European Transport \ Trasporti Europei n. 31 (2005): 67-82

A model of residential location choice with endogenous housing prices and traffic for the Paris region

André de Palma [1*], Kiarash Motamedi [2], Nathalie Picard [2], Paul Waddell [3]

[1] *University of Cergy-Pontoise and Ecole Nationale de Ponts et Chaussées, France.*
[2] *University of Cergy-Pontoise, France.*
[3] *University of Washington, USA.*

Abstract

There is a growing interest in the development and the use of large-scale planning models. In this paper, we describe the first step of a project to integrate UrbanSim, a dynamic microsimulation land use model, and METROPOLIS, a dynamic traffic model. This is the first attempt, to our knowledge, to integrate a dynamic land use model and a dynamic traffic model. We briefly describe the two models and propose a unified framework for their integration. Within this integrated framework we develop a model of residential location choice, with endogenous housing prices and traffic. The study area for this research is the Ile-de-France (Paris region), for which we provide empirical results.

Keywords: Land use; Integrated model; Transportation modelling; Paris area.

Introduction

In metropolitan regions throughout the world, increasing population and urban expansion generate increased transportation congestion and rising housing prices. The need to coordinate land use policies with transportation investments has been widely recognized, but the task remains difficult for both technical and political reasons. Politically, the coordination of transportation and land use is difficult because land use decisions are controlled by local governments that by nature have a parochial mandate, whereas transportation investments are generally coordinated at a metropolitan level to ensure efficient coordination of the regional transportation network. Technically, the coordination of land use and transportation is challenging due to the lack of well-integrated models that provide a coherent behavioural basis to model not only the effects of changing patterns of locations of jobs and households and real estate

[*] Corresponding author: André de Palma (andre.depalma@u-cergy.fr)

investments on transport flows, but also the effect of changes in the transportation system and travel conditions on these location choices.

Though models that reflect the interaction between land use and transportation have been developed and used for at least three decades (Putman, 1983, de la Barra, 1990), the models have been characterized by a high degree of aggregation of space, agents, and time. Prior land use models have represented geography using a very aggregate zone structure, usually with 30 to a few hundred zones. Agents such as households and jobs have been aggregated to a small number of categories. But perhaps most importantly, chronological time has not been explicitly represented in prior land use models, in that they solve for an equilibrium with a given set of inputs, with no path dependence, and an assumption that all agents can adjust instantaneously, with no transaction cost. This approach requires making the assumption, for example, that the effect of building a major transportation facility in a given year will produce all of its effects on real estate development, location choices, and travel behaviour in that same year. Our approach, by contrast, avoids this assumption by representing the partial adjustment of households, firms and developers in annual steps of time, allowing the effects of a major shock such as a change in infrastructure to be spread over multiple years.

These restrictions are important constraints, and have led to recent innovations to overcome these simplifications to allow more behavioural realism in the modelling. This realism is important in making the modelling efforts responsive to current policy questions that require considerable behavioural resolution in order to represent the dynamic short-term and long-term effects of major transportation investments and their interaction with land use policies. Two models exemplify the recent trend towards microsimulation and dynamic temporal representation in the land use and transportation domain. UrbanSim is a simulation model developed since the late 1990's to simulate the spatial and temporal evolution of household location, job location, and real estate supply and prices using microsimulation to allow complete disaggregation in agents, locations, and the representation of time (Waddell et al, 2003). This model has been applied to numerous cities in the United States and Europe, but until now has been connected to traditional four-step travel models that provide static equilibrium traffic assignment, usually for only a small number of time periods during the day. METROPOLIS is a dynamic traffic assignment model that simulates evolving traffic conditions on large-scale networks over the course of a day, representing individual travellers (de Palma and Marchal, 2002).

The major innovation developed in this paper, in addition to the operational integration for the first time of dynamic microsimulation land use and traffic models, is in the treatment of two types of endogeneity in residential location choices. Residential location is clearly interdependent with housing prices, and we develop an econometric specification and estimation methodology that correctly accounts for this endogeneity. We also treat the endogeneity of travel times for work trips with residential location by coupling the land use and traffic models. The paper proceeds as follows. In the next section, we provide an overview of UrbanSim and METROPOLIS and the proposed integrated model architecture. Following this, we describe the region that serves as the basis for this application, the data used for model estimation, and the model specification and empirical results. We conclude with some interpretation of results and discussion of further research.

Integrated model system design

In this section we briefly describe the UrbanSim and METROPOLIS models and our proposed approach to integrate them.

UrbanSim: a path-dependent land use model

UrbanSim is a disaggregate land use model used to simulate the spatial and temporal evolution of land use and the locations of households and jobs within metropolitan areas. It has been developed at the University of Washington since 1996, and released under an Open Source license; see (Waddell et al., 2003). In 2005 a new Open Platform for Urban Simulation (OPUS) has been implemented to support further development of UrbanSim and incorporation of other simulation models and tools (Waddell et al., 2005). UrbanSim simulates year-to-year changes in real estate development and in the location of households and jobs for each geographical unit. Geography has typically been represented using grid cells as small as one hectare, though in the current application to the Ile-de-France we use 1300 Communes, or local municipalities.

