AESTIMUM 61, Dicembre 2012: 133-144

Marko Kryvobokov
Aurélie Mercier
Alain Bonnafous
Dominique Bouf

Laboratory of Transport Economics (LET), Lyon, France
E-mail:
mark.kryvobokov@let.ish-lyon.cnrs.fr
Aurelie.mercier@let.ish-lyon.cnrs.fr
Alain.Bonnafous@let.ish-lyon.cnrs.fr
Dominique.BOUF@let.ish-lyon.cnrs.fr

Key words: *real estate price model, validation, sensitivity*

Parole chiave: *metodo di stima prezzi immobili, analisi di sensitività*

JEL: *R32*

Simulating housing prices in Lyon with UrbanSim: predictive capacity and sensitivity analysis[1]

Housing prices in the Lyon Urban Area are simulated with the land use framework UrbanSim interacting with transportation model. We focus on the Real Estate Price Model of the UrbanSim framework. This OLS regression model of housing prices is calibrated using a nine-year back-casting period. The calibrated model, applied in simulation, provides price dynamics similar to actual one in the centre of Lyon. Sensitivity analysis demonstrates the model's ability to capture changes in employment accessibility on price dynamics. We conclude that the calibrated residential Real Estate Price Model from the UrbanSim application in Lyon is sensitive to changes in accessibility and provides good predictive capacity in the city centre, but underestimates prices in other areas.

1. Introduction

Transportation-land use modelling tools are important instruments capable to simulate urban development in a holistic sense. Recent literature reviews of these frameworks include the studies of Wegener (2004), Hunt *et al*. (2005), Chang (2006) and Iacono *et al*. (2008). In France, where regulation is particularly important in urban planning, transportation-land use models have a significant potential in academic research as well as in practical decision-making.

This study applies the land use simulation framework UrbanSim, whose highly disaggregate structure in comparison with other tools is recognised in the literature (Hunt *et al*. 2005; Iacono *et al*. 2008). This framework designed in the late 1990s at the University of Washington (Waddell 2002; Waddell *et al*. 2003) is distributed as an open source (UrbanSim Project 2011). UrbanSim is an operational modelling framework with multiple applications in the US as well as in other countries (Waddell *et al*. 2007; Felsenstein and Ashbel 2010; Löchl and Axhausen 2010; Di Zio *et al*. 2010). To our knowledge, French experience with UrbanSim in-

[1] The study is a part of the project PLAINSUDD (Innovative Numerical Platforms of Urban Simulation for Sustainable Development) sponsored through French ANR (number ANR-08-VD-00). Provision of data on real estate prices by Perval and Pierre-Yves Péguy and calculation of coordinates of real estate objects by Nicolas Ovtracht is acknowledged. The authors thank Mats Wilhelmsson for his suggestion of a price lag.

cludes the Paris and Lyon applications (de Palma *et al*. 2005; Patterson *et al*. 2010). Transportation model is external in UrbanSim. There are the following recent examples of integration of UrbanSim with transportation models: with the aggregated zone activity-based travel model (Waddell *et al*. 2010) and with the disaggregated "agent-based" MATSim (Nicolai *et al*. 2011).

Applying dynamic disequilibrium modelling approach, UrbanSim simulates on a yearly basis the spatial distribution of households, jobs, real estate development and real estate prices with specific models. Household and employment location choices are modelled with multinomial logit, while real estate construction and valuation are provided with ordinary least square (OLS) regression. In this paper, we focus on the residential Real Estate Price Model (REPM), which updates housing prices across an area under study. This model is an important component of the simulation framework, which, on the one hand, uses inputs from other models (including a transportation model), and, on the other, each simulation year provides other models with updated prices. The aim of the study is to calibrate the REPM in such a way that provides acceptable predictive capacity and sensitivity to the outcome of a transportation model. The former point, crucial for realistic simulation, is addressed with historical validation. In the latter issue, which is an important feature of land use and transport interaction (LUTI), accessibility indices are exploited.

