Netw Spat Econ (2014) 14:47–65
DOI 10.1007/s11067-013-9200-z

Estimation of Bid Functions for Location Choice and Price Modeling with a Latent Variable Approach

Ricardo Hurtubia · Michel Bierlaire

Published online: 2 August 2013

Abstract A new approach for the estimation of bid-rent functions for residential location choice is proposed. The method is based on the bid-auction approach and considers that the expected maximum bid of the auction is a latent variable that can be related to observed price indicators through a measurement equation. The method has the advantage of allowing for the estimation of the parameters of the bid function that explain the heterogeneous preferences of households for location while simultaneously adjusting the expected maximum bid to reproduce realistic values. The model is applied and validated for a case study on the city of Brussels. Results show that the proposed model outperforms other methods for bid-rent estimation, both in terms of real estate prices and spatial distribution of agents, especially when detailed data describing the real estate goods and their prices is not available.

Keywords Location choice · Bid function · Auction · Real estate · Rent

1 Introduction

Land use models are an increasingly used tool for forecasting the evolution of cities and evaluating the potential effects of urban interventions such as real estate developments, modifications to the transport system and changes in urban or transport policy. They are of particular relevance for the field of transport modeling, since long-term

R. Hurtubia (✉)
Department of Urbanism, Facultad de Arquitectura y Urbanismo,
Universidad de Chile, Portugal 84, Santiago, Chile
e-mail: rhurtubia@fau.uchile.cl

M. Bierlaire
Transport and Mobility Laboratory, School of Architecture,
Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne,
Lausanne, Switzerland

travel demand is explained in a large amount by the spatial distribution of agents and activities in a region. The distribution and agglomeration of agents (households and firms) describes trip generation and attraction in spatially disaggregated terms and can be used to generate origin-destination trip matrices or, in the case of agent-based travel models, can be used to define the location of the activities involved in the tours of traveling agents. Besides this, the spatial distribution of agents and activities in a city is one of the main sources of a wide variety of externalities such as congestion, pollution or social segregation and, simultaneously, is one of the main factors that affect the value of land and real estate goods. Modeling the location choice of the different agents that interact in a city is, therefore, one of the main objectives of any land use model.

Forecasting location choice requires forecasting prices and vice versa. It is widely accepted (Alonso 1964; Fujita 1989; Fujita et al. 1999; Glaeser 2008) that location choice depends heavily on the prices of locations while, simultaneously, real estate prices are determined in part by the location preferences of agents. This endogenous interdependence makes the modeling of prices and location a particularly complex task.

Location choice and real estate prices have been traditionally modeled under two different main assumptions regarding the way the market operates: the choice approach and the bid-auction approach. Under the choice approach (McFadden 1978; Anas 1982), agents select the location that maximizes their utility under the assumption that they behave as price takers. In most cases, under the choice paradigm, prices or rents are modeled exogenously through a hedonic model (Rosen 1974). The bid-auction approach (Ellickson 1981) assumes that real estate goods are traded in an auction market, where the best bid for a particular location determines both the located agent and the price or rent of the good.

In the field of urban economics, the bid-auction model has been used mostly as an alternative to hedonic models for the estimation of prices and marginal willingness to pay for attributes of real estate goods. The original model proposed by Ellickson (1981) considered an Extreme Value distribution of the willingness to pay that each agent has for a particular location. This generates a Logit model, conditional on the location, that can be estimated via maximum likelihood. The estimation process assumes that every located agent was the best bidder for the location. However, since the under-determined nature of the Logit model does not allow to find absolute estimates of the willingness to pay, Ellickson's model is only able to estimate relative rents and relative willingness to pay for groups of homogeneous agents.

Improving on Ellickson's work, Lerman and Kern (1983) proposed a method that maximizes the likelihood of an agent being the best bidder for his observed location while, simultaneously, maximizing the likelihood of his bid being equal to the observed transaction price. This method solves the original problem of under-determination in Ellickson's approach, generating absolute estimates of rents or prices and the associated willingness to pay for the location attributes. However, implementing Lerman and Kern's approach requires information that, in general, is not easy to collect: the actual price or rent paid for a particular real estate good, together with the attributes of the best bidder and the location or good itself . Moreover, as in the case of Ellickson's approach, the method imposes a simplification of the bid function, aggregating agents into homogeneous groups of bidders and estimating a single, linear in parameters, bid function for each of them.

