Bid auction model for simultaneous determination of location and rent in land use microsimulation models

Ricardo Hurtubia * Francisco Martínez † Michel Bierlaire *

September 19, 2011

XV Congreso Chileno de Ingeniería de Transporte
3-6 de Octubre, 2011
Santiago, Chile

Abstract

A method for the (micro)simulation of location choice by agents in a urban context is proposed. The method is based in the bid-auction approach for land use modeling, which assumes that rents can be estimated as the expected maximum bid in an auction. The method allows for period-wise simulation of location choice where rents are adjusted depending on the household's perceptions of the market conditions. The location can be modeled both as an auction or as a direct choice, depending on the market conditions. This allows for the simulation of both demand and supply surplus scenarios in a consistent way. A new estimation methodology for bid-auction choice models is also proposed, where a price indicator is included in the maximum likelihood process. The method generates estimates that are able to reproduce both the observed locations and observed prices at the base year of the simulation.

*Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, ricardo.hurtubia@epfl.ch
†División Transporte, Departamento de Ingeniería Civil, Universidad de Chile

1 Introduction

Land use models are an increasingly used tool for evaluation and forecasting of the effects of urban interventions such as real estate developments, modifications to the transport system and changes in urban policy. Among these, microsimulation models are becoming more relevant due to the possibility of representing individual agents and their complex interactions in a simple, yet robust and flexible, way.

Modeling the location choice of the different agents that interact in a city is one of the main objectives of any land use model. Location choice and real estate prices have been traditionally modeled under two different paradigms: the choice approach and the bid-auction approach. Under the choice paradigm, households select the location that maximizes their utility, with prices being determined exogenously through a hedonic model. The bid-auction approach assumes that real estate goods are traded in an auction market, where the best bid for a particular location determines both the located household and the price of the dwelling.

Both bid-auction and choice approaches work under the assumption that prices will be properly estimated only under equilibrium conditions. In the choice case the hedonic approach for modeling prices implicitly assumes that the (equilibrium) market values of the attributes of a location are represented in the parameters of a regression. The bid-auction approach can only determine prices when all households have interacted in all the possible auctions, achieving a state where no household can improve his situation by changing its location.

The underlying equilibrium assumption makes hard to implement either approach directly in a microsimulation context, where equilibrium is never solved but, instead, a dynamic process approximates the equilibrium conditions by simulating all the individual interactions in the market. For operational reasons, microsimulation models usually favor a choice approach, estimating the hedonic price model for a base modeling period and ignoring the equilibrium assumption. This means that prices are insensitive to changes in the market conditions (e.g. income distribution across the population, supply or demand surplus), making the market values of each of the attributes of a location constant in time.

On the other hand, the bid auction approach can handle the effect of changes in the market conditions because prices are a function of the preferences of the households, and bids can be adjusted to react against an increase or decrease of supply/demand. However, this approach has only been implemented in aggregated, equilibrium based, models.

This paper proposes a method to model location choice and real estate prices simultaneously in a microsimulation context. The method is based on the bid-auction approach and estimates both location and prices as a function of the households' preferences. The proposed approach does not require solving for equilibrium, but estimates the maximum bid in each period by simulating the underlying auction process.

Given exogenous supply levels, households adjust their preferences (and their willingness to pay) as a reaction to the (observed) market conditions. Demand surplus generates a more competitive market from the demand point of view, with households competing for a scarce number of locations.This triggers an increase in the willingness to pay of all households and, therefore, the rents of real estate goods. On the opposite case, a supply surplus scenario will generate a more competitive market from the supply point of view, with developers/owners willing to sell/rent the dwellings at lower prices, given the reduced demand. This is translated in a reduction of the willingness to pay of all households and the consequent reduction in rent levels.

The type of auction also depends on the market conditions. In the case of demand surplus the markets behaves like a classical auction, with several agents bidding for few goods. In this case the location choice is modeled following a maximum bid probability. When the market interactions take place under supply surplus conditions the market also clears through an auction, but in this case several goods compete to be chosen by few agents. In this case the location choice is modeled following the classical maximum utility choice probabilities.

In order to implement this model, the estimated parameters of the bid function should be able to forecast both the location choice distribution and the prices (as the expected outcome of the auction process) at the base period of the simulation. To achieve this, a novel approach for model estimation is proposed, where the traditional maximum log-likelihood methodology is modified in order to account not only for observed location, but also for observed prices at a given base period. The method is based in the Generalized Random Utility model, originally proposed by Walker and Ben-Akiva (2002) and integrates an indicator in the log-likelihood function in order to include an additional measurement relationship in the model.

The paper is organized as follows: Section 2 describes the main theory behind the bid-auction approach to location choice modeling. Section 3 explains the choice approach and how it can be consistently used in a bid-auction framework. Section 4 proposes a model that combines the bid and choice approaches in a microsimulation context. Section 5 describes the proposed estimation method for location and prices and shows some results for a real city case study. Section 6 describes an experiment with synthetic data to analyze the reaction-capacity of the model to dynamic changes in the real estate market conditions. Finally, Section 7 concludes the paper and identifies possible further research.

2 The bid approach

Since Alonso (1964), the real estate market has been understood as an auction market, where households bid their willingness to pay for a particular good (residential unit) which is assigned to the best bidder. This process simultaneously defines the price of the good, understood as the maximum bid in the auction process.