The rest of the paper is structured as follows. The next section contains the overview of the Lyon UrbanSim application. Section 3 is devoted to the preparation of data on real estate prices. Section 4 describes in detail the residential REPM calibration applying a back-casting period. Land use simulation takes place in Section 5. Sensitivity analysis is described in Section 6. The final section concludes.

2. The UrbanSim application

In Lyon, the prototype UrbanSim application was gridcell-based with one gridcell per zone (Patterson *et al*. 2010). After comparison of simulation at different geographical levels, the municipal level of a zone-based application was chosen (Bonnafous *et al*. 2010). In the current study, the zone-based UrbanSim version 4.3.2 is applied, where municipalities are used as zones.

Geographically, the area of study is the Lyon Urban Area with the 1999 population of 1.6 million inhabitants. It consists of 304 municipalities (communes and arrondissements). The cities of Lyon and Villeurbanne and the urbanized belt around them are named the Greater Lyon. Lyon and Villeurbanne compose the central part of the urban area. Lyon consists of nine districts (arrondissements), which are considered in our study as municipalities. The available data include a number of residential units and area covered by different land use types (water, industry, etc.) in municipalities.

Data on population are available due to the last general census conducted in France in 1999. This is the base year in our dataset. In the available synthetic population 1999 based also on the household travel survey 2006, there are 662,249 households. In the applied version of UrbanSim, each household lives in a building. However, because of the lack of available data about buildings, in our application we have created one fictional building in each municipality.

For each household, there are socio-economic data, including number of persons, number of cars and income group, to whom it belongs. The employment data 1999 in municipalities contain a number of jobs by economic sector.

UrbanSim models are estimated for the base year 1999. Simulation is run for subsequent years. The Household Location Choice Model distributes population applying in its utility function, among other variables, housing price interacted with household income groups (see Bonnafous *et al.* 2010), while the distributions of jobs (by sector) and real estate development are simulated with simple OLS regression models. Thus, in the final specification, a number of residential units is a linear regression function of number of households. In simulation, a number of residential units in each municipality is updated with this regression under the condition that this number is higher than in a previous year. In the historical districts Lyon 1 and Lyon 4 located between the two rivers, there is no free space for new residential development, and new construction in fact does not take place; therefore a number of residential units is not updated there. The four-step transportation model, external to UrbanSim, is provided by the MOSART[2] platform developed in Visum from PTV (Crozet *et al.* 2008; Bonnafous *et al.* 2011). The next Section describes the process of data preparation for the residential REPM.

3. Data preparation

The available data for the residential REPM contains the sale prices of 9,231 apartments and 2,176 individual houses sold in the Lyon Urban Area in 1997-2008. The idea to use a one-year temporal lag in the model (see Section 4) needs the housing prices recalculated to 1999 and 1998 for model estimation.

There are 9,958 apartment and housing sales in 1999-2008 and 1,449 apartment sales (in Lyon and some municipalities of the Greater Lyon) in 1997-1998. Applying the annual indices calculated previously for apartment prices from the OLS regression model (see Kryvobokov 2010), the prices of the former group are recalculated to 1999, while the prices of the latter group to 1998.