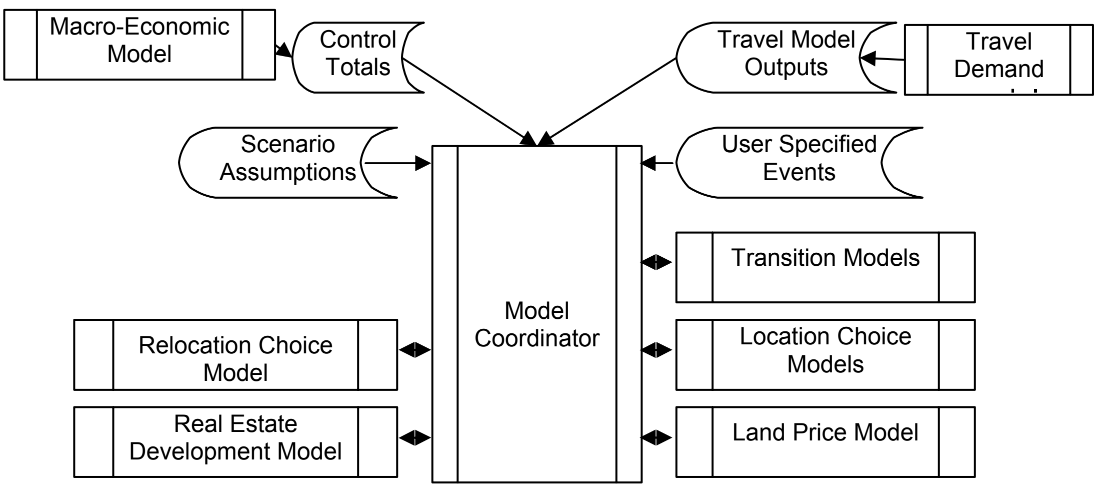

Figure 1: UrbanSim architecture.
Adapted from Waddell et al., 2003.

The principal modules in UrbanSim are presented schematically in Figure 1. Models of choice processes such as location of households, jobs and new real estate development use Discrete Choice Models (Multinomial Logit for standard version).

UrbanSim is typically interfaced with an external travel model system (normally a standard four-step travel model), which generates trip distribution and utility patterns used in UrbanSim to measure patterns of accessibility. Accessibility for the home to work and possibly other purposes is computed using a variety of alternative measures. One such measure uses the composite utility of travel from a particular origin to all destinations:

$$A_i^O = \sum_{j=1,J} E_j . e^{L_{ij}} \quad \text{or} \quad A_i^D = \sum_{j=1,J} P_j . e^{L_{ij}} \ ,$$

where $A^{O}{}_{i}$, is the accessibility of cell i as origin and $A^{D}{}_{i}$ as destination, P_j stands for the population that comes to the cell to work there, E_j represents the jobs to which the people go and L_{ij} is the logsum which is the surplus of the travellers.

The choice of agents (households and jobs) to relocate during a given year is modelled in the Relocation Choice Model, as a probability that depends on agent characteristics. Household age and income are the principal factors to predict differences in relocation rates, and employment sector is used to measure differing propensities to relocate jobs. New and moving agents choose locations from the existing available real estate in the Location Choice Models, using multinomial logit specifications.

Real estate development, or the construction of new housing and non-residential floor space, has typically been modelled in UrbanSim as a multinomial logit transition model, where we predict the probability that a particular location will experience one of many types of real estate development events in a year. This specification has been recently changed to reflect the real estate development process as a location choice for a developer with specialized projects.

Real estate prices are important in the model in that they capitalize locational amenities such as accessibility, and strongly influence the spatial distribution of households in the housing market and firms in the non-residential real estate market. In the current research, we use a semi-hedonic regression model that predicts housing prices as a function of location characteristics, demand and supply.

METROPOLIS: a dynamic transportation model

METROPOLIS is a fully dynamic transportation model that is particularly adapted for large networks. It is a mesoscopic event based model and uses a multi-agent methodology with a disaggregated representation of travellers. On the other hand, the supply system relies on a macroscopic formulation that computes travel time in function of the flow condition of the link. It has been developed since the 90's and its main application on the Paris region has been the QUATUOR project (THEMA/TT&R, 98-02). It models the mode, departure time and route choices. The Logit formula is used for these models. The dynamic assignment procedure can be deterministic or stochastic.

The generalized cost function is:

$$C(t_d) = \alpha tt(t_d) + \beta \max\left[0, t^* - t_a\right] + \gamma \max\left[0, t_a - t^*\right], t_a = t_d + tt(t_d).$$

The generalized cost function *C(.)* includes the schedule delay cost terms in $(t^*\text{-}t_a)$, where t^* is the desired arrival time. Moreover t_a and t_d denote arrival and departure times and *tt* is the travel time corresponding to the mode (private vehicle or public transport). The operator can enter some distributions for α (VOT), β and γ (penalties for arrival too late or too early) and t^*. These behavioural parameters are the only information that is necessary in addition to data required by the classical static traffic models, such as network topology, link characteristics and O-D matrices.

The other transportation modes (mainly public transit) are modelled in an aggregated and static way. For any pair of origin and destination zone centroids the travel time should be given in a matrix form. The trip cost is the sum of a constant part p_{PT} that represents the ticket fare or constant penalty of using public transport and a linear travel time-dependent part, $\alpha_{PT}*tt_{PT}$. We denote the part of private cars by P_{PV}, the generalized cost of private vehicles by GC_{PV} and the mode choice heterogeneity factor by μ_m. The mode choice is described by a binary logit model:

$$P_{PV} = \frac{\exp\left[-GC_{PV} / \mu_m\right]}{\exp\left[-GC_{PV} / \mu_m\right] + \exp\left[-C_{PT} / \mu_m\right]}, \text{ where } C_{PT} = \alpha_{PT} * tt_{PT} + p_{PT}.$$

The output of METROPOLIS that we mainly use in this project is the surplus for any traveller's category and for any O-D pair. As the departure time choice is modelled by a continuous Logit model, the surplus is given by:

$$L_{ij} = -\mu^T.\ln \sum_{k=PV,PT} \int_{T_0}^{T_1} \exp(-C_{ij}^k(u)/\mu^T)du,$$

where $C^k_{ij}(t)$ is the time dependent generalized travel cost between zones i and j and where μ^T denotes the departure time choice heterogeneity parameter (see above), PV represents private vehicles and PT represents public transit (see de Palma and Marchal, 2002 for details).

