate variables are also exogenously estimated for each zone, they depend on environmental condition, crime, leisure facilities availability and others.

Attributes related to saturation of available spaces have been directly included into systematic utility functions. So, all the elements of constraints' arrays could be considered set to infinity value (in practice no-constraints). As it has been stated in previous section, this ensures existence of the equilibrium solution and should leads to benefits with regard to equilibrium uniqueness. It results that:

$V^{i}_{14} = V^{i}_{14}(Acc\_W_i(\mathbf{ee}_5, \mathbf{h}_1), Size\_W_i, Price\_W_i, Dummy\_W_i) \Rightarrow \mathbf{ee}_4 = m_4(\mathbf{ee}_5, \mathbf{h}_1);$

$V^{i}_{25} = V^{i}_{25}(Acc\_RS_i(\mathbf{ee}_4, \mathbf{h}_1), Size\_R_i, Price\_R_i, Dummy\_R_i) \Rightarrow \mathbf{ee}_5 = m_5(\mathbf{ee}_4, \mathbf{h}_1);$

$V^{i}_{36} = V^{i}_{36}(Acc\_RS_i(\mathbf{ee}_7, \mathbf{h}_1), Size\_S_i, Price\_S_i, Dummy\_S_i) \Rightarrow \mathbf{ee}_6 = m_6(\mathbf{ee}_7, \mathbf{h}_1);$

and

$V^{ji}_{47} = V^{ji}_{47}(Acc\_H_i(\mathbf{ex}_3, \mathbf{ee}_5, \mathbf{ee}_6, \mathbf{h}_1), Size\_HL_i, Price\_HL_i, Dummy\_H_i, h_1^{ji}, StHL(\mathbf{ee}_7)).$

$V^{ji}_{57}= V^{ji}_{57}(Acc\_H_i(\mathbf{ex}_3, \mathbf{ee}_5, \mathbf{ee}_6, \mathbf{h}_1), Size\_HMH_i, Price\_HMH_i, Dummy\_H_i, h_1^{ji}, StHMH(\mathbf{ee}_7))$

$V^{ji}_{67} = V^{ji}_{67}(Acc\_H_i(\mathbf{ex}_3, \mathbf{ee}_5, \mathbf{ee}_6, \mathbf{h}_1), Size\_HMH_i, Price\_HMH_i, Dummy\_H_i, h_1^{ji}, StHMH(\mathbf{ee}_7))$

$\mathbf{r}_{4,7} = \mathbf{r}_{5,7} = \mathbf{r}_{6,7} = r_{4,7}(\mathbf{ex}_1, \mathbf{ex}_2, \ldots, \mathbf{ee}_6) \Rightarrow \mathbf{ee}_7 = m_7(\mathbf{ex}_1, \mathbf{ex}_2, \ldots \mathbf{ee}_6, \mathbf{ee}_7, \mathbf{h}_1)$

In order to complete the definitions of the systematic utility parameters (β's) have to be associated to each attribute: values assumed after calibration are shown in the following table:

Table 6 - Attributes of location choices

Dec. Maker	Activity	SYSTEMATIC UTILITY PARAMETERS					
		1st	2nd	3rd	4th	5th	6th
1	4	0.25473	0.74955	-0.25300	0.99186	---	---
2	5	0.24190	0.89323	-0.41075	0.35622	---	---
3	6	0.50590	0.83630	-0.24900	1.11130	---	---
4	7	0.89650	1.05090	-0.24300	0.04570	1.8234	-1
5	7	0.03000	1.00570	-0.31500	0.03040	1.5941	-1
6	7	0.01000	1.10720	-0.12500	0.05670	2.9168	-1

Calibration of the model

The calibration of the parameters has been obtained by comparing model results against the observed urban system pattern. The basic idea is that the model should predict the observed distribution of endogenous activity locations. Parameters are tuned until a satisfying reproduction is obtained. The sum of squared errors between predicted and observed data has been minimised. In formal terms:

$$\forall\, a \in X \qquad \beta_a = \operatorname{argmin}\left([\hat{m}_a(\beta_a) - \hat{\mathbf{ee}}_a]^T \times [\hat{m}_a(\beta_a) - \hat{\mathbf{ee}}_a] \right) \tag{20}$$

where β_a , $\hat{\mathbf{ee}}_a$, $\hat{m}_a(\beta_a)$ and ($\hat{\mathbf{E}}_x=[\ldots, \hat{\mathbf{ee}}_a^T, \hat{\mathbf{ee}}_b^T, \ldots]^T$ are respectively: the set of parameters involved in the location map of activity $a \in X$; the observed location array of activity $a \in X$ and the predicted location array. Parameters β_a influence systematic utilities ($V^{ji}_{\delta a}$), which, in turn, influence m_a through choice functions $p^{ji}_{\delta a}$.