The willingness to pay, from an economic point of view, can be derived from the classical consumer's problem of maximum utility, given income constraints:

$$\max_{x,i} U(x, z_i) \tag{1}$$

$$\text{s.t. } px + r_i \leq I$$

In the previous problem, the consumer maximizes his utility by choosing a vector of continuous goods (x) and a discrete location (i), described by a set of attributes (z_i). The budget constraint states that the total amount spent in goods (with price p) plus the price of the selected location (r_i) must be smaller that the consumer's available income (I). Solving the problem on x and assuming equality in the budget constraint, the problem can be re-written as

$$\max_{i} V(p, I - r_i, z_i) \tag{2}$$

where V is the indirect utility function, conditional on the the location. Conditional on the level of maximum utility ($\overline{U}$), the indirect utility can be inverted in the rent variable:

$$r_i = I - V^{-1}(\overline{U}, p, z_i) \tag{3}$$

Under the auction market assumption, the rent variable can be understood as the willingness to pay for a particular location, therefore the bid function B can be expressed as:

$$B_{hi} = I_h - V_h^{-1}(\overline{U}, p, z_i) \tag{4}$$

In the bid function, the index h has been included to take into account heterogeneity in preferences within different households. If we assume bids to be random variables, with an extreme value distributed error term, it is possible to express the probability of a household (h) being the best bidder for a particular location (i) as follows:

$$P_{h/i} = \frac{\exp(\mu B_{hi})}{\sum_g \exp(\mu B_{gi})} \tag{5}$$

Under the auction market assumption, the price or rent of a good will be the maximum bid. The extreme value distribution assumption allows to express the expected maximum bid for a particular location as the logsum of the bids

$$r_i = \frac{1}{\mu} \ln\left(\sum_g \exp(\mu B_{gi})\right) \tag{6}$$

The bid approach has been traditionally implemented in equilibrium based models like MUSSA (Martínez, 1996), where rents can only be determined when bids have been adjusted to ensure that each household is located somewhere and in not more than one location. This means that the utility level (and therefore the bid level) of each household should be adjusted to ensure that:

$$\sum_i P_{h/i} = 1 \tag{7}$$

The previous condition is only possible when an absolute equality between supply (the number of location alternatives) and demand (the number of households) holds, meaning that:

$$\sum_h \sum_i P_{h/i} = H = S \tag{8}$$

with H the total number of households and S the total number of locations.

3 The choice approach

The choice approach (McFadden, 1978; Anas, 1982) assumes that households choose the location that maximize their utility. The utility a household perceives is the indirect utility function (2) and can be defined as a function of the attributes of the location ($V_{hi} = f(z_i)$). Assuming an extreme value distribution for the error term of the utility function, the probability of a household h choosing a location i is:

$$P_{i/h} = \frac{\exp(\mu V_{hi})}{\sum_j \exp(\mu V_{hj})} \tag{9}$$

It is possible to demonstrate that, under the assumption of an auction market, the location where the agent is the highest bidder is also that of the maximum surplus or maximum utility (Martinez, 1992, 2000). This assures that the auction outcome yields an allocation consistent with maximum utility behavior of consumers. The consumer surplus is defined as the difference between the willingness to pay for a good and the actual price of the good. If the utility is written in terms of consumer surplus it will take the following form:

$$V_{hi} = B_{hi} - r_i \tag{10}$$

Replacing (10) in (9), the probability of a household h choosing a location i is:

$$P_{i/h} = \frac{\exp(\mu(B_{hi} - r_i))}{\sum_j \exp(\mu(B_{hj} - r_j))} \tag{11}$$

If prices are the outcome of an auction process and the market clears, the distribution of households across locations obtained through (11) will be the same as the distribution obtained from (5) when supply and demand are equilibrated. Otherwise, the choice approach is only valid when there are more alternatives than decision makers (Martinez, 1992).

4 Bid rent model for microsimulation

Microsimulation of land use requires a representation at the individual level of the location choice and price formation processes. This means that each household is paired to a location in a sequential way.

The choice approach is straightforward to implement in a microsimulation context because it provides the individual location probabilities and rents are calculated exogenously (and independently) for each dwelling following an hedonic model without requiring any assumption about equilibrium between supply and demand. However, implementing the choice approach requires the assumption that supply will always satisfy (or exceed) demand, so the allocation process can be simulated by drawing a location for each household. The order in which the allocation happens can only be assumed to be random, drawing the location for each household at a time and making selected location unavailable for future choices. If a demand surplus scenario happens, a choice approach will only be able to deal with this by randomly selecting households that will not be located. The choice approach also presents the drawback of using hedonic prices, which make the price formation process exogenous to the location choice problem and independent of changes in the market conditions. An analysis of the disadvantages of using hedonic prices and the differences between them and maximum bid prices can be found in Hurtubia et al. (2010).

Implementing a bid approach is not straightforward, because prices can only be determined if a supply-demand equilibrium is achieved and bids are adjusted to this. The complexity comes from the fact that equality between demand and supply is usually not guaranteed in a microsimulation (because of an independent supply generation process). Also, the bid approach traditionally assumes that each location "chooses" a household through the auction process, therefore making hard to simulate scenarios with supply surplus (there is no clear rule to decide which locations are not used). This difficulties are addressed and partly overcome in Martínez and Hurtubia (2006), but in an aggregated, quasi-equilibrium context.

