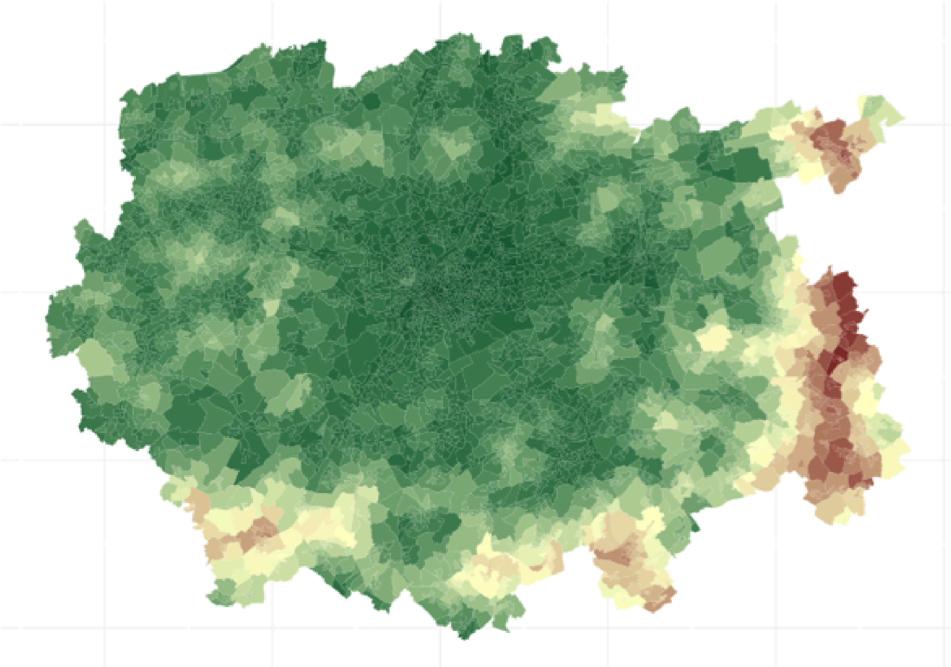


The Integrated Land-Use and Transport Model of Brussels

Dimitrios Efthymiou

Ricardo Hurtubia

Michel Bierlaire

Constantinos Antoniou

Transp - OR, EPFL

April 2013

STRC | 13th Swiss Transport Research Conference

Monte Verità / Ascona, April 24 – 26, 2013

Transp - OR, EPFL

The Integrated Land-Use and Transport Model of Brussels

Dimitrios Efthymiou
PhD Candidate
LoTE, NTUA
15780 Athens
phone: +41 21 693 93 29
fax: +30 210 7722629
defthym@mail.ntua.gr

Ricardo Hurtubia
Assistant Professor
Departamento de Urbanismo, Univ. de Chile
Portugal 84, Santiago, Chile
phone:
fax:
rhurtubia@fau.uchile.cl

Michel Bierlaire
Professor
Transp - OR, EPFL
1015 Lausanne
phone: +41 21 693 25 37
fax: +41 21 693 80 60
michel.bierlaire@epfl.ch

Constantinos Antoniou
Assistant Professor
LoTE, NTUA
15780 Athens
phone: +30 210 7722783
fax: +30 210 7722629
antoniou@central.ntua.gr

April 2013

Abstract

The Land-Use Model UrbanSim and the agent-based traffic simulation model MATSim, are being further developed and integrated under the framework of the Sustaincity project, that is funded by the European Union FP7. UrbanSim differs from the other Integrated Land-Use and Transport models, as it adopts an approach of dynamic disequilibrium, it can forecast in different time scales, and it requires extremely disaggregate spatial information. UrbanSim applies 1) Location/relocation choice models, that determine the spatial allocation of the new or relocating agents (households and jobs); 2) Transition models, that describe the evolution of the agents' demand per each simulation period; 3) Development models, that determine the generation of new real estate supply (residential and non-residential buildings) and 4) Price models, that compute the prices of real estate. MATSim generates the zone to zone impedance matrix and compute the car, walk and bike accessibilities per spatial level, based on the 'logsum' function.

In this paper, the authors present in detail the steps required for a successful implementation of an Integrated Land Use and Transport model in UrbanSim and MATSim. The study begins with the presentation and interpretation of the choice, development and real estate price model estimations, for the Brussels (Belgium) case study (zone-level project). In order to verify the model's performance, the results of a base-case scenario are validated with real data. A land-use policy scenario is then applied and evaluated using disaggregate spatial indicators. The study closes with a description of the difficulties and the challenges that were encountered.

Keywords

Integrated Land-Use and Transport Models, Brussels, Location Choice

1 Introduction

Numerous regional economic, social and environmental characteristics, increase the size and the complexity of urban systems. In order to handle the complexity and create an understandable framework of their interactions, the research and policy communities invest on the development of integrated land-use and transport models, also known as Land-use/Transport Interaction Models (LUTI). In these models, the marriage of two methodological approaches is achieved: the theory of the urban complexity, which sometimes makes the interaction between the land-use and transportation clearer, and the mathematical modeling and simulation techniques, which help to better understand the situation, which is needed in order to make the right policy planning decisions (Waddell, 2002). The latter is enhanced by the integration of the Geographical Information Systems (GIS) for the visualization of the outcomes of the alternative simulated scenarios (e.g Shaw and Xin, 2003).

The need for accurate forecasting of the direct and indirect effects –many of which are usually unobservable– of land and transport policies, on the society, environment, economy, the land-use changes, and the transport patterns of the community, has led to the development of many different LUTI models around the globe.

The technical difficulties that arise from their implementation, such as the restricted data available from the authorities, and the need for more powerful computers to handle the computationally demanding simulations, were made clear from the beginning. Despite these restrictions, the effort was continued with significant improvements. LUTI models have been applied in a number of cities in Europe, the Americas and Asia: cities of different size, population and spatial characteristics, for evaluation of policies and investments. However, these applications were mainly made for research purposes, while the development and application of LUTI models remained a demanding procedure and entails risks that the authorities and the private sector usually hesitate to undertake.

Nevertheless, many uncertainties about their efficiency and their forecasting capabilities remain unanswered, preserving the ground under the LUTI models fertile for criticism. Hunt *et al.* (2005) make an extensive review of known LUTI models, and point out the pros and cons of each, and Curtis (2011) reviewed the inefficiencies of the current integrated land-use and transport models in measuring the accessibility of public transport.

The objective of this research is to develop an Integrated, agent-based, Land Use and Transport Model (LUTI) for Brussels (Belgium), evaluate its ability in forecasting the socioeconomic characteristics of the metro area, and measure the effect of land use policies in transport accessibility.

The individual location choice, transition, development, and real estate price models were estimated and simulated in UrbanSim. The simulation of the traffic conditions was performed using the agent-based model MATSim, which has been integrated in UrbanSim within the framework of the European FP7 research project SustainCity (`http://www.sustaincity.org`). In accomplishing this objective, sub-objectives have been set, including the identification of the strengths, the weaknesses and the gaps of this model, and examine the margins of improvement, so as to be applicable even with less detailed data, which is usually the case in a zone-level UrbanSim project. The remainder of this paper is structured as follows: Following the introduction, the current LUTI models are presented, emphasizing on the characteristics of UrbanSim (Waddell, 2002). The analysis of the case study begins with a description of the data needed and their sources, continues with the estimation and interpretation of the location choice and real estate price models, and closes with the validation of the results of a baseline scenario, using observed data.