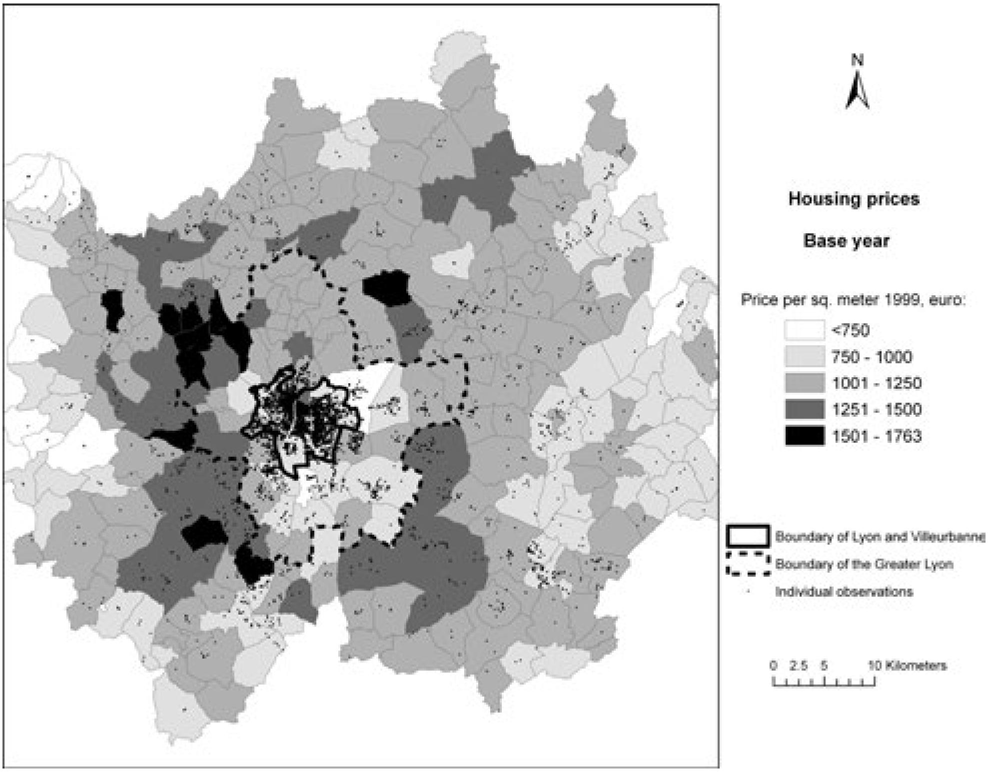
Because of insufficient geographical coverage of the latter group (e.g. only one observation in Villeurbanne), the following calculations are done. Apartment prices 2001 in Vileurbanne (63 observations) and housing prices 2002 (111 observations located outside the Greater Lyon) are recalculated to 1998, excluded from the former group and added to the latter. Each of the groups covers almost the whole Lyon Urban Area (see Figure 1 for the former one).

---

[2] *MOdélisation et Simulation de l'Accessibilité aux Réseaux et aux Territoires* (Modelling and Simulation of Accessibility to Networks and Territories).

In both cases the prices represented in points are interpolated to raster with ArcGIS Spatial Analyst. The Inverse Distance Weighted method is applied with 12 neighbours, power 2 and cell size of 10 meters. With the Zonal Statistics tool of Spatial Analyst, average price per square meter is calculated in each of 304 municipalities for 1998 and 1999. The descriptive statistics is presented in Table 1, while Figure 1 shows the geographical distribution of average prices 1999.

Figure 1. Individual observations and average prices per square meter 1999.



In the described approach to data preparation, price indices refer to apartments, while they are applied to both apartment and housing market segments. It is the reason of a bias in the regression model. At the stage of price interpolation from points to raster, there is one more source of errors. While Lyon and Villeurbanne are quite well covered by observations, not all residential areas in the suburbs have enough data. For example, in some prestigious western suburbs (Écully, Dardilly) there are no observations at all. This drawback in addition to that connected with indices decreases spatial consistency and reliability of the result (see Briassoulis 2001).

One more important limitation of the REPM refers to the UrbanSim modelling approach. In its modelling framework, there is no place for individual real estate observations. In the zonal application in the Lyon Urban Area, there are only 304 municipalities, at which level the initial individual prices are aggregated. Consequently, structural attributes of apartments and houses cannot be included into the list of explanatory variables that gradually decreases their number (see also Löchl and Axhausen 2010).

Urban theory strongly suggests that accessibility has crucial impact on real estate prices. The accessibility index used in this study is calculated applying the MOSART platform with the following formula:

$$A_i = \sum_{j=1}^{n} E_j e^{-\beta C_{ij}} \text{ , (1)}$$

where:
E_j – number of jobs in zone j;
C_{ij} – travel cost between zone i and zone j, which includes monetary cost and value of time;
β – cost sensitivity parameter;
n – number of zones.