Integrated model architecture

We present in Figure 2 the architecture of the integrated system. The key information transferred between the traffic and the land use model is the travellers' surplus matrix. To make a complete loop, we should feed a revised O-D matrix to the traffic model that is based on the new geographical distribution of population and jobs. This cycle is reproduced by time step that can correspond to one or more years according to the evolution of transportation system conditions and projects.

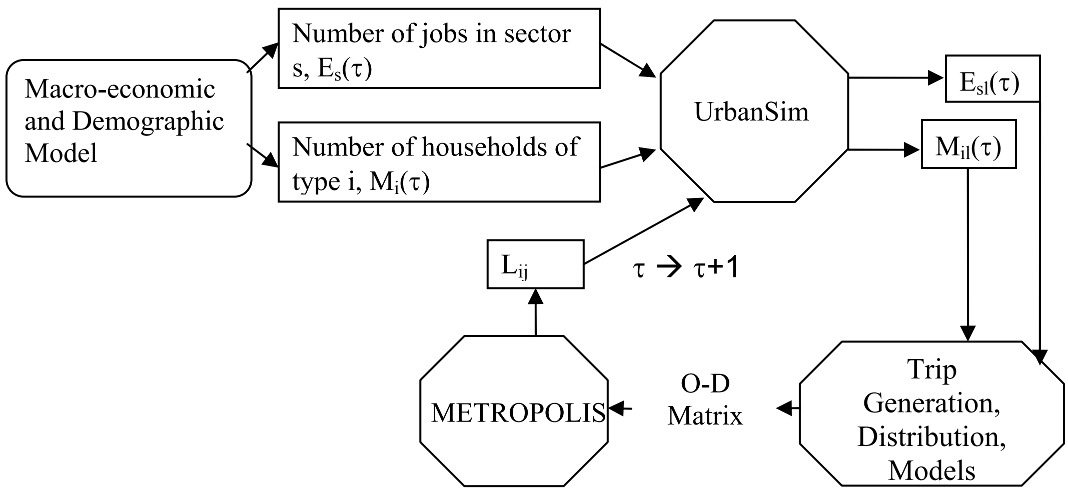

Figure 2: Architecture for the integrated model.

UrbanSim assigns locations (l) to jobs $E_s(\tau)$ and households $M_s(\tau)$ generated by the macro-economic model in year τ, resulting in $E_{sl}(\tau)$ and $M_{sl}(\tau)$, which are used as inputs to the travel model. Since the Origin-Destination (O-D) matrix is not directly provided by UrbanSim, a three step travel model is needed to build the O-D matrices that represent the trips generated by population and activities and their distribution on origins and destinations. This module develops an O-D matrix using three steps:

- Trip emissions and attractions (by zone and travel segment),
- Trip distribution (zone to zone by travel segment),
- Mode choice: private vehicles and public transport.

We use the specific demand model developed for the Paris region by IAURIF and which was calibrated with the last Global Transportation Survey in 2001.

Data produced by METROPOLIS cannot be used directly by UrbanSim. A data preparation module is developed to convert these results to logsums and travel times for use by UrbanSim. In previous applications of UrbanSim with four-step travel models, the computation time of the travel models has prohibited the coupling of the models, usually by running the travel model only once in every 5 simulation years. Travel model run times of 18 hours or longer for one simulation year are not uncommon, mostly due to the computational burden in the traffic assignment component. Due to the computational efficiency of METROPOLIS, however, we interface the models every simulation year, providing a significant improvement in the model realism over prior integrated model applications.

Descriptive analysis of the study area

The Ile-de-France

The Paris area, namely Ile-de-France Region, embraces Paris and its suburbs. The city of Paris has about 2 million inhabitants, on a regional total of 11 million. The total number of jobs is 5.1 million. It covers 4,610 sq. miles (12,000 sq. km). Ile-de-France Region occupies 2% of the surface of France and represents 19% of the population, 22% of the jobs and 29% of the GDP of the country. There are 3 administrative divisions in Ile-de-France: 1 “région”, 8 “départements” (counties) and 1300 “communes” (municipalities). In addition, we consider the 3 counties around Paris as close suburbs or “inner ring” and the 5 counties far away from Paris as far suburbs or “outer ring”.

The land use is composed of built-up areas (30%), green areas (20%) and rural areas (50%). The public transportation network consists of:

- A main radial railway network, especially the RER lines (high speed train service between Paris and the suburbs),
- A subway network that provides comprehensive and timely service in the city centre,
- A bus network to complement the rail services.

The road network is organised into a hierarchy that is densely interconnected and often congested. The principal road network of the region is composed of 590 km of motorways and 250 km of expressways, with a total of 4,500 lane-km. Road traffic flows attain the highest levels known all over the country. Despite the occasional rush-hour traffic jams, traffic conditions are on the whole remarkably good for a metropolis of this size, since the average duration of all car trips is 19 minutes, and commute trips by car average 25 minutes (EGT, 2001). The mode market shares for the home based work trips (2001) are: 50% private vehicles, 36% public transit and 14% bicycle or walk. Over the last twenty years, the public transportation mode share has decreased by 6% in the region, due principally to ongoing suburbanization of the region.