2 UrbanSim

The LUTI model UrbanSim, was initiated by the Department of Urban Design and Planning of the University of Washington in the end of 1990s (Waddell, 2002) and has been further developed by the University of California, Berkeley and other major institutions. UrbanSim's main characteristics are that it adopts an approach of dynamic disequilibrium, it can forecast in different time scales, and it needs extremely disaggregate spatial information (Waddell, 2002).

Urbansim can be applied in three different levels of spatial disaggregation: zones, gridcells or parcels. One of its characteristics that can be regarded both positive and negative is the need of extremely disaggregate data. The model requires major and minor data tables, such as households, buildings and jobs within the area of interest. Urbansim integrates the following types of models: 1) *Transition models*, that describe the evolution of the demand agents (households and jobs) for each simulation period; 2) *Development models*, that describe the generation of new real estate supply (residential and non-residential buildings); 3) *Relocation models*, where the agents' decision of moving from their current location is simulated; 4) *Location Choice Models*, describing the spatial allocation of the new or relocating agents across the different, already existing or new alternative locations; 5) *Price models*, that compute the prices of real estate (Waddell, 2002, Waddell *et al.*, 2007). The models interact in each simulated period, generating a new state of the system that is used as a starting point for the simulation in the next period.

All location choice models in UrbanSim are based on a *multinomial logit approach*, where decision makers (agents) choose from a sample of available alternatives (buildings or locations) selecting the one that provides maximum utility given its attributes and price. Market clearing is treated with a *first come first served* approach (Waddell, 2010) meaning that, when two agents select the same location, conflict is solved by randomly selecting one of them.

Real estate price models are also of high relevance because they describe the market value of the traded goods. In UrbanSim, real estate prices are modeled using a hedonic regression of property value per surface unit of the building, and its environment and market-level vacancy rates (Waddell and Ulfarsson, 2003), as following:

$$\ln(p_{vit}) = \alpha + \delta\left(\frac{Q^s_v - Q^c_{vt}}{Q^s_v}\right) + \beta X_{vit} \tag{1}$$

Where $\ln(p_{vit})$ is the natural logarithm of price of land per surface unit for development type v at location i and time t, Q^c_{vt} is the current vacancy rate at time t, Q^s_v i is the long-term structural vacancy rate, X_{vit} is a vector of building and location attributes, and α, δ and β are estimated parameters.

In each period, new households and firms are generated by the Transition models. Simultaneously, new supply is generated by the Development models and distributed within locations in the city. Relocating and new agents enter the market and choose their location following the distribution defined by the location Choice models. At the end of each period, prices are computed and all location and building attributes are updated to enter as the main input to the next period simulation.

One of the main characteristics of UrbanSim is the independent estimation process for each of the involved submodels. This is a practical advantage that simplifies the implementation of the model but also implies strong assumptions about the behavior of agents and the interdependence of the decision processes that takes place in the city. The modular structure and open source nature of the code makes feasible to customize UrbanSim for several different circumstances and conditions, although this may require advance knowledge of the software.

UrbanSim is probably the most widely applied LUTI model. Examples of UrbanSim case studies include: Springfield (Oregon), Salt Lake City (Utah), Seattle (Washington) and San Francisco (California) (`http://www.urbansim.org`). The European Union's research project SustainCity (Sustaincity, 2009) aims to apply UrbanSim in three European Metropolitan areas: Paris (France), Zurich (Switzerland) and Brussels (Belgium). Before SustainCity, significant knowledge for the European conditions was gained from two case studies in Lausanne (Patterson and Hurtubia,

2008) and Brussels (Patterson and Bierlaire, 2010, Patterson *et al.,* 2010), where the model was applied with aggregate data, and Zurich (Löchl *et al.,* 2007). Among the main tasks of this project is to integrate the agent-based traffic simulation model MATSim (`http://www.matsim.org`) with UrbanSim, to implement a tool for policy evaluation based of the Social Welfare Function (De Palma *et al.,* 2010) and to integrate a "Sustainability submodel".

MATSim is an agent-based framework used for travel demand modeling. It is based on the individual person's travel schedule, while it integrates information about his travel patterns, such as: mode used, time of departure, time of duration per trip and time of return. MATSim selects the optimal routes applying iterative optimization process, an algorithm suggested by Charypar and Nagel (2005), which considers random initial departure time and duration for each activity.

3 Accessibility and land-use

The accessibility measurements are gaining an increasing interest in policy evaluation, since they indicate the ease with which the activities can be reached. The accessibility affects the household location choice (e.g Vandenbulcke *et al.,* 2009) and as a result the house prices (e.g Medda, 2012, Ibeas *et al.,* 2012). Gutiérrez and Urbano (1996) tried to forecast the resulted increase of accessibility in Europe, after the implementation of the Trans-European road network, using an indicator that is based on the impedance from country to country and the GDP. Linneker and Spence (1996) measured the impact of the resulted accessibility after the construction of the M25 London Orbital Motorway, in the regional development. Geurs and Wee (2004) make an extended review of the accessibility indicators used for land-use and transport strategies. They identify four types of components in the current accessibility indicators, that are based on: 1) land-use (e.g. the supply and demand of the opportunities distributed spatially); 2) transportation (e.g. travel time); 3) temporal (e.g. availability of the opportunities in day); 4) individual (e.g. personal characteristics). Moreover, they identify that there are four basis perspectives on measuring the accessibility, namely: 1) infrastructure-based; 2) location-based; 3) person-based; 4) utility-based. According the same research, accessibility is being used as a way of measuring the operationalization, the interpretability and communicability, as a social or economic indicator. Currently, there are two types of utility-based accessibility measurements. Ben-Akiva and Lerman (1985) suggested the *logsum*, which uses the demoninator of the multinomial logic model. This measures the accessibility of the complete choice set. In this research, we are focusing on this, since it is the one that has been recently implemented in the adept-based model MATSim. The second, is the based on the doubly constrained entropy model (Martinez, 1995). Banister and Berechman (2001) suggest they accessibility is the engine behind the economic growth of an area, after the implementation of a policy, because leads to employment increase and factor productivity. Vandenbulcke *et al.* (2007) provide an extended report on analyzing static accessibility indicators for different transportation models in Belgium. More recently, Martinez and Viegas (2013) applied an methodology that is used in Botany to model the distance-decay functions for accessibility assessment in transport studies.

4 Case Study Setup

4.1 Data Collection

One of the main characteristics of UrbanSim is the requirement of extremely disaggregate data. The main datasets needed are the agents that generate the demand (jobs and households) and the available supply (buildings), while historical data of the developed projects (development events history) and the shapefiles that define spatial levels are integral parts needed for a basic UrbanSim run.

The main data sources used for this application were the 2001 Belgium Population Census and the Belgium Land Registry (a cadastre of real estate goods). Both datasets were obtained at an aggregate level from the Belgian Statistical Authority (SPF - Economie, `http://www.economie.fgov.be`). Aggregate data regarding employment by activity type and commune was collected from the ONSS (`http://www.onssrszlss.fgov.be`) and INASTI (`http://www.rsvz.be`)databases, from the same source. Additionally, individual level data for households and persons was obtained from the travel survey MOBEL (Hubert and Toint, 2002), performed in the area of study during 2002.

The data refer to different levels of spatial disaggregation. These levels are the "communes", that divide the Brussels Metropolitan Area in 151 units, and the "zones" that divide it in 4945 units. Each zone belongs only to one commune.