The simultaneous location choice and price estimation method of Lerman and Kern has been applied, among others, by Gross (1988), Gross et al. (1990), Gin and Sonstelie (1992), McMillen (1997) and Chattopadhyay (1998) to estimate bid-rent functions in several urban case studies. This literature shows that, in general, accounting for the location preferences of consumers (as it is done in the bid-auction approach) generates better results than hedonic price models, thanks to the possibility of estimating the willingness to pay of different groups of agents and, therefore, providing information about consumer behavior. Despite this, the bid-auction approach has not been extensively applied due to a more complex estimation process than standard hedonic models and the already mentioned expensive data requirements. Moreover, the emphasis has been put in estimation of prices and marginal willingness to pay, giving little attention to the location choice distribution and with scarce validation of the resulting model when forecasting prices or locations.

Muto (2006) analyzed location choice results when using Lerman and Kern's method, finding significant and systematic deviations in the results when compared with observed location distributions for the city of Tokyo. This result suggests that, while Lerman and Kern improve over Ellickson's model by estimating absolute rents, it does so at the cost of worse location forecast capabilities.

The bid-auction approach is particularly attractive for location choice modeling since it provides an explicit explanation of the market clearing process that generates the transaction prices (or rents in the case of the rental market) of real estate. This has motivated the development of several land use models that base their location choice process on the bid-auction approach. Examples of this are RURBAN (Miyamoto and Kitazume 1989), MUSSA (Martínez 1996), IRPUD (Wegener 2008) ILUTE (Salvini and Miller 2005) and, to some extent, UrbanSim (Waddell et al. 2003). In these models, the bid-auction approach has been applied with a focus on modeling the spatial distribution of agents (households and firms) in a city, most of the times using Ellickson's approach to find the relative willingness to pay of different households for the attributes of a location. In these models, if done, the adjustment of the bid functions to absolute levels is done in the context of a market clearing process, separated from the original estimation.

Besides the theoretical appealing, the bid-auction approach is attractive for location choice modeling from an econometric point of view, because it does not have the price endogeneity problems usually found when using the choice approach. Endogeneity occurs because the price is highly correlated with unobserved attributes of the location, therefore complicating the estimation of parameters. In the worst case, if descriptive attributes of the location are omitted, price endogeneity may lead to wrong estimates of the price elasticity and proper estimation will require the use of correcting mechanisms like the Control Function method (Guevara and Ben-Akiva 2006). Because the price of the location does not enter the bid function as a variable, the bid-auction approach does not present price endogeneity issues.

The relevance and advantages of the bid-auction approach motivates the search for estimation methods that allow for consistent estimation of both location choice (maximum bid probabilities) and price distributions without the need of individual level price data. At the same time it is interesting to explore the possibility of estimating bid-rent models where the bidding agents don't have to be aggregated in homogeneous groups or regimes and where bid functions are not constrained to be linear in parameters. This paper proposes a method for the estimation of bid functions that maximizes the likelihood of the observed maximum bids while simultaneously adjusting the bid levels to observed average prices or zonal price indicators. The main assumption behind the proposed method is that, as observed many times in practice, real estate goods are traded in auctions that don't take place explicitly. This implies that the outcome of the auction (the expected maximum bid) is a latent construct that can not be observed but is, however, structurally related to the transaction price. This assumption implies that the potential bid of all agents affects the final price of a real estate good, regardless if they are active in the market (currently looking for a location) or not.

The structure of the proposed model is inspired by the Generalized Random Utility Model (Walker and Ben-Akiva 2002) and defines structural and measurement relationships for two latent variables: the bid of each agent and the expected outcome of an unobserved auction (that we call auction price). The bid is, as usual, structurally dependent on characteristics of the decision maker and attributes of the location and it is measured by a non-linear regression over observed location choices. The auction price is structurally defined as the expected maximum bid in the auction and therefore depends on the bids of all agents. It is measured through a linear regression over observed average zonal prices.

The paper is organized as follows: Section 2 describes the bid-auction approach to location choice modeling. Section 3 reviews the literature on estimation of bid-rent function and analyzes the advantages and drawbacks of the different existing methods. Section 4 describes the method proposed in this paper and Section 5 describes a case study where the method is implemented, validated and compared with other methods. Finally, Section 6 concludes the paper and identifies future lines of research.

2 The Bid Approach to Location Choice

Since Alonso (1964), the real estate market has been understood as an auction market, where agents (households and firms) bid their willingness to pay for a particular good (residential unit, land, etc.) 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:

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

$$s.t.\ px + r_i \leq I$$

In this 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. Given the maximum utility level ($\overline{U}$) a consumer can achieve, the indirect utility can be inverted in the price variable:

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

Under the auction market assumption, the price or rent variable (r_i) of Eq. 3 can be understood as the willingness to pay for a particular location (Jara-Díaz and Martínez 1999), therefore the bid function B can be expressed as:

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

The bid, or bid-rent, function can be understood as the maximum rent (or price) a household can pay for a particular dwelling, while enjoying a fixed utility level $\overline{U}$ (Fujita 1989). In Eq. 4 the index h has been included to take into account heterogeneity in preferences within different households. Ellickson (1981) showed that the bid defined by Eq. 4 can also be written directly as a function of the location attributes ($B_{hi}(z_i)$) and proposed to account for the unobserved heterogeneity in preferences across households by adding a random term,

$$\tilde{B_{hi}} = B_h(z_i) + \varepsilon_h = B_{hi} + \varepsilon_h \tag{5}$$

The probability of a residential unit or location i being occupied by h is the probability of that particular household being the best bidder for the location among all the other bidding households:

$$P_{h/i} = Prob\left\{\, B_{hi} + \varepsilon_h > B_{h'i} + \varepsilon_{h'},\ \forall h' \neq h \right\}$$

If the error terms follow an Extreme Value distribution, the best bid probability can be expressed as a Logit model (McFadden 1978):

$$P_{h/i} = \frac{\exp(\mu B_{hi})}{\sum\limits_{g \in H} \exp(\mu B_{gi})} \tag{6}$$

where H is the set of households bidding for location i.

Under the auction market assumption, the price or rent (r_i) of a good will be the maximum bid and it can be expressed as the following expectation:

$$r_i = E\left(\max_{h \in H}(B_{hi})\right) \tag{7}$$

The extreme value distribution assumption allows to express the expected maximum bid for a particular location as the logsum of the bids, in the same way the logsum represents the expected maximum utility in a traditional maximum utility discrete choice problem (Ben-Akiva and Lerman 1985):

$$r_i = \frac{1}{\mu} \ln \left(\sum_{g \in H} \exp(\mu B_{gi}) \right) + C \tag{8}$$

where C is an unknown constant indicating that the absolute value of the bids cannot be measured. This happens because the Logit model is under-identified and, while relative bids are enough to calculate the best bidder probability of Eq. 6, they do not necessarily relate to real prices or rents.

3 Estimation of Bid-Rent Functions

The first work on estimation of bid-rent functions was developed by Ellickson (1981) who introduced stochasticity in the bid function specification and proposed for the first time the conditional probability of a household being the best bidder for a location (see Eq. 6). The original formulation by Ellickson considers a linear in parameters bid function and is estimated via maximization of the following likelihood function:

$$\mathcal{L} = \prod_{i \in S} \left(\prod_{h \in H} \left(P_{h/i} \right)^{y_{hi}} \right) \tag{9}$$

where y_{hi} is a binary indicator that assumes the value of one if household h is observed to be located in dwelling i and zero otherwise. The term $P_{h/i}$ corresponds to the best bidder probability of Eq. 6.

Ellickson's method had as main objective the estimation of the willingness to pay for housing attributes by different agents, as an alternative to the hedonic rent model originally proposed by Rosen (1974). However, Ellickson's method only allows to estimate relative parameters because the scale parameter (μ) cannot be identified and, as depicted in Eq. 8, rent estimates are known only up to an undefined constant.

A method accounting for observed prices in the estimation to adjust the bids level was first proposed by Lerman and Kern (1983), as a direct extension of Ellickson's model. The method is based on estimating the joint probability of a household being the best bidder for a particular location and of that particular bid being equal to the observed transaction price or land rent (R_i). As a probability, this event can be expressed as:

$$P_{h/i} = \text{Prob} \left\{ B_{hi} + \varepsilon_h = R_i \text{ and } B_{hi} + \varepsilon_h > B_{h'i} + \varepsilon_{h'}, \ \forall h' \neq h \right\} \tag{10}$$

Lerman and Kern's approach considers that the land rent has exactly the same value as the maximum (winning) bid. If the error terms are Extreme Value distributed, the probability of Eq. 10 can be written as:

$$P_{h/i} = f(R_i - B_{hi}) \prod_{h' \neq h} F(R_i - B_{h'i}) \tag{11}$$

with the density (f) and cumulative distribution (F) functions given by:

$$f(\varepsilon) = \mu \exp(-\mu\varepsilon) \exp(-\exp(-\mu\varepsilon)) \tag{12}$$

and

$$F(\varepsilon) = \exp(-\exp(-\mu\varepsilon)) \tag{13}$$

Therefore the likelihood function that needs to be maximized in order to estimate the parameters of B_{hi} is:

$$\mathcal{L} = \prod_{i=1}^{S} \left(-\mu \exp(-\mu (R_i - B_{hi})) \prod_{h'=1}^{H} \exp(-\exp(-\mu (R_i - B_{h'i}))) \right)^{y_{hi}} \tag{14}$$

where H is the total number of households participating in the auction and S is the total number of dwellings in the market. The term y_{hi} is a binary indicator that assumes the value of one if household h is observed to be located in dwelling i and zero otherwise. According to Lerman and Kern, the parameters of Eq. 14 can only be consistently estimated if the bid function is linear in parameters and a full set of constants is included in the bid function B_{hi}.