Socio-economic characteristics

Turning to a description of the socio-economic characteristics of the study area, and specifically focusing on households that have recently moved, Figure 3 presents the number of households according to the year of their last move. The mode is at 1998, the year just before the census. It should be noted that this cannot be seen as the distribution of how long the people live in housing units, since this duration is truncated here. These data confirm that many households remain in their locations for periods of ten years or longer, and support the argument that a partial adjustment approach is more plausible than a full, instantaneous adjustment to equilibrium.

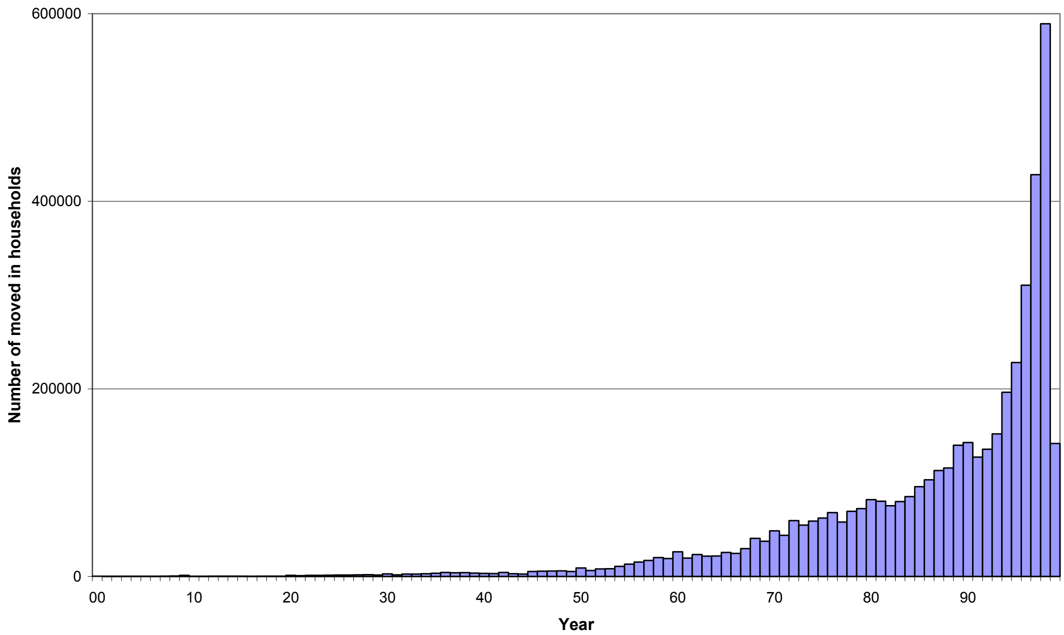

Figure 3: The distribution of the last move in year for the households living in Paris area in 1999.
Source: Census 1999.

The principle of multi-cored structure has been adopted for urban organisation to stem the rapid expansion of the agglomeration and to decentralize the jobs. So since 1965, the outer suburbs were structured around five poles, or “new cities”. Accordingly, 44% of the population surplus recorded in Ile-de-France between the 1975 and 1999 censuses, settled in these areas.

The analysis focuses on “recent movers”: households who settled or moved in the region recently, that is during year 1998. Among the 4,510,369 households living in the study area in March 1999, 589,355 moved during year 1998. Most of them (71%) are male headed. The “poor households” (that is, the 33% households in the region with the lowest per capita income, defined as household income divided by the square root of the number of persons in the household) are unevenly distributed in the region: only 26% of households living in district 78, located west of Paris; are poor, whereas this fraction goes up to 41% in district 93, located north-east of Paris. These same two districts contain the highest (38% in district 78) and the lowest (21% in district 93) proportions of rich households.

Single-person households are highly concentrated in Paris city (52% of households in Paris are single). Between 25 to 30% percent of households of all the counties have two members. The larger families are better represented in rural counties in the far suburbs. 25% percent of households have no working member. Among them, 28% percent live in Paris city. Near 50% of the families in far suburbs have two or more workers. Foreign households are concentrated in district 93 (19%), and are less represented in the larger ring (9%). 25% of households have a young head. They have a bigger share in Paris center and in district 92 (31% and 27%) and their part is uniform in other counties (23%).

Housing prices

Table 1 shows housing price data based on average prices of housing sales transactions for 1998 in each commune, using a weighted average of transactions of new and existing single-family houses and apartments. The data show important differences in average housing prices by district: prices are higher in Paris, intermediate in the close suburbs and decline in the more distant suburbs. In addition, prices are higher in the western part of the area than in the eastern part.

Table 1: Prices by district.

Sub-region	*District*	*Average*	*Standard Deviation*	*Minimum*	*Maximum*
Paris	75	294,500	165,241	83,939	694,375
Close Suburbs (inner ring)	92 (West)	247,556	205,038	66,966	1,198,950
	93 (North)	115,709	49,055	47,876	259,163
	94 (South)	144,098	74,603	53,356	373,499
Far away suburbs (outer ring)	78 (West)	135,122	65,714	38,112	373,815
	91 (South)	114,826	46,740	24,719	332,338
	95 (North)	104,375	41,670	25,154	241,692
	77 (East)	91,539	37,220	18,028	253,827

Source: Author’s computations from notaries’ database.

Accessibility

Figure 4 presents the average travel time in minutes from any city in the region by private vehicles and by public transit. The Paris boroughs are at the left of the figure and the cities in the farthest Parisian suburbs are at the right hand side. The correlation between private vehicle and public transit average travel times is 0.97.