For zone-level projects, as the one described here, UrbanSim makes the following assumption: The total number of buildings in a zone is aggregate per building type. In other words, each zone contains one representative building per type, while the real number of units (buildings) of each type in the zone, is included in a field named "number of units" in the buildings table. In our building data for Brussels, there are 14 categories of building types: four residential (detached, semi-detached, attached and apartments) and 10 non-residential (industrial, governmental, educational, quarrying, warehouses, office, shops, hotels/bar/restaurants, industrial). The main attributes of the buildings are: price (average value per unit), type, structural characteristics (residential/non-residential sqft etc), and capacity.

The households table contains information about the size, the number of workers, cars, income, location of residence, etc. Finding a complete dataset of households is usually impossible. For the purpose of this research, a synthetic population was generated (Farooq *et al.,* 2012).

Since the employment data were available at the commune level, they were distributed to zones through Monte Carlo simulation, following the observed distribution of available non-residential surface. The number of job-sectors was assumed to be equal with the number of types of non-residential buildings, meaning that each type of job could be located in a specific type of building (e.g. jobs in industrial sector can be located in industrial buildings).

The integrated version of UrbanSim with MATSim requires a road network, which in our case was acquired from the Open Street Maps (openstreetmaps.org), and a list of commuters (workers) with information about their household (origin) and job (destination). This dataset was structured using the synthetic population (households) and the observed OD for Brussels.

Finally, UrbanSim requires a dataset with the residential and non-residential projects (from 1990 to 2001), to be used for the estimation of the building development model, and other smaller, less disaggregate data, such as the annual employment control totals, annual household control totals and target vacancies.

4.2 Preliminary Model Estimations

4.2.1 Introduction

All the models were estimated in the OPUS/UrbanSim software platform. Figure 1 depicts the interaction between the land-use and transport models and sub-models used in this case study. One of the main advantages of using UrbanSim for model estimation, is that in OPUS the user can create interaction variables using different data tables. Several different specifications were examined, given the available data and following what the literature and urban economic theory suggests as explanatory variables for each of the modeled phenomena (Picard and Antoniou, 2011). Final specifications were selected following estimate-significance and theoretical-consistency criteria. For all choice models a linear-in-parameters utility function specification like the following was chosen:

$$V_i = \sum_k \beta_k x_i^k \tag{2}$$

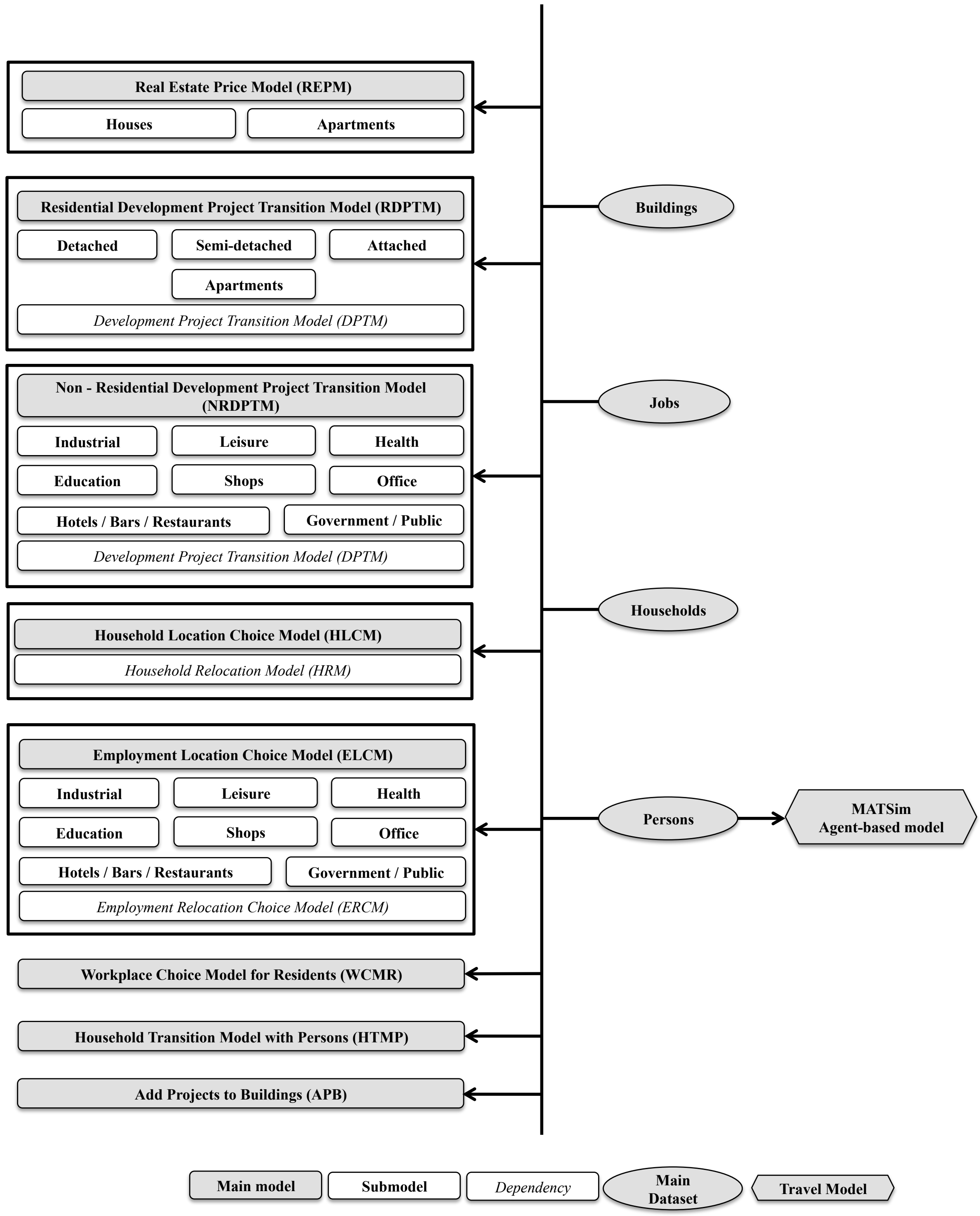
Where β_k is the k-th parameter to estimate and x_i^k is the k-th attribute of alternative i. For some models, like the household location choice model (see Table 2), x_i may be replaced by $x_{in} = x_i \cdot x_n$, describing an interaction between an attribute of the alternative i and a characteristic of the decision maker n.

The choice probabilities of a multinomial logit model (MNL) which is used in UrbanSim, is given by:

$$P(i|n) = \frac{e^{V_{in}}}{\sum_{j \in C_q} e^{V_{jn}}} \tag{3}$$

Where C_q is the choice set.

Figure 1: The Integrated Land-use and Transport Model of Brussels



4.2.2 MATSim for UrbanSim

MATSim for UrbanSim (Nicolai, 2012, Nicolai and Nagel, 2011) is able to use a single accessibility measurement per zone, instead of OD matrices. These accessibility indicators are calculated by the "logsum" (the logarithm of the denominator of the MNL probability) as follows:

$$A_i = \frac{1}{\beta_{Scale}} \cdot ln(\sum_{j=1}^{J} (W_j \cdot exp(-\beta_{Scale} \cdot c_{ij}))) \tag{4}$$

where, A_i is the workplace accessibility at location i, $i \in I$ the origins, $j \in J$ the destinations, β_{Scale} is a scale factor related to the scale of a logic model, W_j is a weight giving the number of jobs at location j, $exp(-\beta_{Scale} \cdot c_{ij})$ is a deterrence function, c_{ij} is the generalized travel cost from location i to location j.