Lerman and Kern's method has been applied to estimate the real estate rents and the different agent's willingness to pay for particular attributes of housing units in several instances. For example, Gross (1988) applied the model on the city of Bogota, Colombia, finding that the bid-auction approach performs better than hedonic models when forecasting rents and marginal willingness to pay. Gross et al. (1990) and Gin and Sonstelie (1992) applied the model to the cities of Philadelphia and Baton Rouge (Louisiana) respectively, finding reasonable rent estimates. Chattopadhyay (1998) applied the model to the city of Chicago, finding that the rent estimates do not differ much from those of a hedonic model, but have the advantage of providing estimates of the willingness to pay for different groups of agents. Muto (2006) expands Lerman and Kern's model by incorporating an instrumental variable in the estimation and estimates the model for the city of Tokyo, obtaining reasonable results for rent forecasting but a significant bias for location choice. In all the applications reported in this literature review, agents are grouped in homogeneous groups, therefore considering h as a type of agent instead of an individual household or firm. The estimation is done over a sample of locations for which detailed information on the attributes and individual transaction price is available.

An alternative way of estimating bid-rent functions can be derived from the two stage estimation procedure originally proposed by Lee (1982) and adapted by Dubin and McFadden (1984) for the particular case of electric appliances and energy consumption. In this method a choice model is estimated in a first stage, obtaining parameters for the endogenous price function that are adjusted to observed prices in a second stage. In the particular case of bid-rent functions, the choice model is the maximum bidder probability described by Eq. 6 and the adjustment of the bid-rent function is done through the estimation of an hedonic price model where, besides the bid function itself, an instrumental variable is used as an explanatory element. The instrumental variable is obtained via regression of the price against attributes of the location that appear to be correlated with the price but not correlated with the error term in the agent's bid function. The two-stage model has been applied to the bid-rent problem and compared to Lerman and Kern's approach by McMillen (1997). Results show significant differences between the estimates of both approaches and suggests that Lerman and Kern's approach generates distorted results when implemented over data with selection bias problems. As in the bid-auction approach, the two-stage approach requires the aggregation of agents into a restricted number of homogeneous agents.

Similar methods for simultaneous estimation of price functions and location of agents can be found in the literature on locational sorting models (Tiebout 1956). Based on the choice approach and generally focused on the estimation of willingness to pay through hedonic prices, these models also relate individual location choices with aggregate outcomes of the location choice of all agents in equilibrium (Epple 1987) and spatial agglomeration phenomena (Bayer and Timmins 2005). A specific example of this line of research in the context of spatial choice is the work of Bayer et al. (2007), where household preferences for school is analyzed under the effect of unobserved (and endogenous) neighborhood characteristics.

4 Latent Variable Approach for Bid-Rent Function Estimation

We propose a new approach for the estimation of the bid-rent function. We assume that real estate goods are traded in auctions, but that these auctions never take place explicitly. This means that the potential bid of all agents is latent and determines the price of the good, but only in relative terms. We call the outcome (or expected maximum bid) of this latent auction the "latent auction price". To adjust the latent auction price to the level of real prices it must be related to price indicators through a measurement relationship. 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), allowing for simultaneous estimation of the parameters of the bid function and of the price model.

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 latent auction price. The dashed lines represent measurement relationships and the continuous lines describe structural relationships.

The proposed model is different from Lerman and Kern's model because it does not impose the bid of the located household to be equal to the observed price but, instead, imposes a linear relation between the latent auction price and a price indicator. An advantage of this approach is the fact that the price indicator (although it would be preferable) does not have to be the actual price of the transaction but, instead, it can be a much simpler and coarse proxy of price, like the zonal average price or rent by type of location.

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 (Eq. 6). The structural relation of the latent auction price with the observed

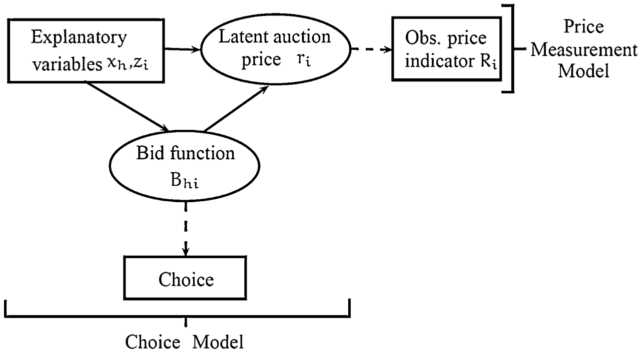

Fig. 1 Model structure

attributes of the location and the agents is given by the expected maximum bid, which is defined by the logsum expression of Eq. 8. A new measurement relationship is considered in this formulation, assuming there is a linear relation between the latent auction price (r_i) and the observed prices (R_i), expressed as the following equation:

$$R_i = a + \gamma r_i + \eta. \tag{15}$$

Assuming a normal distribution for the error term η, a probability density function $f(R_i|r_i)$ with mean zero can be defined for the measurement relation of Eq. 15 as follows:

$$f(R_i|r_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{R_i - a - \gamma r_i}{2\sigma^2}\right) \tag{16}$$