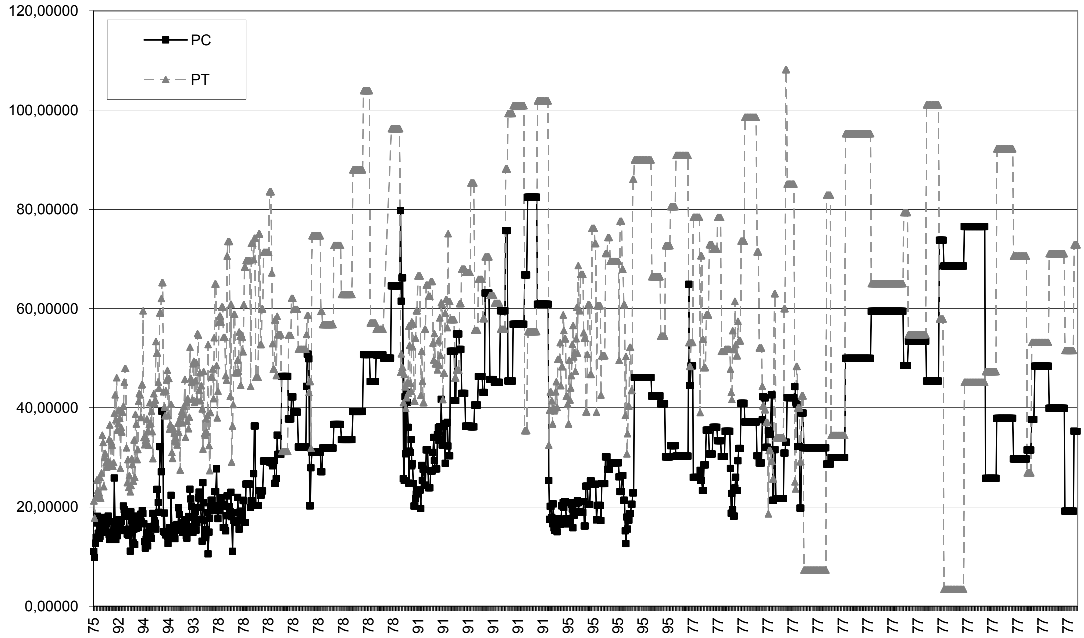

Figure 4: The average travel time for people travelling from any city.
Source: METROPOLIS simulation results.

Residential migration patterns

Table 2 and figure 5 present the origin and destination rings and counties for the moves during 1998. A majority of the moves has taken place in the same district, although households who move to Paris rather come from outside Ile-de-France. The most important part of the moves has been into Paris from outside of the region. After Paris, the outsiders go mostly to 92 in the close suburb that provides almost the same features as Paris. We do not observe out-migration from the Ile-de-France.

Table 2: The distribution of moves between different rings (origin by destination).

Current District		***Origin district***				***Total***
		Outside	**Paris**	**C. S.**	**F. S.**	
Paris	Frequency	77.579	67.027	18.192	18.023	180.821
	Percent	**42.9%**	**37.1%**	**10.1%**	**10.0%**	**30.68%**
Close Suburbs	Frequency	61.135	22.633	103.205	20.168	207.141
	Percent	**29.5%**	**10.9%**	**49.8%**	**9.7%**	**35.15%**
Far Suburbs	Frequency	49.936	9.299	23.967	118.191	201.393
	Percent	**24.8%**	**4.6%**	**11.9%**	**58.7%**	**34.17%**
Region	Frequency	188.650	98.959	145.364	156.382	589.355
	Percent	**32.01%**	**16.79%**	**24.66%**	**26.53%**	**100.00%**

Source: Census, 1999.

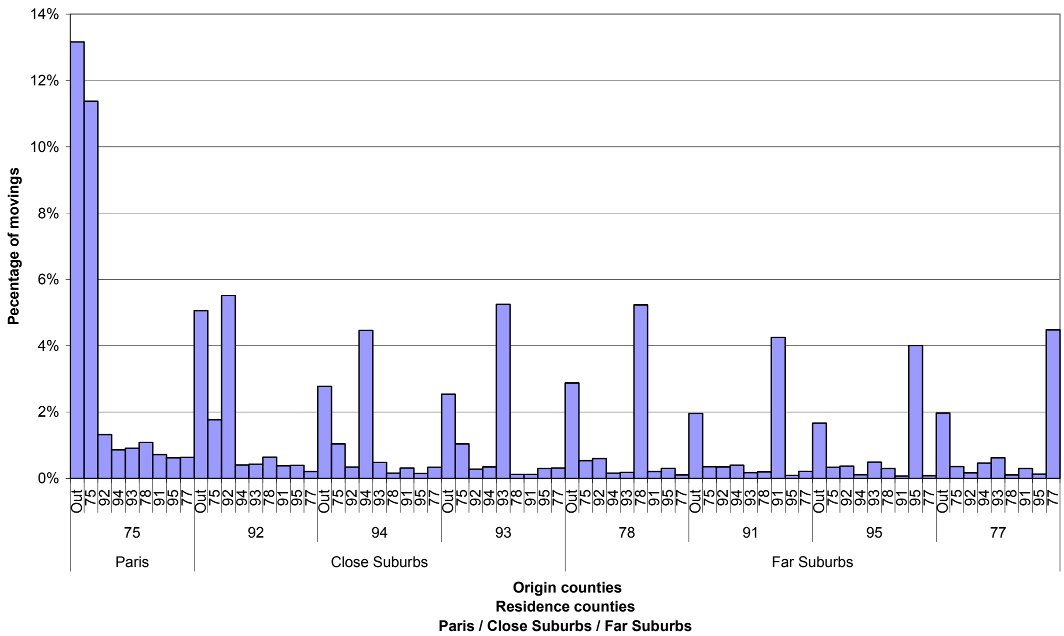

Figure 5: The distribution of moves between different counties en 1998.
Source: Census, 1999.

Model specification

In this section we develop the specification of the model components that comprise the focus of the paper: household residential location choice and housing price.

Household residential location choice model.