The generalized travel cost c_{ij} is:

$$c_{ij} = (\alpha \cdot ttime) + (\beta \cdot ttime^2) + (\gamma \cdot ln(ttime)) + (\delta \cdot tdistance) + (\epsilon \cdot tdistance^2) + (\zeta \cdot ln(tdistance)) + (\eta \cdot tcost) + (\theta \cdot tcost^2) + (\iota \cdot ln(tcost)) \tag{5}$$

where *ttime* is the travel time in minutes, *tdistance* is the distance in meters, *tcost* is the monetary travel cost, α to ι are the marginal utilities

In this study, the default values of MATSim were used: $\beta_{Scale} = 1$ and $\alpha = -12$. The other values were set to zero.

Home and work locations are distributed randomly on the nodes within each zone, in order to avoid that all the household and workplace locations are attached at the same link of the road network. Another option would be to be distributed to the nodes in a given distance from the zone centroid, however, because of the uneven sizes of zones in our case, the first method was selected.

4.2.3 Real Estate Price Model

The real estate price model (REPM) used in UrbanSim, is a semi-log linear regression based on Ordinary Least Squares (Franklin and Waddell, 2003, Waddell and Ulfarsson, 2003). It predicts the average value per unit, for every year of the simulation (eq. 1).

The results for the real estate price model are shown in Table 1. Two submodels were estimated for the REPM: one for the houses (detached, semi-detached and attached), and one for the apartments. There were no available observations of non-residential real estate prices, and therefore no model was estimated for this case.

The price of houses and apartments is positively affected by the car accessibility of the zone and the percentage of green areas in the commune. Moreover, sociodemographic characteristics that have a positive impact is the percentage of households with high income in the commune, and the logarithm of the population density.

In order to avoid endogeneity issues, an instrument variable De Palma and Picard (2005) was included in this model's specification. This is the communal housing tax, which is a percentage of the dwelling's price per year, and has a negative effect.

The real estate price model presented here is mostly based on location (neighborhood or commune) attributes. The only building-specific attribute used is the residential m^2, which is positive and in both sub-models. Despite this problem and the fact that the literature shows that prices are largely explained by attributes of the buildings (e.g. (Löchl and Axhausen, 2010, Efthymiou and Antoniou, 2013)), the presented models are still able to capture land use effects that should be relevant for the modeling purposes.

Table 1: **Real Estate Price Model**

			Houses (n=14835)			
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
constant	-			11.5407	0.0135	857.94
$\beta_{\text{car-acc}}$	Car accessibility	zone	%	0.0020	0.0005	4.09
β_{green}	Green area score	commune	0 to 1	0.1349	0.0125	10.81
$\beta_{\text{income-high}}$	Percentage of high income (>3) households	commune	%	0.0260	0.0004	60.02
β_{tax}	Housing tax	commune	%	-0.0681	0.0014	-47.75
$\beta_{\text{pop-den}}$	Logarithm of population density	commune	ln(pop/hectare)	0.0591	0.0011	56.33
β_{sqm}	Surface	building	m^2	0.0005	5.29e-05	8.751
$\boldsymbol{R}^2$=0.59						
			Apartments (n=4945)			
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
constant	-			11.2914	0.0306	368.69
$\beta_{\text{car-acc}}$	Car accessibility	zone	%	0.0046	0.0011	4.09
β_{green}	Green area score	commune	0 to 1	0.4128	0.0290	14.24
$\beta_{\text{income-high}}$	Percentage of high income (>3) households	commune	%	0.0225	0.0010	22.67
β_{tax}	Housing tax	commune	%	-0.0334	0.0033	-10.13
$\beta_{\text{pop-den}}$	Logarithm of population density	zone	ln(pop/hectare)	0.0020	0.0011	1.82
β_{sqm}	Surface	building	m^2	0.0002	0.0001	1.89
$\boldsymbol{R}^2$=0.31						

4.2.4 Household Location Choice Model (HLCM)

The household location choice model estimation results are presented in Table 2. All parameters are statistically significant and have the expected signs. Price has a negative effect in the utility for all households, no matter their income level. The presence of high income households attracts other households of high income but makes locations less attractive for low income households. Households with university degree holders prefer to be located in zones with high ratio of university degree holders. This is consistent with the expected social agglomeration and segregation effects, usually observed in residential location.

Car accessibility increases the utility of car-owning households. Households with workers prefer to be located closer to the central business districts and households without owning cars, select locations close to the rail stations. Communes with with percentage of green areas are more attractive.

The spatial alternative-specific constant that accounts for unobserved attributes, indicates the attractiveness of the central locations. This constant is active when the location is inside the Brussels Capital Region.

4.2.5 Employment Location Choice Model (ELCM)

The employment location choice model is subdivided in eight submodels, one for each type of economic activity. Table 3 shows the estimation results for each submodel. Jobs in the agricultural and mining sectors are not considered for modeling purposes.

Employment location choice of each sector is positively affected by the density of jobs of the same sector in commune. The logarithm of non-residential surface and the car accessibility in zone, have a positive impact, when significant. The estimation of the industry jobs location sub-model shows that there is a negative effect of the density of jobs in the zone, which is observed to be the case only in this particular sub-model. Office jobs prefer to locate in zones with agglomeration economies and therefore favor density of jobs of the same type. Job density also has a positive effect in the utility for office jobs, probably because office jobs are service providers and prefer to locate near potential clients. Retail jobs also have benefits for the agglomeration economies and therefore the presence of jobs of the same type and of jobs in general have a positive effect in their location preferences. Finally, it is noted that health and leisure activities prefer to be located in communes with high percentage of high income households.

Table 2: **Household Location Choice Model**

(n=48526)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{car-access}}$	Households with car * Car accessibility	household * zone	0 or 1 * (logsum)	0.0106	0.0036	2.95
β_{educ}	Household with high education level * Ratio of university degree holders in zone	household * zone	0 or 1 * ratio	3.6401	0.1301	27.97
β_{green}	Green area score	commune	0 to 1	0.1924	0.0733	2.62
$\beta_{\text{income-low}}$	Households at income class 1 or 2 * Ratio of hh with high income over all hh	household * zone	0 or 1 * ratio	-2.8948	0.2206	-13.12
$\beta_{\text{income-high}}$	Households at income class 4 or 5 * Ratio of hh with high income over all hh	household * zone	0 or 1 * ratio	4.8074	0.3712	12.95
β_{workers}	Households with workers * Log distance from CBD	household * zone	0 or 1 * ln(meters)	-0.1453	0.0117	-12.44
β_{rail}	Households without cars * Distance from rail station <1000m	household * zone	0 or 1 * meters	0.3681	0.0287	12.82
β_{price}	Logarithm of transaction price	building	log(euros)	-1.0298	0.0354	-29.09
ASC_{BCR}	Central Brussels area (central communes)	commune	0 or 1	0.8738	0.0196	44.64
Log-likelihood=–95751						

Table 3: **Employment Location Choice Model**

Industry (n=13943)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{job-den}}$	Logarithm of jobs density	zone	ln(jobs/hectare)	-0.0627	0.0084	-7.43
β_{sam}	Logarithm of non residential surface	building	ln(m^2)	1.2514	0.0105	118.65
$\beta_{\text{ind-den}}$	Density of jobs in industry sector	commune	jobs/hectare	0.0782	0.0028	27.47
Log-likelihood=-13634						