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:

$$\mathcal{L} = \prod_{i \in S} \left(\prod_{h \in C_i} \left(P_{h/i} \cdot f(R_i|r_i) \right) \right)^{y_{hi}} \tag{17}$$

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 set of households that participate in the auction for i. If no 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 Eq. 17 will be the set of parameters (β) for 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 Eq. 15 can be used to estimate the expected prices as a function of the latent auction price.

5 Brussels Case Study

The model is estimated for the residential market of the city of Brussels. Data was collected from three main sources: the 2001 Belgium National Census the 2000 Brussels Land Registry Record and a travel survey to household performed in year 2000 (MOBEL). The study area considers an extended metropolitan region, including 151 communes that contain a total of 4945 zones, denoted by the index i. Dwelling alternatives are classified in 4 types (isolated, semi-isolated and attached houses and apartments), denoted by the index v. Data adds to a total of 1274701 residential units or location alternatives, characterized by their average physical and land use attributes by type of dwelling and zone (vi), which are calculated from the Census and the Land Registry. The area of study contains a total of 1267998 households, characterized by their size, income level, number of active workers, education level and number of vehicles in the household. The whole region has an aggregated vacancy rate (supply surplus) of 0.5 %. The estimation is done over a sample of 1367 observations of located households from the travel survey.

We consider several household characteristics, as well as location and dwelling attributes, that are reported in the literature as relevant to the residential location choice and price formation process. For example the size and income level of the household (Clark and Onaka 1983; Clark et al. 1997), the education level of the members of the household (Gabriel and Rosenthal 1989; Walker and Li 2007), the number of active workers (Horner 2004; Levy 2003; Waddell 1993) and the number of vehicles in the household (Bhat and Guo 2007) were considered as characteristics of the decision makers. The locations are described by traditional land use variables like the presence of different types of economic activities (represented as number of jobs by each type) and average income in a zone (Clark et al. 2006; Lo and Jim 2010), accessibility measures and presence of public transport facilities in a zone (Martinez 1995; Bhat and Guo 2007). We combine these attributes in bid functions with specifications where household characteristics are interacted with location attributes, attempting to capture the heterogeneity in the willingness to pay for location attributes across households. Following this, the specification shown in Table 1 was finally considered for the bid function B_{hvi}, which can be interpreted as the willingness to pay of household h for a dwelling of type v in zone i.

The variable surface$_{vi}$ is the average surface of a residential unit of type v in zone i and it is interacted with the logarithm of the number of individuals in the household in order to account for the bigger (but diminishing) demand for space of large households. The building types consider three types of house (fully-detached, semi-detached and attached) and apartments. A dummy for houses is interacted with a dummy for large households (with more than 2 members), under the hypothesis that bigger households (probably families) will have a tendency to prefer houses over apartments. The percentage of households of high income (more than 3100 Euros per month) in a zone is interacted with a dummy that indicates if household h is of mid-high income in order to capture socioeconomic agglomeration by income level. In a similar but opposite way, the percentage of low income level (earning less than 1860 Euros per month) households in a zone is interacted with a dummy indicating if the bidding household is of high income, in order to capture segregation. The

Table 1 Bid function specification

Parameter	Variables
ASC_I	Alternative specific constant for each income level I of the bidding households
β_{surf}	surface$_{vi}$ (m^2) × log(size$_h$) (number of people)
β_{house}	is_house$_{vi}$ (dummy) × size2$_h$ (dummy)
β_{sup}	high_educ$_i$ (%) × high_educ$_h$ (dummy)
β_{high_inc}	high_inc$_i$ (%) × mid/high_income$_h$ (dummy)
β_{low_inc}	low_inc$_i$ (%) × high_income$_h$ (dummy)
β_{trans0}	PT_accessibility$_i$ (facilities/km^2) × 0_cars$_h$ (dummy)
β_{trans2}	PT_accessibility$_i$ (facilities/km^2) × 2_cars$_h$ (dummy)
β_{access}	car_access$_i$ (generalized travel cost) × 1_cars$_h$ (dummy)
β_{indu}	industry$_i$ (jobs/m^2) × high_income$_h$ (dummy)
β_{office}	office$_i$ (jobs/m^2) × workers$_h$ (dummy)