The commune *j, j=1,…,1300* contains C_j dwellings (housing units). We assume that all the dwellings *i* located in commune *j* have the same observable attributes, and therefore the same expected utility $V_i^h = V_j^h$ for household *h, h=1,…,N*. The total number of dwellings in Ile-de-France is denoted by *I*.

The probability for household *h* to choose a dwelling *i* is given by the Multinomial Logit formula:

$$\mathbb{P}_i^h = \frac{\exp\left(V_i^h\right)}{\sum_{i'=1}^{I} \exp\left(V_{i'}^h\right)}, \quad (1)$$

Since all the dwellings located in *j* have the same expected utility (since we do not have information on structural attributes of the housing units), and the same probability of being selected, Equation (1) implies that the probability that household *h* selects commune *j* is:

$$\mathbb{P}_j^h = C_j \mathbb{P}_i^h = \frac{C_j \exp\left(V_j^h\right)}{\sum_{j'=1}^{J}\left(\sum_{i' \text{ in } j'} \exp\left(V_{i'}^h\right)\right)} = \frac{\exp\left(V_j^h + \log\left(C_j\right)\right)}{\sum_{j'=1}^{J} \exp\left(V_{j'}^h + \log\left(C_{j'}\right)\right)}, \qquad (2)$$

Under the IIA (Independence of Irrelevant Alternatives) assumption, one can obtain consistent estimates of β (the preference parameters of V_j^h) by selecting a random sample of alternatives, with uniform sampling of alternatives, provided the correcting term $\log\left(C_j\right)$ is added to the likelihood. However, more efficient estimates can be obtained with importance sampling of alternatives, that is if the probability that alternative j is included in the choice set is proportional to C_j, provided a second correcting term $-\log\left(C_j\right)$ is added to the likelihood. Since the two terms $\log\left(C_j\right)$ and $-\log\left(C_j\right)$ exactly compensate, no correcting factor is necessary to obtain consistent estimates of β when importance sampling is used (see Ben-Akiva and Lerman, 1985 for details on this section).

Housing price model and endogeneity

One of the major factors affecting location choice is the price, which we predict using a semi-hedonic regression on the natural log of total price:

$$\ln P_j = X_j \lambda_1 + S_j \lambda_2 + D_j \lambda_3,$$

where the demand D_j and supply S_j levels are explicitly taken into account (in log form) and in which X_j is the vector of local characteristics.

Housing price depends on the supply and demand for housing, and demand depends on price, so the two models (location choice and price) should be estimated jointly in order to correct for the potential bias implied by the endogeneity of prices P_j. The bias is tested for and corrected using the method proposed by Blundell and Smith (1989), which simply consists of introducing the residuals from the price equation in the location choice equation.

We therefore develop an iterative procedure in which prices depend on estimated demand and demand depends on observed price and price residual. We denote by $X_i^h = X_j^h$ the vector of commune attributes (except price), possibly crossed with household characteristics, and we assume a linear formulation for expected utility: $V_i^h = X_i^h \beta + P_j \delta_h$, where β denotes a vector of parameters, to be estimated, and δ_h corresponds to the marginal utility of price, which may depend on household characteristics. The expected demand for commune j is then: $D_j = \sum_{h=1}^{N} \mathbb{P}_j^h$. The vector of other commune attributes influencing price is denoted by Z_j, and we assume a log-linear formulation, so that the price equation is of the form: $P_j = Z_j \gamma + D_j \lambda + \varepsilon_j$.

In order to test and correct for the endogeneity of prices, Equation (1) is replaced by:

$$\mathbb{P}_j^h = \frac{\exp\left(X_j^h \beta + P_j \delta_h + \varepsilon_j \eta\right)}{\sum_{j=1}^{J} \exp\left(X_{j'}^h \beta + P_{j'} \delta_h + \varepsilon_{j'} \eta\right)}, \qquad (3)$$

where no correction factor is necessary if importance sampling is used.

Empirical results

Application of METROPOLIS to the Paris region

To model the Paris region, we have used the road network coded by IAURIF (Institute for urban planning and development of the Paris Ile-de-France region). The coded network had 606 zones and 17957 links. The morning peak hour O-D matrix includes 970,000 trips. We have multiplied the peak hour O-D matrix by relevant factors to cover the whole day. The trips were divided by their purpose: work trips and others.

To obtain the dynamic behavioural parameters a survey, MADDIF, was conducted, (Fontan, 2003) in 2000. 4200 individuals were surveyed by telephone. It provided the distribution of schedule delay penalties, the Logit heterogeneity parameter proportional to VOT (value of time, for which the official value of 12.96 Euros per hour was taken) and the distribution of desired arrival time.

Table 3: The Paris Region Transportation Model Predictions.

Variable	*Values*
Travel time [min]	32.50
Free flow travel time [min]	24.34
Congested delay [min]	8.16
Early arrival delay [min]	28.28
Late arrival delay [min]	19.07
Average velocity [km/h]	28.90
Early arrivals percentage[%]	47.25
Late arrival percentage [%]	33.15
Average cost [€]	9.47
Free flow travel time cost [€]	5.26
Waiting time cost [€]	1.76
Average schedule delay cost [€]	2.45
Traffic volume [million de veh. x km]	63.70
Average travelled distance [km]	17.51
Number of links passed by a traveller	17.61
Congestion index	28.85

Source: METROPOLIS simulation results.

Housing price

The estimated coefficients for housing price model are presented in table 4. The R^2 for the model is 0.53. We obtain the expected signs for demand and supply but they are not exactly opposed. A purely structural equation (results not reported here, available on request) with only supply and demand gives coefficients exactly opposed, which means that the price only depends on the supply/demand ratio, and not separately on supply and demand. Once covariates are added, however, the coefficients on demand and supply are no longer equal in absolute terms, because the covariates are more correlated with demand.