We propose a model where, at each period of time, the auction for each good is simulated, therefore obtaining rent levels that reflect the competition between different bidders for the good. The adjustment accounts for the effect a supply or a demand surplus will have on the bids. We solve the allocation problem by proposing a different market clearing solution depending on the supply/demand surplus conditions of the scenario.

We assume the bid function to be composed of two elements, therefore, for a particular period t:

$$B^t_{hi} = b^t_h + b_{hi}(z^t_i, \beta) \tag{12}$$

where b^t_h is the adjustment component that relates the bid with the utility level of the household and b^t_{hi} is the hedonic part of the bid expressing the value a household h gives to the attributes (z_i)

of a location i through a set of parameters β. The functional form of (4) implies the assumption of a quasi-linear underpinning utility function which allows to the additive decomposition and simplifies the interpretation of each element Martínez and Henríquez (2007). We assume the preferences of households remain constant in time, therefore the value of the hedonic part of a bid for a particular pair (b_{hi}) will remain constant in time unless the attributes of the location (z_i^t) change from one period to the next. It is reasonable to expect changes market conditions from one period to the other (population , income levels, available supply, etc.) making the utility term b_h reacts to these changes, therefore having different values in each period.

The adjustment of b_h follows the logic of households increasing or decreasing their bids depending on the conditions of the auction (or, in more general terms, the market). In each auction, if there is a demand surplus, households will try to outbid other households until reaching an expected average outcome of winning auctions that allows to locate "somewhere" (although it does not ensure their location). Similarly, in the presence of supply surplus, households will reduce the level of their bids because they can reach an expected number of winning auction that allows to locate somewhere with smaller bids.

In each period, the knowledge of the state of the market comes from the observed rents from previous periods (r_i^t). We assume that households also observe the available supply (S^t) and know the number of households looking for a location in each period (H^t). However, we assume they don't observe the bids of other households (therefore our system represents a sealed-bid auction). Considering this information, each household estimates the value of b_h^t required to make the expected number of winning auctions equal to one.

$$\sum_i P_{h/i}^t = \sum_{i \in S^t} \frac{\exp\left(\mu(b_h^t + b_{hi}(z_i^t))\right)}{\sum\limits_{g \in H^t} \exp(\mu B_{gi}^{t-1})} = 1 \tag{13}$$

Since households can't observe the bids of other households in t we assume they observe the bids in the previous period $(t-1)$. This is equivalent as observing the rents in the previous period since, following (6), the denominator of (13) can also be expressed as:

$$\sum_{g \in H} \exp(\mu B_{gi}^{t-1}) = \exp(\mu r_i^{t-1}) \tag{14}$$

The previous expression implies the assumption of myopic households, that, being unable to forecast the future equilibrium rents, use the available historic information of past rents as a proxy.

Clearing b_h^t from (13) and assuming that only rents from the previous period can be observed, we obtain:

$$b_h^t = -\ln\left(\sum_{i \in S^t} \exp\mu\left(b_{hi}^t(z_i^t) - r_i^{t-1}\right)\right) \tag{15}$$

The adjustment of (15) is similar to the one proposed by Martínez and Donoso (to appear) with the difference of considering that only the households looking for a dwelling and the available units have an effect in the bid level correction. After the adjustment of $b_h$is calculated it's possible to calculate the location probabilities and rents in t.

4.1 Allocation process

The location of households is determined through a Monte Carlo simulation, following probability distributions that will depend on the general conditions of the market, regarding demand or supply surplus. The number of located households or used dwellings may differ from the total number of active households or locations in the market. We denote the set of located households in a period as $\hat{H}^t \subseteq H^t$ and the set of used locations in the same period as $\hat{S}^t \subseteq S^t$.

In a period with demand surplus it is impossible to allocate all households because of the insufficient supply. As explained before, households will increase the level of the bid as a reaction to this. However, some of the households will be outbid in every auction and remain un-located. The market conditions make more appropriate to use the bid probabilities ($P_{h/i}$) to simulate the allocation of households to dwellings. It makes sense to do this location wise, drawing households following (5), as if each location was selecting the best bidder from the pool of remaining households. Under this assumption, the expected number of winning auctions for a particular household h in period t is given by $\sum\limits_{i \in S^t} P^t_{h/i}$.

In the opposite case, in a period with supply surplus, not all the dwelling will be used. Therefore a choice probability ($P_{i/h}$) seems more appropriate to simulate the allocation of dwellings to households, drawing a dwelling following (11) for each household. In this case, the expected number of households choosing a particular dwelling i is given by $\sum\limits_{i \in H^t} P^t_{i/h}$.