Office (n=14937)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{gov-den}}$	Density of jobs in public sector	commune	jobs/hectare	-0.0212	0.0033	-6.36
$\beta_{\text{off-den}}$	Density of jobs in private sector (office)	commune	jobs/hectare	0.0152	0.0031	4.93
$\beta_{\text{job-den}}$	Logarithm of jobs density	zone	ln(jobs/hectare)	0.6641	0.0094	70.19
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	-0.0057	0.0005	-12.08
β_{sqm}	Logarithm of non residential surface	building	ln(m^2)	0.5227	0.0072	72.36
Log-likelihood=-22791						

Retail (n=3886)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{car-access}}$	Car accessibility	zone	logsum	0.0384	0.0110	3.50
$\beta_{\text{ret-den}}$	Density of jobs in retail sector	commune	jobs/hectare	0.1643	0.0371	4.43
$\beta_{\text{job-den}}$	Logarithm of jobs density	zone	ln(jobs/hectare)	0.0780	0.0153	5.09
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	-0.0036	0.0016	-2.19
β_{sqm}	Logarithm of non residential surface	building	ln(m^2)	0.8906	0.0174	51.24
Log-likelihood=-6443						

Hotels/Bar/Restaurants (n=2013)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{car-access}}$	Car accessibility	zone	logsum	0.0427	0.0133	3.21
$\beta_{\text{job-den}}$	Logarithm of jobs density	zone	ln(jobs/hectare)	0.3854	0.0169	22.82
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	-0.0076	0.0011	-7.10
β_{sqm}	Logarithm of non residential surface	building	ln(m^2)	0.3377	0.0142	23.73
$\beta_{\text{hbr-den}}$	Density of jobs in hotels/bar/restaurants	commune	jobs/hectare	0.2018	0.0193	10.44
Log-likelihood=-4923						

Government and public service (n=8471)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{off-den}}$	Density of jobs in private sector	commune	jobs/hectare	0.0125	0.0019	6.69
$\beta_{\text{job-den}}$	Logarithm of jobs density	zone	ln(jobs/hectare)	0.7523	0.0129	58.37
$\beta_{\text{pop-den}}$	Logarithm of population density	commune	ln(pop/hectare)	-0.0045	0.0006	-7.69
β_{sqm}	Logarithm of non residential surface	building	ln(m^2)	0.5081	0.0115	44.25
Log-likelihood=-10973						

Education (n=3775)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{edu-den}}$	Density of jobs in education sector	commune	jobs/hectare	0.2208	0.0157	14.08
$\beta_{\text{job-den}}$	Logarithm of jobs density	zone	ln(jobs/hectare)	0.1824	0.0161	11.37
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	-0.0075	0.0010	-7.65
β_{sqm}	Logarithm of non residential surface	building	ln(m^2)	0.8405	0.0183	46.04
Log-likelihood=-5995						

Health (n=5099)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{high-inc}}$	Percentage of households in high income scale (>3)	commune	%	0.0564	0.0057	9.86
$\beta_{\text{hea-den}}$	Density of jobs in health sector	commune	jobs/hectare	0.1832	0.0107	17.10
$\beta_{\text{job-den}}$	Logarithm of jobs density	zone	ln(jobs/hectare)	0.3708	0.0116	32.00
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	-0.0129	0.0010	-13.04
β_{sqm}	Logarithm of non residential surface	building	ln(m^2)	0.4908	0.0119	41.29
Log-likelihood=-10493						

Leisure activities (n=1315)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{high-inc}}$	Percentage of households in high income scale (>3)	commune	%	0.0837	0.0127	6.58
$\beta_{\text{leiz-den}}$	Density of jobs in leisure sector	commune	jobs/hectare	0.2978	0.0181	16.43
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	0.0133	0.0012	11.17
β_{sqm}	Logarithm of non residential surface	building	ln(m^2)	0.6327	0.0230	27.56
Log-likelihood=-2349						

4.2.6 Residential Development Project Location Choice Model (RDPLCM)

Estimation results for the Residential Development Location Choice model are presented in Table 4. The models are estimated over data for real estate developments that took place in the ten year period previous to the base year and, therefore, are not representative of all existing supply in the city. All types of residential development tend to agglomerate and therefore have a positive parameter for the logarithm of the number of buildings of the same type. The dwelling categories 'semi-detached' and 'attached' were grouped for the purpose of this research, because of their similar characteristics.

Residential buildings are developed in zones with high price number of residential units. The population density of the commune has a positive impact for semi-detached, attached and apartments, but is insignificant for detached houses.

4.2.7 Non-Residential Development Project Location Choice Model (NRDPLCM)

The Non-Residential Development Project Location Choice Model (NRDPLCM) models the location of the developed non-residential projects. Eight sub-models were estimated in UrbanSim, one for each of the building types: 1) industrial, 2) office, 3) shops, 4) hotels/bar/restaurants, 5) government and public service, 6) education, 7) health, 8) leisure activities. Since there were not a significant number of development projects in the past, sub-models regarding quarrying and agricultural buildings were not estimated.

The estimation results for the location choice model of non-residential real estate developments are shown in Table 5. New non-residential supply tends to locate in places that already show agglomeration and with high concentration of other activities in general. Locations with good car and public transport accessibility tend to be attractive for the location of new developments.

Development of projects of buildings that host private services is negatively affected by the population density of the commune and the logarithm of the total population in the zone, and positively by the logarithm of total number of jobs in zone. Another factor that affects the development of retail buildings is the number of jobs in the zone. The more jobs, the more preferable the zone is for the development of such infrastructure. The number of citizens in a zone is a positive determinant of the location, while the density at a commune level is negative.

4.2.8 Workplace Choice Model for Residents (WCMR)

This model assigns jobs to workers of the households. For its estimation, a table with each individual person with information about its household and at the base year (household_id and job_id), was created. The WCMR contains a single variable, the car accessibilities per zone, which is positive.

Table 4: **Residential Development Project Location Choice Model**

Detached (n=59558)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
β_{price}	Logarithm of price of detached houses	building	ln(euros)	1.5334	0.0259	59.31
β_{units}	Logarithm of number of detached house units	building	ln(sum)	1.6578	0.0049	338.21
Log-likelihood=-160082						
Semi-detached and Attached (n=20119)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
β_{price}	Logarithm of price of semi-detached and attached houses	building	ln(euros)	0.3013	0.0427	7.06
β_{units}	Logarithm of number of semi-detached and attached house units	building	ln(sum)	1.1172	0.0068	164.97
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	0.4097	0.0109	37.48
Log-likelihood=-58729						
Apartments (n=5119)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
β_{price}	Logarithm price of apartments	building	ln(euros)	0.1823	0.0764	2.38
β_{units}	Logarithm of number of apartment units	building	ln(sum)	0.1823	0.0764	2.38
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	1.0609	0.0124	85.62
Log-likelihood=-12286						