A decrease in average travel time significantly increases the price: 10 minutes less travel time to work imply a 2.8% increase in housing price. The price is very sensitive to socio-economic structure of the commune: a 10% increase in the proportion of one-member households causes a 50% increase of the price. A similar 10% increase for the proportion of two-member households results in a 19% increase of the price (note that effects of households with one or two members should be interpreted with reference to the omitted fraction of households with three or more members). Similarly, the fraction of households with no or only one working member has a positive effect on the price. Strangely enough, the fraction of foreign households has a positive effect on price. We should notice however, that the data do not distinguish the nationality of the foreigners, and make no difference between OECD countries and third world ones. Finally, we notice the negative and highly significant effect of the proportion of low and intermediate income families on the price.

Table 4: housing price estimation results

Variable	*Coefficient*	*Standard error*	*t-statistic*	*p-value*
Intercept	11.02668	0.12800	86.14	<.0001
Log(Supply)	-0.04791	0.02466	-1.94	0.0522
Log(Demand)	0.09918	0.02244	4.42	<.0001
Average travel time from j to work (minutes)	-0.00280	0.00085119	-3.28	0.0011
% households with 1 member	5.09136	0.37884	13.44	<.0001
% households with 2 members	1.87960	0.34135	5.51	<.0001
% households with no working member	1.25241	0.30954	4.05	<.0001
% households with 1 working member	0.82300	0.33762	2.44	0.0149
% poor households	-6.63187	0.50316	-13.18	<.0001
% households with medium income	-4.54311	0.33102	-13.72	<.0001
% households with a foreign head	1.58406	0.36279	4.37	<.0001

Source: Authors estimations' results (using SAS).

Location choice

Table 5 contains the results of the residential location choice model estimation. With a pseudo-R^2 of 22% this model has a moderate explanatory power. This estimation has been performed on a 20% sample of total moved households in order to improve the computational tractability of the model.

We notice the very significant role of the "same district as before" variable. This shows the strong preference of the households to move in the same district or neighbourhood in which they lived before. Testing the effect of the distance from last residence may be interesting but it was not possible with our available data. The Paris dummy variable has a negative coefficient, indicating that, ceteris paribus, the households who live in Paris and decide to move have a slightly higher probability of relocating to a district outside Paris than do residents living outside Paris. Note that this is consistent with the intra-metropolitan migration patterns shown in Table 2 and Figure 5, and with general expectations that households moving into the region, and new households formed within the region locate initially within Paris, and may relocate to suburban neighbourhoods later. Note, however, that some of the other variables in the model, such as better accessibility in Paris, tend to have effects that at least partially offset this suburbanization preference, while others, such as housing prices, tend to reinforce it.

As expected, housing price has a negative effect on location preference for a commune. This effect increases with the age of the household head and decreases as the household income increases. The older heads of households are more sensitive to price and the richer households are less sensitive to it. Since price is entered using three variables to capture average effects as well as interactions with age and income, the combined effects are complex. We note that the average price effect as well as the age and income interactions, all have expected signs. However, for a small subset of the population, namely very young and very rich households, the net price effect from the interaction of these three coefficients would be predicted by this model to show a slight positive preference for higher prices in communes where they the neighbouring households are in the same socio-economic category and which have more amenities.

Table 5: residential location choice estimation results

Variable	*Coefficient*	*Standard error*	*t-statistic*	*p-value*
Same district as before move	2.5461	0.009353	272.24	<.0001
Paris	-0.2988	0.0267	-11.19	<.0001
Log(Price)	-1.7285	0.1009	-17.14	<.0001
Log(Price)* (Age-20)/10	-0.0653	0.004695	-13.92	<.0001
Log(Price)* Log(Income)	0.1783	0.0100	17.78	<.0001
Number Railway stations	-0.0129	0.002838	-4.56	<.0001
Number Subway stations	0.007070	0.001300	5.44	<.0001
Average travel time from j, commuting (TC)	0.000561	0.000483	1.16	0.2457
TC*(Dummy female)	-0.006842	0.000697	-9.82	<.0001
Average travel time from j, by private car (VP)	-0.001391	0.000481	-2.89	0.0038
Distance to highway [km]	-0.003392	6.273E-7	-5.41	<.0001
% households with 1 member * 1 member in h	2.6327	0.0851	30.95	<.0001
% households with 2 members* 2 members in h	0.9366	0.3060	3.06	0.0022
% households with 3+ members* 3+ member in h	3.2437	0.0810	40.03	<.0001
% hh with no working member * no working member in h	6.1790	0.2287	27.02	<.0001
% hh with 1 working member * 1 working member in h	0.3384	0.1455	2.33	0.0201
% hh with 2+ working member * 2+ working member in h	0.7132	0.1078	6.61	<.0001
% hh with a young head	-0.0147	0.1335	-0.11	0.9122
% hh with a young head * young head in h	4.7947	0.1351	35.50	<.0001
% poor households	0.3853	0.1706	2.26	0.0240
% households with a foreign head * foreign head in h	6.2094	0.1622	38.28	<.0001
% households with a foreign head * French head in h	-2.7905	0.1007	-27.70	<.0001
Total employment [/1000]	-0.0001349	2.348E-7	-0.57	0.5657
Density (Population/Surface) [1000 persons/km]	-0.004621	1.0479E-6	-4.41	<.0001
Log(Population)	0.0931	0.005506	16.90	<.0001
% change in population, 1990 to 1999	0.0931	0.0168	5.54	<.0001

Source: Authors estimations' results (using SAS).