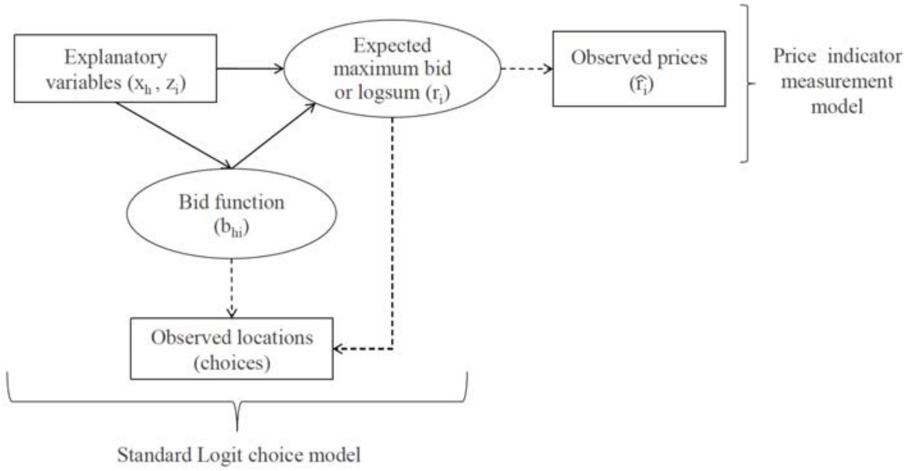
In any market, the transactions are usually bounded by structural characteristics of the involved agents. In the case of the real estate market the constrains are given by the maximum feasible bid for each household (usually determined by the income level) and the reservation price (or minimum feasible rent) of each location. For simplicity, these constraints are ignored in the current formulation of the model, meaning that prices can go has high or low as required by the adjustment of (21). This means that, in the case of demand surplus, all dwellings will be used ($\hat{S}^t = S^t$) while only a fraction of the total households will be located ($\hat{H}^t \subset H^t$). Similarly, in the case of a supply surplus scenario, all households are expected to be located ($\hat{H}^t = H^t$) while only a fraction of the dwellings will be used ($\hat{S}^t \subset S^t$).

Introducing a constrained behavior in the bidding/selling process requires to define thresholds which trigger the exclusion of a household or a dwelling from the transaction. This would allow the existence of (more realistic) scenarios where some households are not located while, at the same time, some dwellings are not occupied. The inclusion of the thresholds should generate a non-compensatory location probability, which can be modeled using models like the Constrained Multinomial Logit (Martínez et al., 2009). An example of the use of non-compensatory probabilities for location choice, but in the context of equilibrium models, can be found in Martínez and Hurtubia (2006)

5 Estimation

Implementation of the proposed model requires to estimate the parameters of the bid function for a base period. The estimation should maximize the likelihood of the observed location pattern but, at the same time, it should ensure that the expected maximum bid (6) of each location is proportional to the real observed prices. For this we propose a model formulation based on the latent variable approach for discrete choice (Walker and Ben-Akiva, 2002; Walker and Li, 2007), where the price is not directly affecting the willingness to pay, but is related as an indicator of the expected maximum bid (or logsum) through a measurement equation. Figure 1 shows the structure of the proposed model. Boxes represent observable data like the attributes of households and locations, transaction prices and observed locations. Circles represent unobservable variables (or latent constructs) like the willingness to pay (bid) and the expected maximum bid. The dashed lines represent measurement relationships and the continuous lines describe structural relationships. The main difference between a traditional Logit model and the proposed model lies

Figure 1: Model structure



in the explicit inclusion of the logsum as a latent variable, that has a measurement relationship with both the choice and an additional indicator (the price).

The Bid function is related to the attributes through the structural equation that defines its functional form: $b_{hi} = f(x_h, z_i, \beta)$. Simultaneously, the measurement relationship between the Bid and the observed location is defined by the choice probability (5). It is worth noticing that, in a traditional Logit formulation, the logsum also intervenes in the measurement as the denominator of the probability.

As described before, the expected maximum bid is related to the observed location through the choice probability (5) and its structural relation with the observed attributes is given by the logsum expression of (6). A new measurement relationship is considered in this formulation, assuming there is a linear relation between the expected maximum bid (r_i) and the observed prices ($\hat{r}_i$), expressed as the following equation:

$$\hat{r}_i = a + \gamma r_i \tag{16}$$

Assuming a normal distribution, a probability density function $f(\hat{r}_i|r_i)$ with mean zero can be defined for the measurement relation of (16) as follows:

$$f(\hat{r}_i|r_i) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{\hat{r}_i - a - \gamma r_i}{2\sigma^2}} \tag{17}$$

The estimation of the proposed model can be done through traditional maximum likelihood but, in this case, the likelihood function is the product of the choice probability and the density function for the price for all observations:

$$L = \prod_{i \in S} \left(\prod_{h \in C_i} \left(P_{h/i} \cdot f(\hat{r}_i|r_i) \right)^{y_{hi}} \right) \tag{18}$$

where $y_{hi} = 1$ if household h is the best bidder for location i and zero otherwise. In the context of the previous equation, S represents the set of available observations for estimation and C_i is the "choice set" for location i, understood as the set of households that participate in the auction for i. If no choice set generation model is available, it is reasonable to assume that all households participate in all auctions, therefore making $C_i = H$ for all i.

The outcome of the maximization of (18) will be the set of parameters (β) from the hedonic part of the Bid function (b_{hi}) and the a, γ and σ parameters of the density function for the price. However, in application, only the choice probability determines the best bidding household, therefore making the location probabilities independent of the price parameters. The measurement equation (16) can be used to estimate the prices from the logsum values (6).