This accessibility measure provides a link between transportation and land use models on our study. The descriptive statistics of this index as well as of some important demographic and real estate attributes is presented in Table 1.

Table 1. Descriptive statistics.

Attribute	Mean	Minimum	Maximum	Std. dev.
Real estate price 1998, euro per square meter	1,060	586	1,720	190
Real estate price 1999, euro per square meter	1,109	601	1,763	195
Employment access, index	48,648	2,577	251,585	51,703
Residential units	2,432	54	63,449	6,585
Residential vacancy rates	0.11	0.02	0.35	0.06
Population	5,293	100	118,589	12,263

4. Model calibration

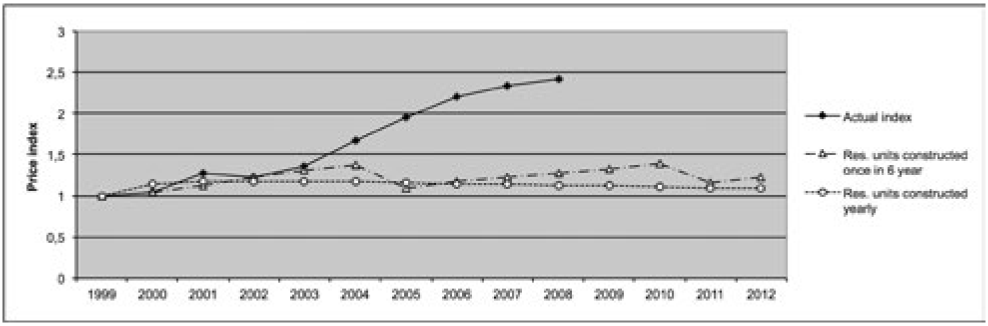
The initial model specification included population, cars to population ratio, vacant residential units, accessibility measure and land use attributes of municipalities. The model goodness-of-fit was 39.2%[3]. Calibrating the REPM, different

3 In Kryvobokov (2010), with more than 4,000 individual observations located in the central part of the Lyon Urban Area, it was obtained the adjusted R^2 of the OLS and GWR models in the

sub-scenarios were tested without actual data on residential development during simulation.

Without new residential construction, simulation fails after eight years – there are no more residential units to allocate a growing population. It was implemented the attempt to add residential units once in six years, i.e. in 2005, 2011, etc. A number of residential units in each municipality was calculated on the base of the ratio of residential units to households in 1999; during simulation, from the updated ratio it was subtracted some ad hoc percentage on order to keep constant an overall vacancy rate. Figure 2 shows that the average housing price in the Lyon Urban Area is very sensitively react to increased residential supply: the years of new construction are characterized by slumps in the simulated price curve, which is very different from the actual price index.

Figure 2. Price index in the Lyon Urban Area.



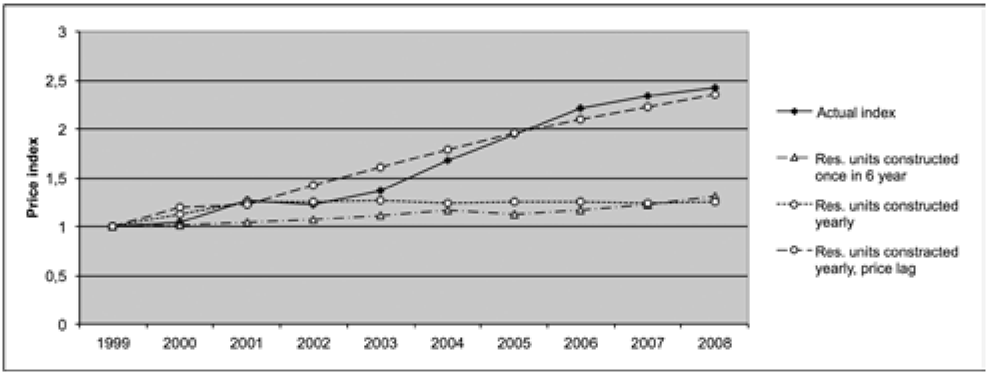
An attempt to update residential supply yearly as a linear OLS regression function of number of households leads to more stable, but slowly decreasing average price, which again differs a lot from the actual index curve (Figure 2).