The relative sensitivity to price is as we would expect, though the potential for a small positive preference for higher prices for this specific subpopulation and sample of locations is likely to be due to some amenities that are not accounted for in the model, rather than an actual preference to may more for housing, ceteris paribus. Increase of the average travel time by public transit decreases the utility of households headed by a woman, though this effect is insignificant for male-headed households.

The number of metro stations in a commune increases the probability of location but the number of railway stations decreases it, after accounting for transit accessibility and other effects. These results may reflect the relative effects of positive and negative externalities associated with metro stations and railway stations. Metro stations are more likely than railway stations to be located within clusters of shopping and service employment or adjacent to major cultural attractions, and railway stations are larger and may be more likely to have negative localized externalities on the immediate neighbourhood, such as traffic, noise, and possibly petty crime. The average travel time by private car and the distance to the highway have a negative effect on the preference for a commune, as expected.

The estimated coefficients corresponding to the socio-economic structure of the commune show a general preference of the households to live with the people in the same social category. This preference is very strong for households without workers, or with a foreign or young head. The households with one worker are less sensitive to the concentration of similar households. Households of French origin tend to avoid locations in which there are higher concentrations of foreign households. The coefficients for the percentage of young head households and the total number of employments are insignificant. Households prefer more populated but less dense communes. The communes that have absorbed more population during the 90-99 period attract still more households. Considering these variables the composition of the population with regard to income does not remarkably influence the location choice of households. But the density of high, middle and low income families can be studied if we don’t take into account the total population, its density and its evolution.

Adding the residuals of the price equation as an explanatory variable, the estimated coefficients change trivially and the coefficient of this new variable is not at all significant. This result confirms that housing price is not endogenous with regard to the location choice model. In other words, the variables used in these two models fully explain the correlation between prices and location choice.

Conclusions

The research on which we report in this paper is from an early phase of a longer-term research collaboration to explore the interaction of land use and transportation. Our particular emphasis is on issues of dynamics, endogeneity and constraints. We have now succeeded in developing and estimating a model of residential location at a commune level for the Paris region, with a rigorous econometric treatment of the endogeneity of housing prices. Further, we have integrated UrbanSim with METROPOLIS, providing the first experience of connecting dynamic models of land use and traffic. By coupling these models we are able to represent the endogeneity of residential location and traffic, given a distribution of job locations. In related research, we are developing employment location choice models and real estate development models for the Paris region, and will address the endogeneity among these choice processes. We are also developing a rigorous theoretical and empirical treatment of the problem of endogenous constraints on the availability of alternatives, where demand may exceed the supply of housing within popular neighbourhoods. In this field, there is a need to distinguish between unconstrained and constrained demand, and traditionally used estimation procedures tend to confound these two concepts (for a complete treatment of this topic, we refer the reader to de Palma, Picard and Waddell, 2006).

This research is in progress, and it is likely to evolve substantially as it moves to completion and into an operational framework for exploring the potential effects of combinations of transportation and land use policies. We hope that this line of research provides a valuable future direction for the integrated treatment of land use and transportation, and advances the state of the field by better representing these as dynamic processes with substantial endogeneity.

References

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

Blundell, R. and Smith, R. (1989) "Estimation in a Class of Simultaneous Equation Limited Dependent Variable Models", *Review of Economic Studies* 56(1): pp. 37-57.

de la Barra, T. (1990) *Integrated Land Use and Transportation Modelling: Decision Chains and Hierarchies*, Cambridge Urban and Architectural Studies, No. 12, Cambridge, UK.

de Palma, A. and Marchal, F. (2002) "Real cases applications of the fully Dynamic METROPOLIS Tool-box: an Advocacy for Global Large-scale Mesoscopic Transportation Systems", *Networks and Spatial Economics* 2 (4): pp. 347-369.

de Palma, A., Picard, N. and Waddell, P. (2006) "Discrete choice models with capacity constraints: an empirical analysis of the housing market of the greater Paris region". Paper presented at the Conference in honor of Kenneth Small.

DREIF (2005) Enquette Globale de Transport 2001, Paris, France.

Fontan, C. (2003) Choix de l'heure de départ et coûts des délais: enquête et estimations en Île-de-France, *Ph.D. Thesis,* University of Cergy-Pontoise, France.

Putman, S. H. (1983) *Integrated Urban Models*, Pion, London, UK.

THEMA/TT&R. (1998 – 2002) *QUATUOR: Outils dynamiques de simulation pour la gestion des déplacements en Ile-de-France*. Report for DRAST/ PREDIT, N°98MT30, N°99MT35, N°00MT66, University of Cergy-Pontoise, France.

Waddell, P., Borning, A., Noth, M., Freier, N. Becke, M and Ulfarson, G. (2003) "Microsimulation of Urban Development and Location Choices: Design and Implementation of UrbanSim", *Networks and Spatial Economics* 3 (1): pp. 43-67.

Waddell, P., Sevcikova, H., Socha, D., Miller, E. and Nagel, K. (2005) "Opus: An Open Platform for Urban Simulation", The Computers in Urban Planning and Urban Management Conference, London, U.K.

Acknowledgements

This study has been performed as part of the SIMAURIF project, in collaboration with the Institute for urban planning and development of the Paris Ile-de-France region (IAURIF), and financed by the French Ministry of transportation (PREDIT) and the Transportation department of Ile-de-France region (DREIF). The third author also acknowledges the National Science Foundation Grants EIA-0121326 and IIS-0534094. We wish to acknowledge the collaboration of Dany Nguyen-Luong and José Moyano. Finally, this paper benefited from the comments of participants in the Second Kuhmo Conference, Moshe Ben-Akiva, Daniel McFadden, and Stef Proost.