5.1 Estimation results for a real case study

The model is estimated for the residential market of the city of Brussels. The study area considers an extended metropolitan region, including 151 communes that contain a total of 4945 zones (i). Dwelling alternatives are classified in 4 types (ν) adding to a total of 1274701 residential units or location alternatives (νi). The area of study contains a total of 1267998 households, therefore having an aggregated vacancy rate (supply surplus) of 0.5%. The estimation is done over a sample of 1007 observations of located households considering the following bid function:

$$b_{h\nu i} = \beta_{surf} \cdot surf_{\nu i} \cdot \ln(N_h) + \beta_{sup} \cdot Q_i^{sup} \cdot N_h^{sup} + \beta_{house} \cdot \lambda_{\nu i}^{house} \cdot N_h +$$

$$\beta_{trans} \cdot Y_i^{trans} \cdot \gamma_h^{cars=0} + \beta_{trans2} \cdot Y_i^{trans} \cdot \gamma_h^{cars>1} + \beta_{comm} \cdot Y_i^{comm} \cdot \ln(N_h) +$$

$$\beta_{off} \cdot Y_i^{off} \cdot W_h + \beta_{green} \cdot Y_i^{green} \cdot W_h \tag{19}$$

where:

- $surf_{\nu i}$ is the average surface of a residential unit in buildings type ν in zone i (calculated from the census). The building types consider 3 types of house (fully-detached, semi-detached and attached) and apartments.
- N_h is the size (number of individuals) of a household.
- W_h is number of active individuals (workers) in a household
- N_h^{sup} is number of persons in the household who achieved a university degree as their maximum education level.
- Q_i^{sup} is percentage of the population in zone i with a superior level education-degree.
- Y_i^{trans} is a measurement of the quality of public transport (accessibility)
- Y_i^{comm}, Y_i^{off}, Y_i^{green} are measurement of the presence of commerce, offices and public green areas respectively

The measurement equation for prices is defined following (16) and using the explicit definition of the maximum expected bid given by (6):

$$\hat{r}_i = a + \gamma \cdot \ln \sum_h \exp(b_{h\nu i}) \tag{20}$$

For the estimation process, the scale parameter μ is assumed to be one and the adjustment term (b_h) of the bid function is assumed to be zero for all h. For the case study, the observed prices are only available as average transaction prices (in € 100'000) at the commune level for 2 aggregated types of residential units (houses and apartments). Despite this, the available observed prices still provide useful information for the estimation process.

Table 1: Estimation results for Brussels

	Standard Logit			Logit with price indicator		
Parameter	**Value**	**Std err**	**t-test**	**Value**	**Std err**	**t-test**
β_{surf}	0.00698	0.00256	2.73	0.000225	0.000162	1.39*
β_{sup}	0.522	0.104	5.04	0.659	0.0721	9.14
β_{trans0}	0.317	0.135	2.35	0.637	0.0744	8.56
β_{trans2}	-0.438	0.151	-2.9	-0.428	0.0854	-5.02
β_{house}	0.439	0.0599	7.32	0.0459	0.00599	7.67
β_{comm}	-1.32	0.273	-4.82	-0.0118	0.0235	-0.5*
β_{green}	-0.336	0.0718	-4.68	0.152	0.0188	8.09
β_{off}	-0.16	0.0885	-1.81*	0.0738	0.0331	2.23
α	-	-	-	-32.3	4.24	-7.61
γ	-	-	-	2.3	0.301	7.66
σ	-	-	-	-2.12	0.0223	-94.88
Final Log-Likelihood	-7011.03			-6387.76 (7091.13**)		
Likelihood ratio-test	232.44			1478.97 (72.23**)		

*parameters not significant at the 95% level

** log-likelihood considering only the choice probabilities

The model was first estimated for a standard Logit specification and, once good estimates were obtained, it was re-estimated including the measurement equation (indicator) for the observed prices and using the likelihood function of (18). The estimation in both cases was done using an extended version of the software package BIOGEME (Bierlaire, 2003; Bierlaire and Fetiarison, 2009); results are shown in Table 1.

For the standard Logit model all parameters are significant with a 95% confidence (with the exception of β_{off} which is significant with a 90% confidence). The signs of the parameters show that the willingness to pay increases with the surface of the dwelling and the size of the household, and that households with members having university degrees prefer to locate in neighborhoods with a high presence of people with a university degree (this, we assume, is correlated with the income level). Households without a car give a positive value to the presence of public transport facilities while households with more than one car actually prefer to locate in regions with low accessibility for public transport. An interesting result is the effect of the presence of commerce, public green areas and office space, with a negative parameter for all of them and decreasing with the size of the household or the number of workers, depending on the case. These negative estimates were originally interpreted as households preferring to locate in peripheral areas of the city, where the density of commerce, public areas and offices is lower. However, this eventually turned out to be an endogeneity problem (Guevara and Ben-Akiva, 2006) as it will be shown next.

When adding the price indicator to the estimation process some of the parameters become insignificant and some change their sign. For example the relevance of the surface of the dwelling is smaller and its parameter is significant only with an 84% confidence. Other estimates like β_{green} and β_{off}, that were originally negative, came out positive in the estimation with the price indicator. The change in the values of the estimates can be explained as an endogeneity effect in the Standard Logit formulation that happens due to the lack of information. The data for estimation shows that bigger households prefer to locate in the outskirts of the urban area, but this is also explained by the lower prices for bigger dwellings in these regions. Therefore, by not accounting for the price, it appears as if households would bid less for places with access to green areas and services (correlated with the presence o office space). When the price indicator is considered, the estimation generates positive parameters for green areas and offices because these attributes are likely to increase the average price in a neighborhood