Thus, the initial specification lacks some important component(s) influencing price dynamics. Literature on real estate price indices suggests error correction models calculating annual real change in price as a function of mortgage interest rate, annual change in real income per capita and other attributes (e.g. Malpezzi 1999). However, interest rate is not changed with location and therefore cannot be exploited in our municipality-based model. Instead, we apply a temporal one-year lag in price in each municipality. Thus, the REPM estimated for 1999 included the price 1998 as one of the explanatory variables. During simulation, the lagged structure is kept dynamic, e.g. in 2001 one of the variables is price 2000, etc. Though a price lag does not make the simulated dynamics much closer to ac-

interval 85-88%, while the pseudo-R^2 of the spatial lag and spatial error models was in the interval 85-86%.

tual index on average, its effect is visible in the central part of the agglomeration, for which the actual index was calculated. Figure 3 shows how the prediction is improved in Lyon 1 when a price lag is added.

Figure 3. Price index in Lyon 1.



The model specification with the price lag is presented in Table 2 and used in simulation. The adjusted R^2 of this OLS regression is 88.7%. A one-year price lag is the most significant variable. The demand side of the housing market is represented by population and the specific variable of the number of vacant residential units, which also reflects the supply of housing stock. Employment accessibility, population, cars to population ratio and percentage of area covered by water positively affect housing price. Vacant residential units and industrial area have negative influence.

Table 2. REPM estimation.

Variable	Coefficient (*t*-value)
Constant	5.055 (34.26)
Log price lag	0.824 (36.67)
Log employment access	0.034 (5.50)
Log population	0.025 (2.84)
Log cars to population ratio	0.320 (2.59)
Log vacant residential units	-0.023 (-3.45)
Log industrial area	-0.002 (-2.61)
Log percent water	0.009 (2.10)

5. Simulation

Actual price index in Lyon and its closest suburbs is known till 2008. During the back-casting simulation 2000-2008, employment accessibility indices are recalculated with MOSART once, in 2006, while the land use models update their outcomes yearly. The dynamics of housing price per square metre is shown in Table 3, where average price in the Lyon Urban Area is reported as well as the price in Lyon 1. In this reference scenario, during the first simulation decade, the average price is growing slowly, which can be explained by a big number of small municipalities, mainly rural ones located on the fringe. In the central part of the agglomeration, price dynamics is pretty similar to actual index as the example of Lyon 1 shows. Figure 4 exhibits housing price growth after nine years of simulation. Note that the highest growth is observed in Dardilly, to the north-west from Lyon, where interpolated price 1999 was lower than actual one due to the lack of data. Figure 5 shows the ratios of the simulated prices 2008 to the average actual prices in that year in the districts of Lyon and some adjacent municipalities, where a number of actual sales 2008 is more than 10. In Lyon 1, the simulated housing prise composes 99% of the actual one. In other municipalities, despite the inclusion of price lag and population in the regression equation, predicted prices are lower than actual ones.