The proposed calibration approach directly operates on m_a instead of (more rigorously) calibrating on $p^{ji}_{\delta a}$. From this point of view the calibration approach can be defined "aggregate". In order to improve the precision of the aggregate calibration, modelled choices have been also calibrated against some transport demand data. In particular, location of residences and workplaces are strictly related to work-purpose mobility. Under the hypothesis that the average number of work-trip made by each household/worker is constant over the whole study area, probability ($p^{ji}_{\delta a}$) that households working in a given zone j locate their residence in zone i is linearly related to the number of work-trips from origin i to destination j. Since some data on origin/destination demand for systematic trips (mostly work-trips) is known, another source of data for calibration result available.

After the calibration the overall model has been validated. In fact, each location map has been independently calibrated in such a way that location of each activity approximates as well as possible its observed location but this is computed (in calibration phase) depending

on observed locations of all other activities. So, it is not ensured that the whole model gives good results when it runs with unknown endogenous locations.
In our case validation leads to good results. The "less performing" location model seems to be housing choices of high-income households, however corresponding RMSE value is 8.61% that could be accepted .The scatter diagram of observed vs. predicted housing choices is shown in the following figure.

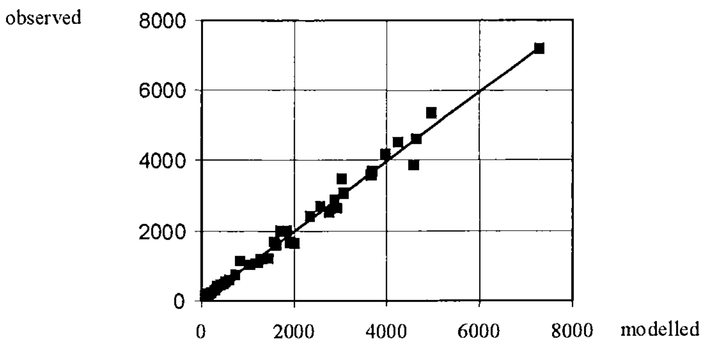

Figure 1 - Observed vs. modelled high-income housing choices

A simulated scenario

The example here given is not aimed to define "good practice" in transportation and/or land-use planning, or to evaluate the effects of "real" policies or, finally, to draw out general learning on transportation and/or land-use impacts. The aim only is to show that the presented model is able to keep the impacts of transportation and/or land-use planning. In particular, the effects of the construction of major new infrastructures are here presented. No Travel Demand Measures (which should complement infrastructural ones) nor land-use policies have been taken into account in this example. The urban system of the Naples' Municipality will be subject in future years to a large enhancement of its transportation system. Particular attention will be done to intermodality (from actual 5 interchanging nodes to 18 in the future and from actual 1 Park-and-Ride node to 16 in the future). Moreover a decisive enhancement in railway transport will be made (from current 25 km to future 112 km of lines). Also 4 new tramway lines will be built.

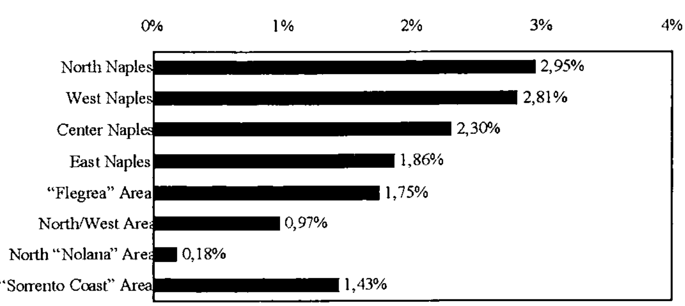

Figure 2 - Active accessibility changes

The previous figure shows the variation of active accessibilities to retails and services when a new equilibrium state is reached. It is evident that accessibility globally increases and it is also not surprising that accessibilities of the internal Naples zones show major increments. It is worth nothing that the simulated measures are only related to transportation system and only to Naples internal zones. Some impacts are shown in figure 3, where only housing location changes and aggregated employment location changes are shown. The overall effect is that, not surprising, Naples internal zones increase both in residents and in employment. The utility of housing in Naples relatively more increases for workers employed in Naples and remain relatively constant for workers employed outside Naples. Moreover firms that locate wholesales, retails and services in Naples gain more accessibility by a major number of person.