Figure 2: Number of people by commune

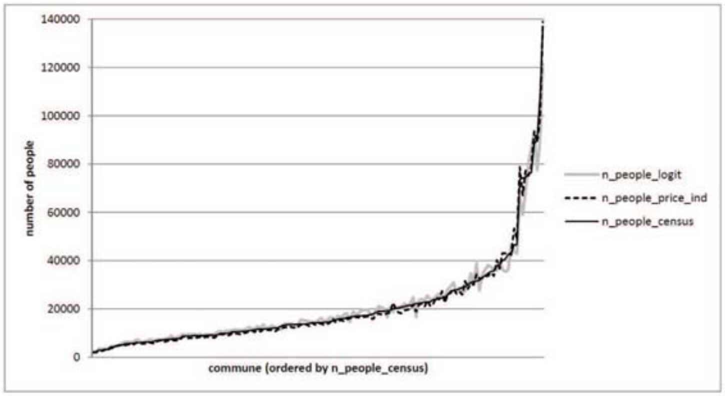
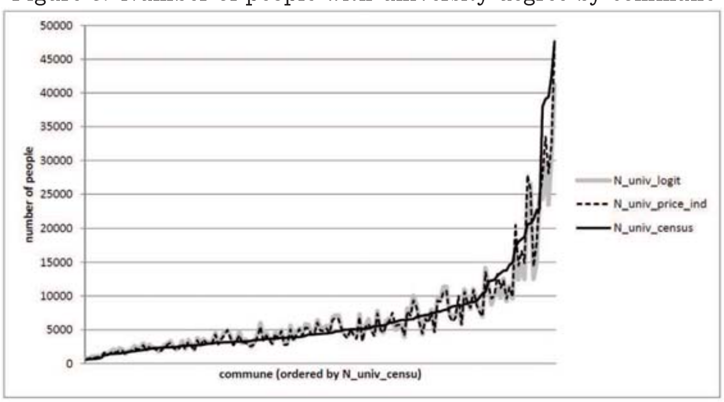


Figure 3: Number of people with university degree by commune



It is not straightforward to evaluate and compare the quality of each model; the different expressions for the likelihood functions make the direct comparison of final log-likelihoods unfair. The ratio test for the "choice" log-likelihood (calculated as the logarithm of sum of the probabilities of the chosen alternatives) is a valid indicator because it considers the same specification for the bid function in both models. This statistic suggests that the Standard Logit performs better than the model with the price indicator. However, this is only valid for the data used in estimation and an expected result because the standard Logit models attempts to fit only to this data set, while the model with a price indicator attempts to fit simultaneously a different set of observations.

A valid comparison is to simulate the location distribution for all the locations in the city with each model, and compare the results with observed statistics. This analysis is performed for two variables: number of individuals in the household and number of individual with university degree. Results are shown in Figures 2 and 3.

Figure 2 shows the results for total number of people, aggregated by commune, obtained with each of the models and compares them with the official statistics coming from the 2001 Belgium National Census. Data is ordered increasingly with the census values. Both models provide reasonable results, with only small deviations from the true values. However, when calculating the fit of each results, the model with price indicator shows a slightly better result with a R^2 of 0.994 (against the census data) versus a R^2 of 0.979 for the standard Logit model.

Figure 3 shows the number of people with university degree by commune. In this case the standard Logit model clearly underestimates the number of persons with university degree for large communes, while the model with price indicator generates estimates which are closer to the real

Figure 4: Logsum vs Prices, Standard Logit

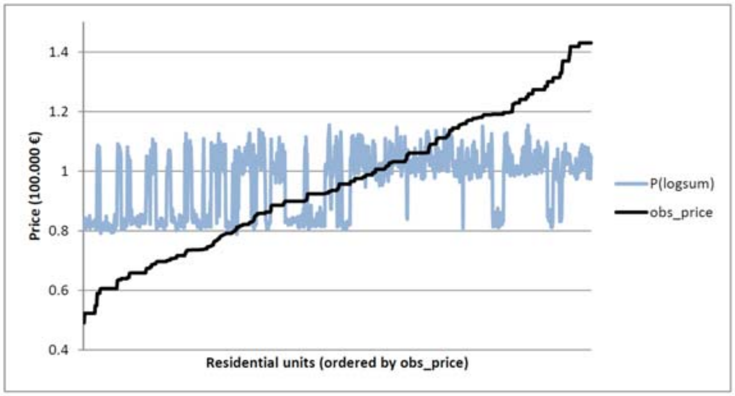
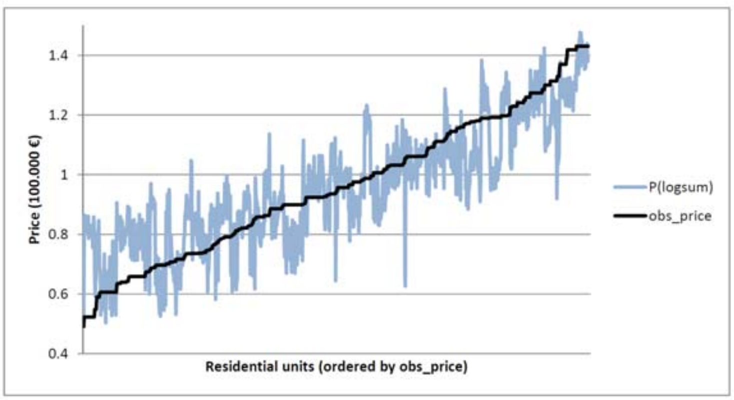


Figure 5: Logsum vs Prices, Model with price indicator



values. The R^2 for the standard Logit in this case is 0.895, while the model with price indicator has a better fit with R^2 =0.925.