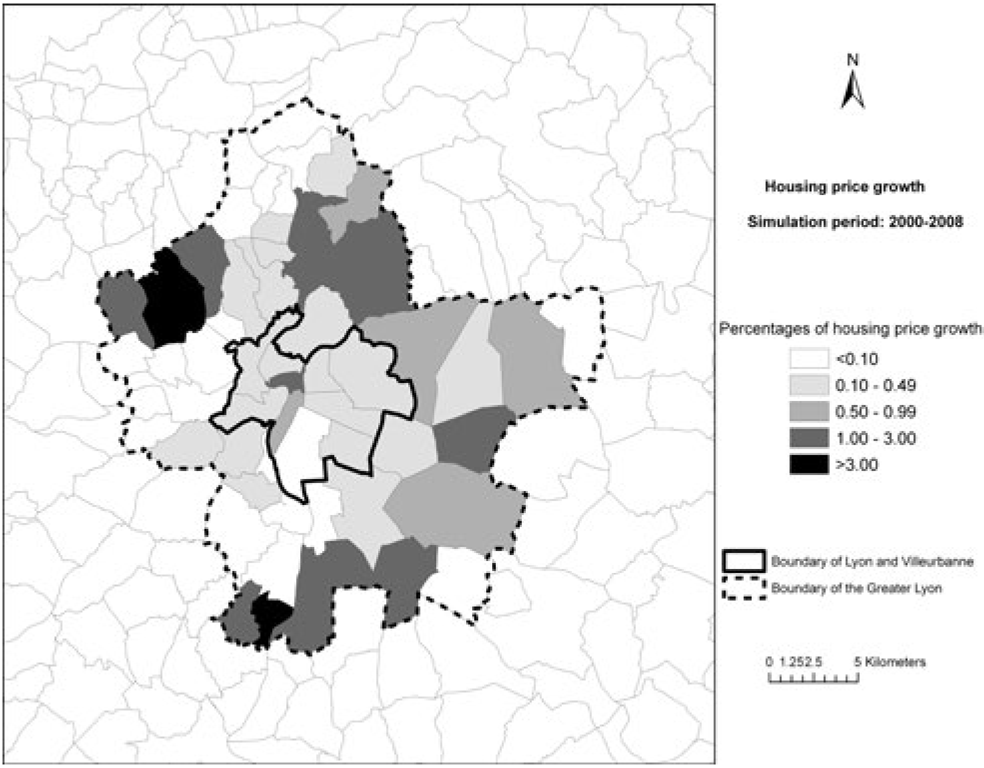
Table 3. Housing price dynamics.

Year	Actual price index	Reference scenario	
		Average	Lyon 1
1999	1.00	1.00	1.00
2000	1.04	0.99	1.02
2001	1.13	1.02	1.22
2002	1.23	1.05	1.42
2003	1.37	1.07	1.61
2004	1.67	1.08	1.79
2005	1.95	1.09	1.96
2006	2.21	1.10	2.10
2007	2.34	1.11	2.23
2008	2.42	1.12	2.35

6. Sensitivity analysis

In this section, we check the sensitivity of the residential REPM to changes in accessibility indices during simulation. For this, the indices from MOSART are in-

Figure 4. Simulated price growth 1999-2008.



creased by 10% and decreased by 10% starting from the second simulation year. In these two alternative scenarios, population, jobs and real estate prices are not only interdependent as they were in the previous simulation, but their dynamics also depends on the changes in the accessibility measures. The changes in the housing price dynamics on average and in Lyon 1 (Table 4) are in line with the base of urban theory: prices increase if accessibility improves and vice versa. When the accessibility indices increase, the average price increases as well, though its dynamics grows by less than 1% during the second and the third years after the change. In the centre of Lyon, the housing price increases a bit faster, reaching the 3% increase in the dynamics by the end of the simulation period. The REPM is more sensitive to the decrease in the accessibility indices: the price index dynamics decreases by 2% on average and by 5% in Lyon 1 in 2008.

7. Conclusion

The Real Estate Price Model as a component of the UrbanSim land use simulation framework is focused in this study. The regression model of housing prices

Figure 5. Ratios of simulated to actual prices 2008.