percentage of people in a zone with university degrees ($texthigh\_educ_i$) is interacted with a dummy indicating that at least one member of the household holds a degree, this is done in order to account for agglomeration by education and as an additional indicator of agglomeration by income. The public transport accessibility variable (PT_accessibility$_i$) was calculated as the density of public transport facilities within a zone and it is interacted with a dummy variable than indicates if the household has no car (to capture the attractiveness of public transport) and with a dummy for two or more cars (to capture the relative disadvantage of living and driving in zones where public transport has priority). The car accessibility measures come from a MATSim model (Rieser et al. 2007) available for the area of study and accounts for the generalized travel cost of visiting all possible destinations in the region by car. They are interacted with a dummy indicating if the household has at least one car. Finally the number of industry jobs in the zone are included to account for the negative externalities this kind of activity generates, while the number of office jobs is included to account for the attractiveness of having services and related job opportunities nearby.

Price data is available as an average by commune (i') and for a simplified classification of dwelling types that aggregates them into houses and apartments (v'). The measurement equation for prices is defined following Eq. 15 and using the explicit definition of the maximum expected bid given by Eq. 8. Instead of price we use the natural logarithm of the price, to capture the diminishing marginal utility of housing attributes (DiPasquale and Wheaton 1996). The resulting expression is similar to a log-log regression for price, a convenient specification due to its good performance for price forecasting when data describing the dwelling is not complete (Cropper et al. 1988).

$$\ln(R_{v'i'}) = a + \gamma \cdot \ln \sum_h \exp(B_{hvi}) \tag{18}$$

Under the bid auction approach, all explanatory variables should be included in the bid, because prices depend directly on them and nothing else. Because of this, no additional explanatory variables are included in the regression. If the set of bids, aggregated in the form of a logsum, is capable to properly explain prices we expect the constant a to be close to zero and γ to be close to one. For the estimation process, the scale parameter μ of the bid probability (Eq. 6) is assumed to be one.

5.1 Estimation Results

The model was first estimated with the method proposed in Section 4, and using the specification of Table 1. For comparison purposes the same model was also estimated with Ellickson's approach. The estimation was performed with an extended version of the software package BIOGEME (Bierlaire 2003; Bierlaire and Fetiarison 2009); results are shown in Table 2, where the first column shows the results using Ellickson's approach while the second column shows the results obtained when using the method proposed in this paper, from now on called "Latent Auction" model.

Table 2 Estimation results for Brussels (Latent Auction)

Parameter	Ellickson			Latent Auction		
	Value	Std err	t-test	Value	Std err	t-test
ASC_1^{***}	0.0	–	–	−3.53	0.301	−11.74
ASC_2	−0.174	0.098	−1.78*	−3.71	0.279	−13.3
ASC_3	−0.858	0.221	−3.89	−4.14	0.273	−15.14
ASC_4	2.29	0.803	2.85	−1.41	0.411	−3.44
ASC_5	2.43	0.818	2.97	−1.37	0.447	−3.06
β_{surf}	0.0031	0.001	3.67	0.002	0.0005	4.0
β_{house}	0.691	0.117	5.89	0.226	0.077	2.94
β_{sup}	1.9	0.169	11.25	0.568	0.093	6.13
β_{high_inc}	4.99	1.99	2.51	5.56	0.793	7.01
β_{low_inc}	−5.49	1.44	−3.81	−4.67	0.796	−5.87
β_{trans0}	0.654	0.335	1.96*	0.312	0.085	3.66
β_{trans2}	−0.615	0.126	−4.9	−0.166	0.088	−1.89*
β_{access}	0.0163	0.0479	0.34*	0.014	0.004	3.25
β_{indu}	0.143	0.816	0.17*	−0.699	0.413	−1.69*
β_{office}	0.207	0.189	1.1*	0.054	0.0256	2.11
a	–	–	–	0.446	2.13	0.21*
γ	–	–	–	1.04	0.146	7.12
σ	–	–	–	−1.89	0.021	−89.37
Final Log-Likelihood		−6080.04**			−6149.91**	

*Parameters not significant at the 95 % level

**Log-likelihood considering only the choice probabilities

***Due to the under-identification of Ellickson's model, the constant for income level 1 is manually fixed to zero

For the Latent Auction model, all parameters show the expected signs. Larger households prefer houses and the willingness to pay for surface increases with the number of people. Households with members having university degrees prefer to locate in neighborhoods with a higher presence of people with a similar education level. Something similar happens with households of mid and high income level, who have a higher willingness to pay for location on zones with high income and assign a negative value to the presence of households of low income . Households without a car give a positive value to the presence of public transport facilities while households with more than one car prefer to locate in regions with high car accessibility and low accessibility for public transport. The presence of industry, as expected, has a negative effect in the price and attractiveness of a location for the Latent Auction model. It is interesting to notice that this variable appears as positive (but extremely non-significant) in the Ellickson estimates. It appears that.the presence of industry, while unable to explain location choice alone, it is capable to explain a decrease in the prices of a location. Therefore, by introducing a regression over prices in the estimation process, it is possible to measure the effect of this variable. In general, thanks to the measurement relationship with both observed location choices and prices, estimates are more significant when using the Latent Auction approach than in the case of Ellickson's method. Despite this, the Latent Auction method has a worse likelihood than Ellickson indicating a worse fit to the estimation data.