Table 5: **Non-Residential Development Project Location Choice Model**

Industry (n=2770)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{car-access}}$	Car accessibility	zone	logsum	-0.0659	0.0106	-6.25
$\beta_{\text{ind-den}}$	Density of jobs in industrial sector	commune	jobs/hectare	-0.0705	0.0068	-10.32
$\beta_{\text{ln-jobs-zone}}$	Logarithm of total number of jobs	zone	ln(sum)	0.4251	0.0128	33.19
Log-likelihood=-10949						
Office (private sector) (n=767)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{car-access}}$	Car accessibility	zone	logsum	0.0906	0.0251	3.61
$\beta_{\text{off-den}}$	Density of jobs in private sector	commune	jobs/hectare	0.0269	0.0081	3.32
$\beta_{\text{pop-den}}$	Population density	commune	pop/acre	-0.0318	0.0054	-5.86
$\beta_{\text{ln-jobs-zone}}$	Logarithm of total number of jobs	zone	ln(sum)	1.1741	0.0341	34.43
$\beta_{\text{ln-pop-zone}}$	Logarithm of total number of population	zone	ln(sum)	-0.1460	0.0275	-5.30
Log-likelihood=-1953						
Shops (n=1466)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{ln-jobs-zone}}$	Logarithm of total number of jobs	zone	ln(sum)	0.4451	0.0204	21.84
$\beta_{\text{ln-pop-zone}}$	Logarithm of total number of population	zone	ln(sum)	0.3899	0.0309	12.62
Log-likelihood=-5451						
Hotels, bar, restaurants (n=107)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{car-access}}$	Car accessibility	zone	logsum	0.2219	0.0812	2.73
$\beta_{\text{hbr-den}}$	Density of jobs in hotels/bar/restaurants	commune	jobs/hectare	0.1600	0.1027	1.56
$\beta_{\text{pop-den}}$	Population density	commune	pop/hectare	-0.0365	0.0126	-2.89
$\beta_{\text{ln-jobs-zone}}$	Logarithm of total number of jobs	zone	ln(sum)	0.7093	0.0805	8.81
Log-likelihood=-359						
Government and public service (n=264)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{ln-jobs-zone}}$	Logarithm of total number of jobs	zone	ln(sum)	0.7184	0.0461	15.57
$\beta_{\text{ln-pop-zone}}$	Logarithm of total number of population	zone	ln(sum)	0.1059	0.0472	2.24
Log-likelihood=-932						
Education (n=140)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{ln-jobs-zone}}$	Logarithm of total number of jobs	zone	ln(sum)	0.3591	0.0578	6.21
$\beta_{\text{ln-pop-zone}}$	Logarithm of total number of population	zone	ln(sum)	0.3539	0.0613	5.77
Log-likelihood=-533						
Health (n=225)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{ln-jobs-zone}}$	Logarithm of total number of jobs	zone	ln(sum)	0.3523	0.0646	5.45
$\beta_{\text{ln-pop-zone}}$	Logarithm of total number of population	zone	ln(sum)	0.5394	0.0715	7.55
Log-likelihood=-840						
Leisure activities (n=970)						
Variable	**Interpretation**	**Level**	**Unit**	**Coefficient**	**SE**	**t-values**
$\beta_{\text{car-access}}$	Car accessibility	zone	logsum	-0.0240	0.0141	-1.70
$\beta_{\text{lei-den}}$	Density of jobs in leisure sector	commune	jobs/hectare	3.2041	0.3811	8.41
$\beta_{\text{ln-jobs-zone}}$	Logarithm of total number of jobs	zone	ln(sum)	0.2371	0.0213	11.14
Log-likelihood=-3867						

5 Basecase Scenario

For validation purposes, a baseline scenario for the years 2001 to 2020 was simulated. The results indicate the strengths and the weaknesses of a zone-level integrated land-use and transport model in predicting the real socioeconomic and transport situation. The time interval for each simulation was set to 1 year for UrbanSim, and 9 years for MATSim (2001, 2010 and 2019). A 10% of the agents that use car for trips to work was randomly selected for the traffic simulation using MATSim.

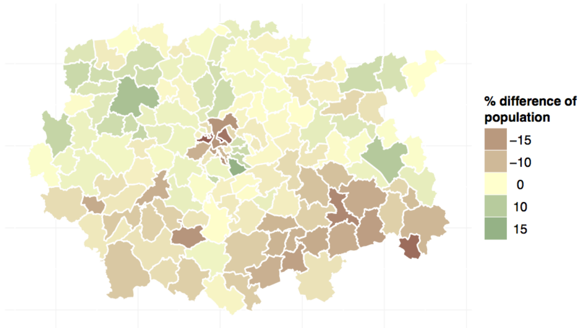
Figures 2(a), 3(a) and 3(b) demonstrate the efficiency of the model in predicting, with considerable accuracy, the real population in the great majority of the communes. However, the model under-predicts the population in the central communes *Brussels*, *Schaerbeek*, *Sint-Jans-Molenbeek* and *Saint-Josse-ten-Noode*, and the south-eastern *Wasseiges*. This can be interpreted by the under-prediction of the house prices in the city center, which leads –according to the household location choice model– less households to be located in that particular region.

Despite the good accuracy prediction of the real estate price model in 2002 3(c), and its success to capture the increased trend of low to medium prices in 2005 and 2008, as indicated in figures 3(d) and 3(e), it fails to predict the higher values in 2005 and 2008. These are mainly observed in the central commune *Brussels*, where the difference goes beyond 20% in 2008, and the south-eastern commune *Lasne*, which has the highest percentage of households with high income, as shown in figure 2(b). Figure 3(f) shows that the real estate price submodel of the apartments is even less accurate in predicting the average transaction prices of 2008, which occurs because of the low R^2=0.31 of this particular submodel. In general, despite the fact that the real estate price model considers socioeconomic dynamics, the hedonic price model partially fails to fulfill its purpose with success, probably because of the following reasons: 1) the hedonic price model does not capture market effects, such as supply or demand surplus (Hurtubia *et al.,* 2012), 2) the market clearing mechanism is oversimplified and may introduce a bias in the location choice of households and jobs, and 3) the ordinary least squares (OLS) fail to capture the spatial autocorrelation, an issue that can be solved by spatial econometric models (Efthymiou *et al.,* 2012*,* Efthymiou and Antoniou*,* 2013). Another reason could be the lack of variables related with building-specific attributes, as they were not available for the particular case study presented in this research.

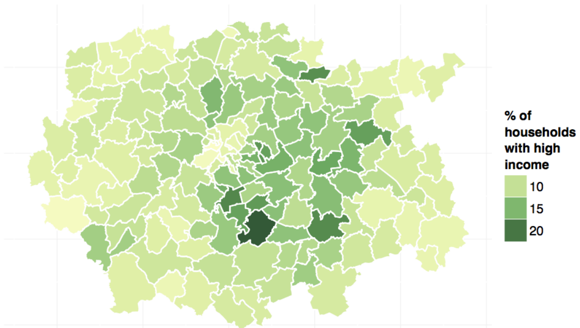
Figure 2(f) shows the car accessibilities per zone, which is the output of MATSim, in the year 2008. For the majority of the zones, the higher the distance from the city center, the lower its accessibility –with some exceptions in the outskirts of the area of study.