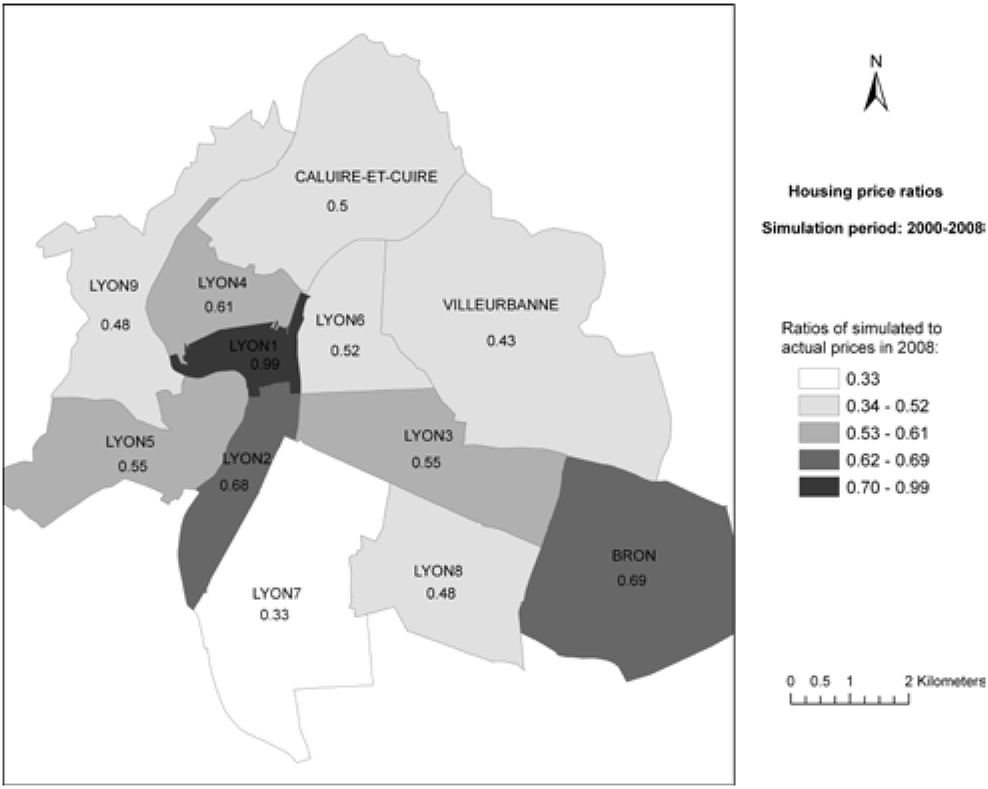


Table 4. Changes in housing price index dynamics.

Year	Accessibility increased by 10%		Accessibility decreased by 10%	
	Average	Lyon 1	Average	Lyon 1
2001	+0.01	0.00	0.00	-0.01
2002	0.00	0.00	-0.01	-0.01
2003	0.00	+0.01	-0.01	-0.01
2004	+0.01	+0.02	-0.01	-0.02
2005	+0.02	+0.02	-0.01	-0.03
2006	+0.02	+0.03	-0.01	-0.03
2007	+0.02	+0.03	-0.01	-0.04
2008	+0.02	+0.03	-0.02	-0.05

is calibrated in the Lyon application during a nine-year back-casting period. The other UrbanSim models accompany this process updating the geographical distributions of households, jobs and housing stock. The transportation model MOSART calculates the accessibility measure interacting with UrbanSim.

In the process of the housing price model calibration, a one-year price lag was found the most significant variable. With this variable, the residential Real Estate Price Model, in spite of its objective problems, in particular the data bias and the zone approach, where there is no place for individual observations, nevertheless, provides a dynamics of simulated prices, which is in line with actual price index in the centre of Lyon, though in other areas, for which actual data exist, price dynamics is underestimated.

The test of sensitivity of the UrbanSim housing price model to changes in the accessibility measure during simulation demonstrates that the output changes correspond to the base of urban theory: real estate prices increase if accessibility improves and vice versa. Thus, this element of the LUTI structure is working correctly and can be used in projects simulating different scenarios of urban development.