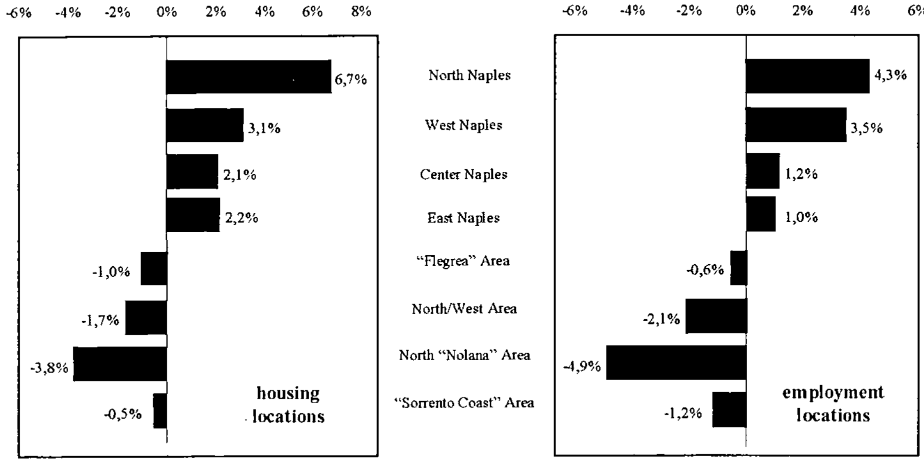

Figure 3 - Housing and employment changes

CONCLUSION AND FURTHER RESEARCH ISSUES

The main feature of the presented model is to establish a disaggregated theoretical architecture in which location models can be framed. Modelling stress is much more focused on "decision-makers" rather than on the aggregate and "ideal" concept of "activity interactions". The main aim is to apply a behavioural approach to decisions related to activity locations; this also is the main originality of the model.

The model formalisation also allows framing the model in the same well-established theoretical contest of transportation models (where behavioural approaches play an important role). Such a theoretical contest "shows the way" in order to deeply investigate theoretical properties of the model (such as: existence and uniqueness of equilibrium). Unfortunately, such an investigation seems to be more difficult than for transportation models, since a larger number of variables are involved and more complex relationships established. In particular, investigation of equilibrium uniqueness conditions seems to be the more urgent further direction of research.

One of the limits of the operative specification here presented is also one of the major challenges for the future. Calibration of the proposed specification, in fact, has faded out the formalisation effort for a behavioural and disaggregated model. Calibration has been done in an aggregate way so that the behavioural interpretation has been lost. In practice, descriptive relationships among variables have been established. The main difficulty for a fully disaggregate and behavioural calibration of the model is related to data availability. However, the article shows that the model can also be calibrated against "traditional" territorial data (observed activity locations) even if proper surveys and related estimation procedures (such as Likelihood estimates) should be used in order to keep disaggregate and behavioural properties.

REFERENCES

Alonso, W.(1964) **Location and land use**. Harvard University Press, Cambridge, Mass.

Ben Akiva, M. and Lerman, S. (1985) **Discrete Choice Analysis: Theory and application to travel demand**. MIT press, Cambridge, Mass.

Cantarella, G.E. (1997) A general fixed-point approach to multi-mode multi-user equilibrium assignment with elastic demand. **Transp. Science 31**, 107-128

Comune di Napoli, (1997), **Piano comunale dei trasporti – Relazione finale**. Assessorato alle infrastrutture di trasporto – Napoli, Italy.

De la Barra, T. (1989) **Integrated Land-use and transport modelling**. Cambridge University Press

Echenique, M. *et al.* (1995) Testing integrated transport and land-use models in the Helsinki metropolitan area. **Traff.Eng.&Control – January**, 20-23

Kim, T. J (1983) A combined Land-use transportation model when zonal travel demand is endogenously determined. T**ransp. Research B 17**, 449-462

Lowry, I. S. (1964) A model of metropolis. **RM-4125-RC**. The Rand Corporation, Santa Monica, Calif.

Mcgill, S. M. (1977) The Lowry model as an input-output model and its extension to incorporate full intersectoral relations. **Reg. Studies 11**, 337-354

Putmam, S.H. (1975) Urban land use transportation models: a state of art summary. **Transp. Research 9**, 187-202.

Wilson, A.G. (1970) **Entropy in urban and regional modelling**. Pion, London.