The parameters of the price regression also show the expected values, with γ being significant and close to one while a is statistically not different from zero. This indicates that the logsum is capable of explaining the prices without the need of scaling or additional constants. This is possible because, unlike Ellickson's approach, the Latent Auction model is not under-identified and allows for the estimation of a full set of constants in the bid functions.

The same specification of Table 1 is estimated using Lerman and Kern's method, maximizing the likelihood function of Eq. 14. Results for this method are shown in the second column of Table 3 (L&K). Again, for comparison purposes, the estimates obtained with Ellickson's method are shown in the first column.

Parameter estimates obtained with the Lerman and Kern method also have the expected signs. Again, the simultaneous estimation of location and prices allows to increase the significance of the β_{indu} parameter, that appears as positive and non-significant in the results obtained with Ellickson's approach. As in the case of the Latent Auction model, the general significance of the parameters is also improved by the inclusion of the regression over prices. Regarding the likelihood values (accounting for location choice only), L&K's method is dominated by both Ellickson's and the method proposed in this paper, however, it generates relatively good rent estimates, as it is shown next.

5.2 Model Price-Fit Analysis

The fit of the estimated prices is a good indicator of the quality of each model. Figure 2 shows the difference between estimated and observed average prices per commune and dwelling type for the estimation data set. Each column in the box-plot graphic shows results for a different model; the box indicates the value of the

Table 3 Estimation results for Brussels (Lerman and Kern)

Parameter	Ellickson			L&K		
	Value	Std err	t-test	Value	Std err	t-test
μ	1.0	–	–	5.25	0.104	50.73
ASC_1	0.0	–	–	10.4	0.031	335.95
ASC_2	−0.174	0.098	−1.78*	10.3	0.027	377.06
ASC_3	−0.858	0.221	−3.89	9.96	0.037	271.4
ASC_4	2.29	0.803	2.85	10.7	0.143	74.56
ASC_5	2.43	0.818	2.97	10.7	0.146	73.31
β_{surf}	0.0031	0.001	3.67	0.001	0.0002	8.08
β_{house}	0.691	0.117	5.89	0.139	0.023	6.03
β_{sup}	1.9	0.169	11.25	0.376	0.026	14.58
β_{high_inc}	4.99	1.99	2.51	4.29	0.251	17.11
β_{low_inc}	−5.49	1.44	−3.81	−1.18	0.262	−4.51
β_{trans0}	0.654	0.335	1.96*	0.135	0.034	3.91
β_{trans2}	−0.615	0.126	−4.9	−0.113	0.024	−4.81
β_{access}	0.0163	0.0479	0.34*	0.006	0.003	1.78*
β_{indu}	0.143	0.816	0.17*	−0.209	0.164	−1.27*
β_{office}	0.207	0.189	1.1*	0.067	0.025	2.71
Final Log-Likelihood		−6080.04**			−6182.26**	

*parameters not significant at the 95 % level

**log-likelihood considering only the choice probabilities

two quartiles of observations that are closer to the median error, the extremes of the column indicate the value of the biggest positive and negative error. Since both the relative and absolute differences are relevant, both statistics are shown, in the left and right-hand plot respectively.

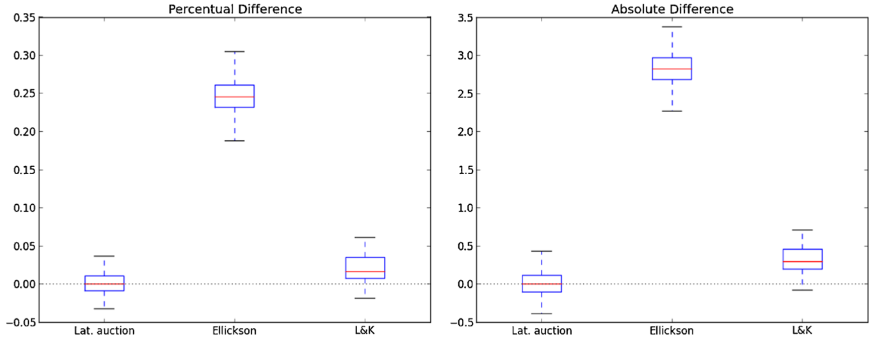

Fig. 2 Estimation fit: natural log of price

Both the Latent Auction and Lerman and Kern methods perform reasonably well. The method proposed in this paper generates estimates that are in 60 % of the cases deviated less than 1 % from the observed prices with a maximum deviation of 4 %. Lerman and Kern also performs well, with 75 % of the estimates deviated less than 4 % and a maximum deviation of 6 %. In both cases, the magnitude of the error is reasonable because the estimated prices were calculated for a wider classification of dwelling types and for a much finer basic spatial unit than those of the observed average prices.