Bibliography

Bonnafous A., Crozet Y., Mercier A., Ovtracht N., Péguy P.-Y. and Puech F. (2011). MOSART et le projet PLAINSUDD : une plate-forme de modélisation et simulation de l'accessibilité pour l'aide à la décision et l'aménagement du territoire. In: Antoni J.-P. (Ed.) *Modéliser la ville : formes urbaines et politiques de transport*, Economica: 186-210.

Bonnafous A., Kryvobokov M., Bouf D. and Ovtracht N. (2010). *Simulating residential location choice in the Lyon Urban Area*. Conference paper, 57th Annual North American Meetings of the Regional Science Association International, Denver, Colorado, 10-13 November.

Briassoulis H. (2001). Policy-oriented integrated analysis of land-use change: An analysis of data needs. *Environmental Management* 27(1): 1–11.

Chang J. S. (2006). Models of the Relationship between Transport and Land-use: A Review. *Transport Reviews* 26 (3): 325-350.

Crozet Y., Marchal F., Ovtracht N., Thiebaut V., Yvon C. and Bischoff P. (2008). *Mise en place d'un outil d'aide à la décision et à l'aménagement du territoire : MOSART (MOdélisation et Simulation de l'Accessibilité aux Réseaux et aux Territoires)*. Conference paper, SAGEO Conference, Montpellier, France, 24-27 June.

De Palma A., Motamedi K., Picard N. and Waddell P. (2005). A model of residential location choice with endogenous housing prices and traffic for Paris region. *European Transport* 31: 67-82.

Di Zio S., Montanari A. and Staniscia B. (2010). Simulation of urban development in the City of Rome: Framework, methodology, and problem solving. *Journal of Transport and Land Use* 3 (2): 85-105.

Felsenstein D. and Ashbel E. (2010). Simultaneous modelling of developer behavior and land prices in UrbanSim. *Journal of Transport and Land Use* 3 (2): 107-127.

Hunt J. D., Kriger D. S. and Miller E. J. (2005). Current Operational Urban Land-use-Transport Modelling Frameworks: A Review. *Transport Reviews* 25 (3): 329-376.

Iacono M, Levinson D. and El-Geneidy A. (2008). Models of Transportation and Land Use Change: A Guide to the Territory. *Journal of Planning Literature* 22 (4): 323-340.

Kryvobokov M. (2010). Is it worth identifying service employment (sub)centres when modelling apartment prices? *Journal of Property Research* 27 (4): 371-390.

Löchl M. and Axhausen K. W. (2010). Modelling hedonic residential rents for land use and transport simulation while considering spatial effects. *Journal of Transport and Land Use* 3 (2): 39-63.

Malpezzi S. (1999). A Simple Error Correction Model of House Prices. *Journal of Housing Economics* 8 (1): 27-62.

Nicolai T. W., Wang L., Nagel K. and Waddell P. (2011). *Coupling an urban simulation model with a travel model – A first sensitivity test*. Working paper, Technical University of Berlin and University of California, Berkeley, 24 p.

Patterson Z., Kryvobokov M., Marchal F. and Bierlaire M. (2010). Disaggregate model with aggregate data: Two UrbanSim applications. *Journal of Transport and Land Use* 3 (2): 5-37.

UrbanSim Project (2011). *Opus: The Open Platform for Urban Simulation and UrbanSim Version 4.3. Users Guide and Reference Manual*. University of California, Berkeley and University of Washington, Seattle.

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

Waddell P., Ulfarsson G., Franklin J. and Lobb J. (2007). Incorporating land use in metropolitan transportation planning. *Transportation Research A* 41 (5): 382–410.

Waddell P., Wang L., Charlton B. and Olsen A. (2010). Microsimulating parcel-based land use and activity-based travel: Development of a prototype application in San Francisco. *Journal of Transport and Land Use* 3 (2): 65-84.

Wegener M. (2004). *Overview of land-use transport models*. In: Hensher D. A. and Button K. (Eds.) Transport Geography and Spatial Systems, Handbook 5 of the Handbook in Transport, Kidlington, the UK: Pergamon/Elsevier Science: 127-146.