A third variable worth analyzing is the predicted price by residential unit. Figure 4 shows the best possible fit between the logsum, obtained with the standard Logit model, for each of the residential units in the area of study. It is possible to see that the logsum does not follow at all the trend of the observed prices.

Figure 5 shows the price estimation results for the choice model estimated with price indicators. The estimated prices (P(logsum)) are calculated following equation (20) and considering the values of Table 1 for the α and γ parameters. The predicted prices, calculated as a function of the resulting logsum, follow the trend of the observed average prices. The noise in the prediction can be explained due to the heterogeneity between the different residential units. While the average price is calculated by house or apartment by commune, the forecast price is calculated for 3 types of houses and apartments and at the zone level, where additional heterogeneity is observed among the dwelling and the zones attributes. A more disaggregated price indicator should allow for a better fit between the logsum and the prices.

These results indicate that the estimation of the choice model including a price indicator allows to better forecast of both the location distribution of agents and the prices.

Table 2: Location in base period

zone	poor hh (P)	rich hh (R)	total supply	rent (r_i)
1 ($z_1 = 0.5$)	281	219	500	1.25
2 ($z_1 = 1.0$)	219	281	500	2.00
total demand	500	500	1000	

5.2 Bid-adjustment component

In the base year, both the utility level (b_h^0) and the hedonic part (b_{hi}^0) of the bid function should be estimated as if a full equilibrium was taking place between the observed households and dwellings. The process then requires to first estimate the parameters of the hedonic part through maximum log likelihood (as described in the previous section) and to adjust the value of b_h^0 following:

$$b_h^0 = -\ln\left(\sum_{i \in S^0} \exp \mu\left(b_{hi}^0(z_i) - r_i^0\right)\right) \tag{21}$$

The solution of the previous equation implies a fixed point problem because the rents (r_i^0) depend on b_h^0, as defined by equation (6). However, little variation is expected if the estimation of the hedonic component of the bid function is properly estimated. Therefore, for the first simulation period, the value of the adjustment components (b_h) should be relatively low (and probably irrelevant) with respect to the value of the hedonic component (b_{hi}). Despite this, it is important to consider the total number of households and available dwellings at the base period in order to introduce the effect of the structural vacancy rate in the rent during the estimation process described in the previous section.

Once the equilibrium bids and rents have been obtained, they are used as the input for the simulation of the first period.

6 Simplified experiment

A simple experiment is conducted to test the properties of the proposed models regarding proper reaction to changes in the market conditions. For this a very simple synthetic city is built considering only two possible zones for location and only two types of households. The zones have only one attribute that characterize them, having a zone with a low value ($z_1 = 0.5$) and a zone with a high value ($z_2 = 1.0$) of the attribute. Households show either a high marginal willingness to pay for the attribute (rich households, with $\beta_R = 2$) or a low willingness to pay (poor households, with $\beta_P = 1$). For simplicity, and in order to allow a better analysis of the reaction to general market conditions of the model, the attributes of the zones remain constant in time. This can be interpreted as not accounting for location externalities in the model.

In the base period the city is perfectly equilibrated, with 500 dwellings in each zone and 500 households of each type. Table 2 shows the location and rents after the equilibrium.

The simulations are done for 20 periods after the base one. Two different scenarios are simulated: one with a supply surplus and one showing demand surplus.

6.1 Supply surplus scenario

In the supply surplus scenario real estate developers predict accurately the total future demand for every period but the first one, where an (arbitrary) overproduction of dwellings take place.

Figure 6: Rents for the supply surplus scenario

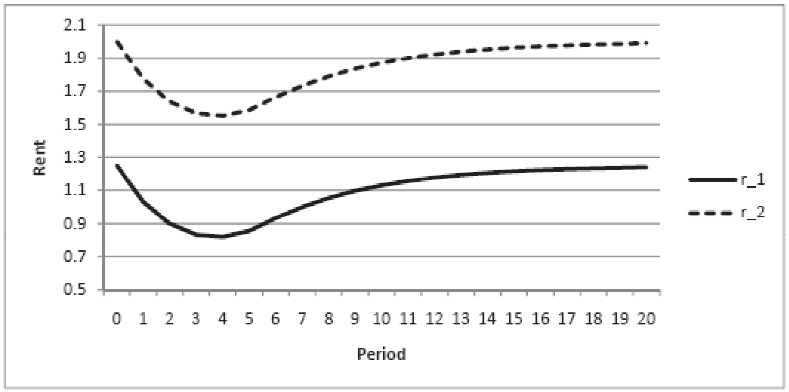
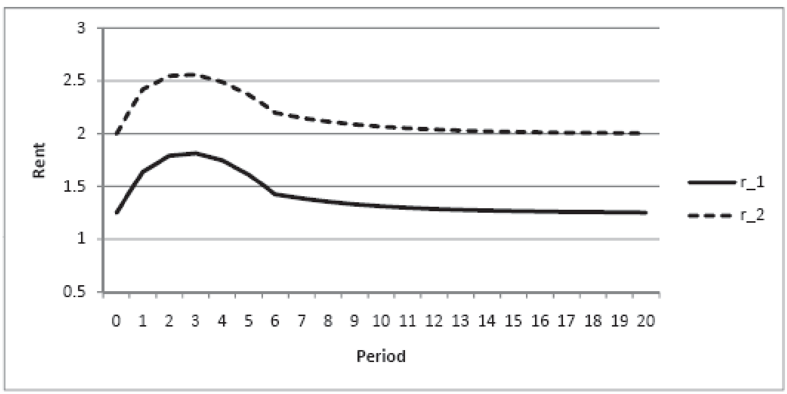