Figure 2: Validation Plots of Basecase Scenario Results

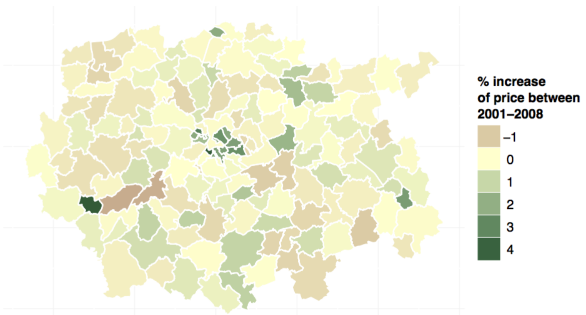
(a) Difference of Observed and Simulated Population in year 2008



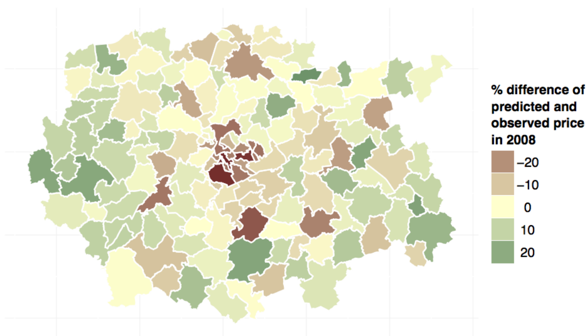
(b) Percentage of Households with High Income in 2008



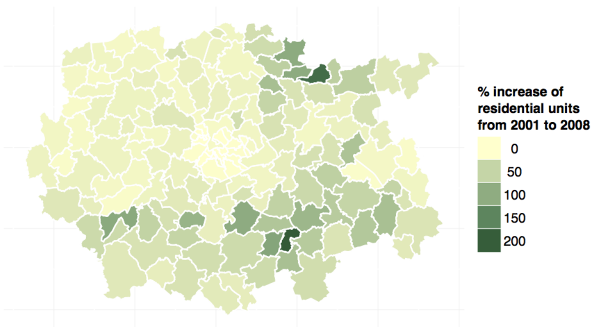
(c) Increase of Price Between 2001 and 2008



(d) Difference of Predicted and Observed Price in 2008



(e) Increase of residential units 2008



(f) Car accessibility in 2008

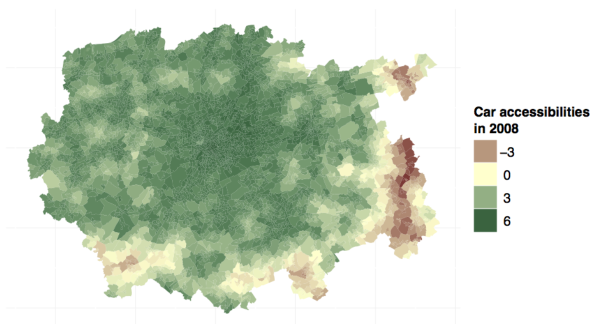
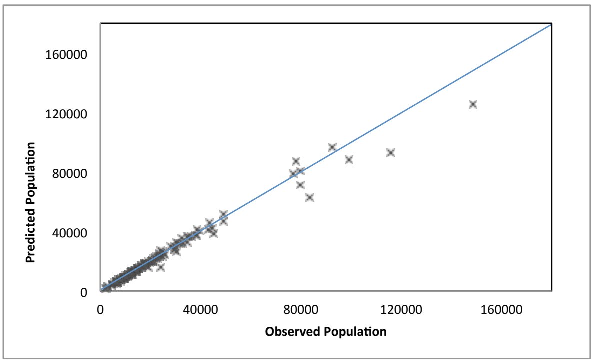
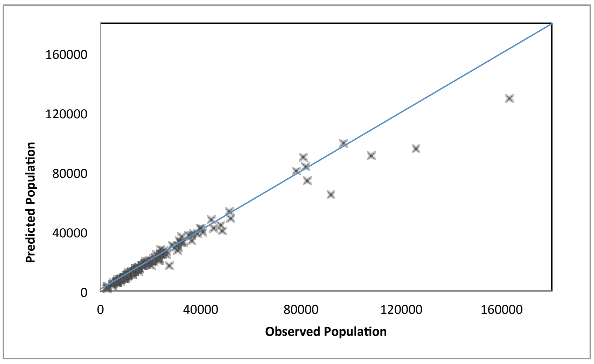


Figure 3: Validation Diagrams of Basecase Scenario Results

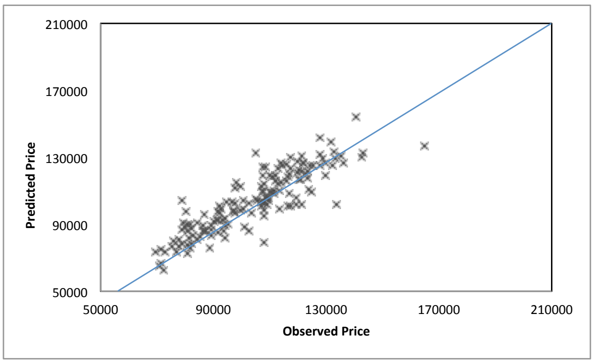
(a) Comparison of observed and predicted population in year 2008



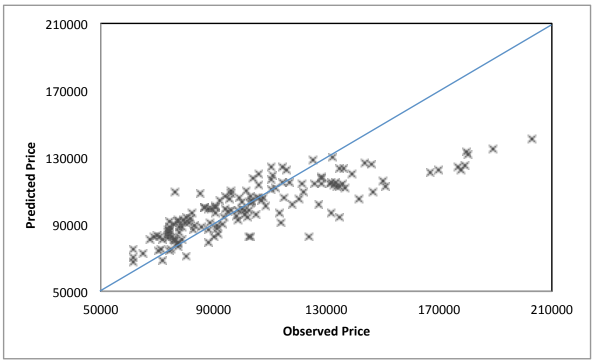
(b) Comparison of observed and predicted population in year 2011



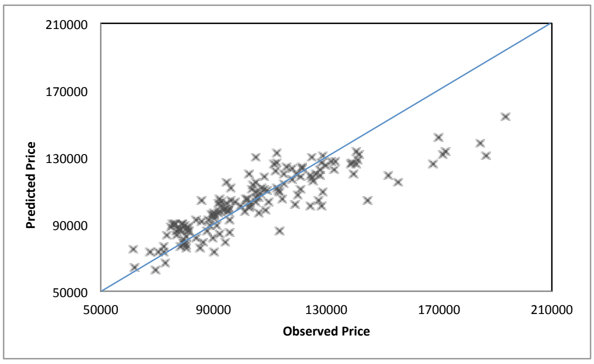
(c) Comparison of observed and predicted prices of houses in year 2002



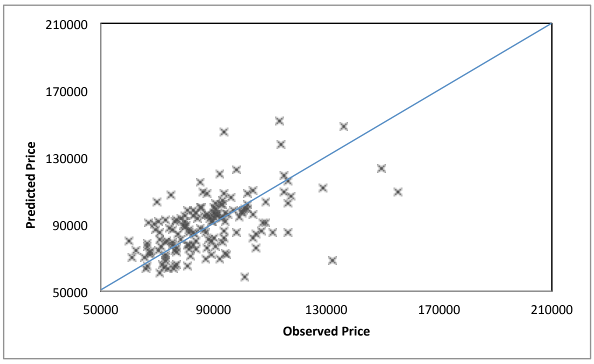
(d) Comparison of observed and predicted prices of houses in year 2005



(e) Comparison of observed and predicted prices of houses in year 2008



(f) Comparison of observed and predicted prices of apartments in year 2008



6 Conclusions and Recommendations

In this paper we present the development and implementation of the Integrated Land-Use and Transport Model for Brussels in Belgium. The model is developed using the agent-based microsimulation platform UrbanSim. The objective is to estimate the individual required models, interpret the estimated coefficients and identify the strengths and weaknesses of the modeling approach. This is done by validating the results of a base-case scenario, comparing them with available observed data and examining the possible margins of improvement.

For the purpose of this research, we used the LUTI model UrbanSim and the agent-based traffic simulation model MATSim. The individual transition, development, relocation, location choice and price models were estimated, and a base-case scenario from 2001 to 2013 was then simulated. Results show that the model succeeds in predicting the spatial distribution of location of new households; however it tends to underestimate the prices for communes were a significant increase was observed. This deviation may be explained by the use of an ordinary least square hedonic model for the real estate prices, making them dependent on attributes of the location but independent of market conditions. Another possible cause of the error in price forecast is the underestimation of dynamic variables that explain the price, like the income distribution or location of other agents in a zone, that could be due to the simplified market clearing process (first come first served) considered by UrbanSim.