Figure 7: Rents for the demand surplus scenario



Demand grows at a constant rate while supply slowly adjusts to match it. Figure 6 show the resulting rents when applying the proposed model. As expected, rents are higher for dwellings in the zone with higher values for the attributes. In the first period, the excess of supply triggers a reduction in the rents that continues for several periods until supply matches demand again (around period 5). After this point, and given the equality between supply and demand, rents increase until they reach the original (equilibrium) levels.

6.2 Demand surplus scenario

The demand surplus scenario is generated by producing a shock in the growth for rich households in the first period. Supply is unable to react immediately to this and does so in a slow manner.

Figure 7 shows the rents for this scenario. The excess of demand generates an increase in the rent which decreases slowly as supply approaches the levels required to satisfy demand. After several periods rents return to the original equilibrium levels

7 Conclusions

The proposed model is able to account for the auctioning process that takes place in each period of a simulation. The advantage of the model lies in the fact that is able to account for changes in the general conditions of the market, like a growth or a reduction of the ratio between demand and available supply. The method is based on a bid approach for location choice modeling. However it simulates the location process of individuals as the outcome of a bid (dwellings selecting the best bidder/household) only when a demand surplus situation is observed. In the case of a supply surplus scenario, the model simulates the location as a choice (households selecting the location that maximizes their utility).

A estimation method that accounts for observed prices was also proposed. Results show that including a measurement relationship between the logsums and the observed prices in the log-likelihood maximization process allows to obtain better estimates of the bid function parameters. The proposed model is able to better forecast the location choice distribution of agents in the city while, simultaneously, generates reasonable forecasts of the prices as a function of the expected maximum bid of the auction process. The differences observed between forecasted and observed prices can be explained by the aggregated nature of the price indicator. A more disaggregated indicator should allow for a better estimation and, consequently, a better fit.

Future work will consist in the implementation of a simulation accounting for location externalities and increasing the heterogeneity in both supply and demand agents. Application to real data and validation will be done in the context of the SustainCity project (www.sustaincity.org), specifically to the city of Brussels.

References

Alonso, W. (1964). *Location and Land Use: Toward a General Theory of Land Rent*, Harvard University Press, Cambridge, Massachusetts.

Anas, A. (1982). *Residential location markets and urban transportation: Economic theory, econometrics, and policy analysis with discrete choice models*, Academic Press, London.

Bierlaire, M. (2003). Biogeme: a free package for the estimation of discrete choice models, *Proceedings of the Swiss Transport Research Conference*, Ascona, Switzerland.

Bierlaire, M. and Fetiarison, M. (2009). Estimation of discrete choice models: extending biogeme, *Proceedings of the 9th Swiss Transport Research Conference*, Ascona, Switzerland.

Guevara, C. A. and Ben-Akiva, M. (2006). Endogeneity in Residential Location Choice Models, *Transportation Research Record: Journal of the Transportation Research Board* **1977**: 60–66.

Hurtubia, R., Flötteröd, G., Martínez, F. and Bierlaire, M. (2010). Comparative analysis of hedonic rents and maximum bids in a land-use simulation context, *Swiss Transport Research Conference 2010*.

Martínez, F. (1996). Mussa: Land use model for santiago city, *Transportation Research Record: Journal of the Transportation Research Board* **1552**(1): 126–134.

Martínez, F., Aguila, F. and Hurtubia, R. (2009). The constrained multinomial logit: A semi-compensatory choice model, *Transportation Research Part B* **43**(3): 365 – 377.

Martínez, F. and Donoso, P. (to appear). A quasi-equilibrium logit model for residential suburbanization, *Computers in Urban Planning and Urban Management 2011 Conference*.

Martínez, F. and Henríquez, R. (2007). A random bidding and supply land use equilibrium model, *Transportation Research Part B: Methodological* **41**(6): 632 – 651.

Martínez, F. and Hurtubia, R. (2006). Dynamic model for the simulation of equilibrium states in the land use market, *Networks and spatial economics* **6**: 55 – 73.

Martinez, F. J. (1992). The bid-choice land use model: an integrated economic framework, *Environment and Planning A* **24**: 871–885.

Martinez, F. J. (2000). Towards a land use and transport interaction framework, *in* D. Hensher and K. Button (eds), *Handbooks in Transport - Hanbook I: Transport Modelling*, Elsevier Science, pp. 145–164.

McFadden, D. (1978). *Modeling the Choice of Residential Location*, In: Spatial Interaction Theory and Residential Location, edited by A. Karlquist et al., pp. 75-96, North-Holland, Amsterdam.

Walker, J. and Ben-Akiva, M. (2002). Generalized random utility model, *Mathematical Social Sciences* **43**(3): 303 – 343.

Walker, J. L. and Li, J. (2007). Latent Lifestyle Preferences and Household Location Decisions, *Journal of Geographical Sytems* **9**: 77–101.