Identification of the causes of this error is matter of future research but some potential ways to improve the modeling results are to include a more realistic market clearing mechanism (Hurtubia and Bierlaire*,* 2012), or the use of spatial autoregression models for the real estate price (Efthymiou *et al.,* 2012).

Acknowledgement

This research was partly funded through the SustainCity project, co-financed by the European Commission within the FP7 program. The first author was also benefited from the Swiss Government Excellence Research Scholarships, during his visit in Switzerland.

7 References

Banister, D. and Y. Berechman (2001) Transport investment and the promotion of economic growth, *Journal of Transport Geography*, **9**, 209–218.

Ben-Akiva, M. and S. R. Lerman (1985) *Discrete Choice Analysis*, MIT Press.

Charypar, D. and K. Nagel (2005) Generating complete all-day activity plans with generic algorithms, *Transportation*, **32** (4) 369–397.

Curtis, C. (2011) Integrating land use with public transport: The use of a discursive accessibility tool to inform metropolitan spatial planning in perth, *Transport Reviews: A Transnational Transdisciplinary Journal*, **31** (2) 179–197, March 2011.

De Palma, A. and N. Picard (2005) Route choice decision under travel time uncertainty, *Transportation Research Part A: Policy and Practice*, **39** (4) 295–324.

De Palma, A., S. Proost and S. van der Loo (2010) Assessing transport investments - towards a multi-purpose tool, *Transportation Research Part B*, **44**, 834–849.

Efthymiou, D. and C. Antoniou (2013) How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece, (in press), *Transportation Research Part A: Policy and Practice*.

Efthymiou, D., C. Antoniou and I. Tyrinopoulos (2012) Spatially aware model for optimal site selection, *Transportation Research Record: Journal of the Transportation Research Board*, **2276**, 146–155.

Farooq, B., M. Bierlaire and G. Flötteröd (2012) A new approach to synthesize heterogeneous agents and their associations for urban microsimulations, paper presented at the *Presentation at the 2nd Workshop on Urban Dynamics*, Termas de Chillán, Chile.

Franklin, J. P. and P. Waddell (2003) A hedonic regression of home prices in king county, washington, using activity-specific accessibility measures, paper presented at the *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*, Washington, D.C.

Geurs, K. T. and B. Wee (2004) Accessibility evaluation of land-use and transport strategies: review and research directions, *Journal of Transport Geography*, **12**, 127–140.

Gutiérrez, J. and P. Urbano (1996) Accessibility in the european union: the impact of the trans-european road network, *Journal of Transport Geography*, **4** (1) 15–25.

Hubert, J. P. and P. L. Toint (2002) La mobilité quotidienne des belges, *Mobilité et Transpors*, **1**.

Hunt, J. D., D. S. Kriger and E. J. Miller (2005) Currentoperational urban land-use-transport modelling frameworks: A review, *Transport Reviews: A Transnational Transdisciplinary Journal*, **25** (3) 329–376, May 2005.

Hurtubia, R. and M. Bierlaire (2012) Estimation of bid functions for location choice and price modeling with a latent variable approach, *Technical Report*, **TRANSP-OR 120206**, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

Hurtubia, R., M. Bierlaire and F. Martinez (2012) Dynamic microsimulation of location choices with a quasi-equilibrium auction approach, paper presented at the *Proceedings of the 12th Swiss Transport Research Conference*, Ascona, Switzerland.

Ibeas, Á., R. Vordera, L. dell'Olio, P. Coppola and A. Dominiguez (2012) Modelling transport and real-estate values interactions in urban systems, *Journal of Transport Geography*, **24**, 370–382.

Linneker, B. and N. Spence (1996) Road transport infrastructure and regional economic development, *Journal of Transport Geography*, **4** (2) 77–92.

Löchl, M. and K. Axhausen (2010) Modeling hedonic residential rents for land use and transport simulation while considering spatial effects, *The journal of transport and land use*, **2** (2) 39–63.

Löchl, M., M. Bürgle and K. W. Axhausen (2007) mplementierung des integrierten flächennutzungsmodells urbansim für den grossraum zürich – ein erfahrungsbericht, *Technical Report*, ETH.

Martinez, F. J. C. (1995) Access: the transport-land use economic link, *Transportation Research Part B*, **29** (6) 457–470.

Martinez, L. M. and J. M. Viegas (2013) A new approach to modelling distance-decay functions for accessibility assessment in transport studies, *Journal of Transport Geography*, **26** (87-96).

Medda, F. (2012) Land value capture finance for transport accessibility: a review, *Journal of Transport Geography*, **25**, 154–161.

Nicolai, T. W. (2012) Integrating matsim to urbansim for the integrating matsim to urbansim for the sustaincity studies brussels and zurich, *Technical Report*, Technische Universität Berlin.

Nicolai, T. W. and K. Nagel (2011) Matsim4urbansim – guide on urbansim usage of the travel model plug-in, *Technical Report*, Transport Planning and Transport Telematics (VSP), Transport Planning and Transport Telematic (VSP), Technical University of Berlin (TUB).

Patterson, Z. and M. Bierlaire (2010) Development of prototype urbansim models, *Environment and Planning B: Planning and Design*, **37**, 344–366, February 2010.

Patterson, Z. and R. Hurtubia (2008) A prototype integrated transportation land-use model for the lausanne region, *Technical Report*, École Polytechnique Fédérale de Lausanne, August 2008.

Patterson, Z., M. Kryvobokov, F. Marchal and M. Bierlaire (2010) Disaggregate models with aggregate data, *Journal of Transport and Land Use*, **3** (2) 5–37.

Picard, N. and C. Antoniou (2011) Econometric guidance for developing econometric guidance for developing urbansim models. first lessons from the sustaincity project, paper presented at the *Proceedings of the 51st European Congress of the Regional Science Association International*, Barcelona, Spain, 30th August - 3rd September 2011 2011.

Shaw, S.-L. and X. Xin (2003) Integrated land use and transportation interaction: a temporal gis exploratory data analysis approach, *Journal of Transport Geography*, **11**, 103–115.

Sustaincity (2009) Micro-simulation for the prospective of sustainable cities in europe, Seventh Framework Program FP7, Annex, October 2009.

Vandenbulcke, G., T. Steenberghen and I. Thomas (2007) Accessibility indicators to places and transports, *Technical Report*, Politique Scientifique Fédérale et SPF Mobilité et Transports, Brussels.

Vandenbulcke, G., T. Steenberghen and I. Thomas (2009) Mapping accessibility in belgium: A tool for land0use and transport planning?, *Journal of Transport Geography*, **17**, 39–53.

Waddell, P. (2002) Urbansim: Modelling urban development for land use, transportation and environmental planning, *Journal of the American Planning Association*, **68** (3) 297–314.

Waddell, P. (2010) Modeling residential location in urbansim, in F. Pagliara, J. Preston and D. Simmonds (eds.) *Residential Location Choice, Models and Applications*, 165–180, Springer Berlin Heidelberg.

Waddell, P. and G. Ulfarsson (2003) Dynamic simulation of real estate development and land prices within and integrated land use and transportation model system, paper presented at the *Proceedings of the 82nd Annual Meeting of the Transportation Research Board*.

Waddell, P., G. Ulfarsson, J. Franklin and J. Lobb (2007) Incorporating land use in metropolitan transportation planning, *Transportation Research Part A: Policy and Practice*, **41**, 382–410.