As expected, Ellickson's method does not perform well in this regard, systematically overestimating the prices. However, it seems to be the best model regarding estimation of the spatial distribution of agents. The log-likelihood, calculated as the logarithm of sum of the probabilities of the chosen alternatives, is a valid indicator of model fit since it considers the same specification for the bid function in all models. This statistic suggests that Ellickson fits better than the Latent Auction model and that both models are better than Lerman and Kern's approach. 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 models introducing a price regression attempt to simultaneously fit to an additional set of observations. Because of this, the result analysis so far does not allow to identify which model is performing better in general and further validation is required.

5.3 Validation

Validation is performed by simulating the location distribution for all the locations in the city with each model, and comparing the results with observed statistics. For this, all the real estate supply is generated from the census data and households are assigned following the different maximum bid distributions obtained with each method. The analysis is performed for three variables: prices, number of individuals in the household and number of individual with university degree. Results are shown in Figs. 3, 4 and 5 as the difference at the commune level of the forecast variables against their observed value.

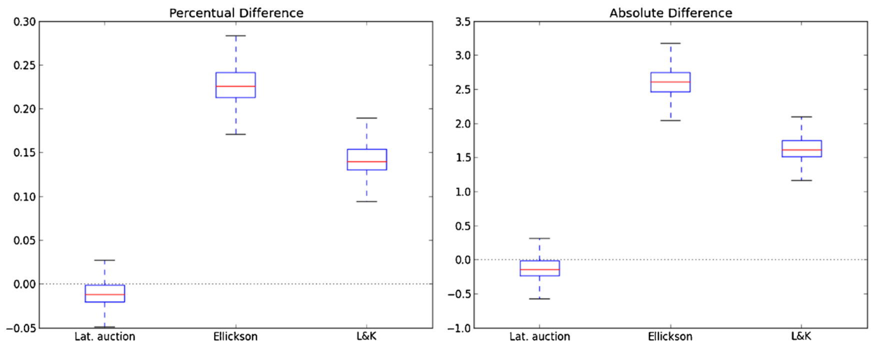

Fig. 3 Error in forecast: natural log of price

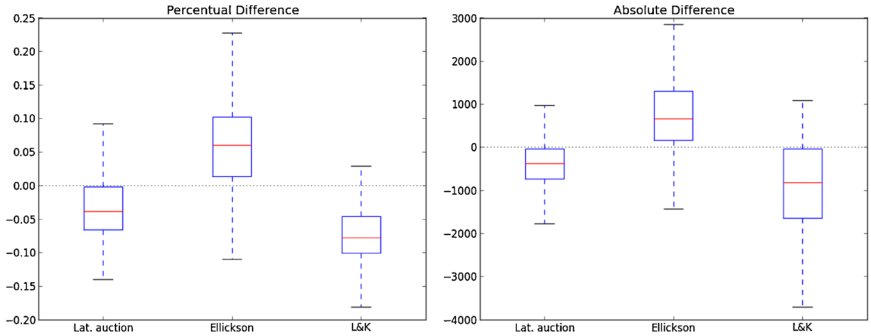

Fig. 4 Error in forecast : number of people by commune

The difference between price forecast and observed average price is shown in Fig. 3. Results show that, when applying the models to a different data set, accounting for the complete market, the Latent Auction approach is superior to Lerman and Kern, where a systematic overestimation occurs. This is probably due to the intensive data requirements of L&K (individual transaction prices), which are not met by the relatively poor nature of the available data.

Figure 4 shows the results for total number of people (the sum of the number of individuals per household), aggregated by commune, against the official population statistics coming from the 2001 Belgium National Census. The Latent Auction model tends to underestimate the population at the commune level, with 50 % of the communes having an absolute deviation smaller than 7 %. Ellickson's model tends to overestimate the population, with a slightly higher deviation while Lerman and Kern's model underestimates this variable.

Figure 5 shows the difference between the forecast of people with university degree by commune against the official statistic from the Census. In this case all models perform similarly well. The Latent Auction and Ellickson models show a tendency to overestimate the variable while Lerman and Kern tends to underestimate. It's worth noticing that, at the absolute level, the Latent Auction model outperforms the forecast of the other models.

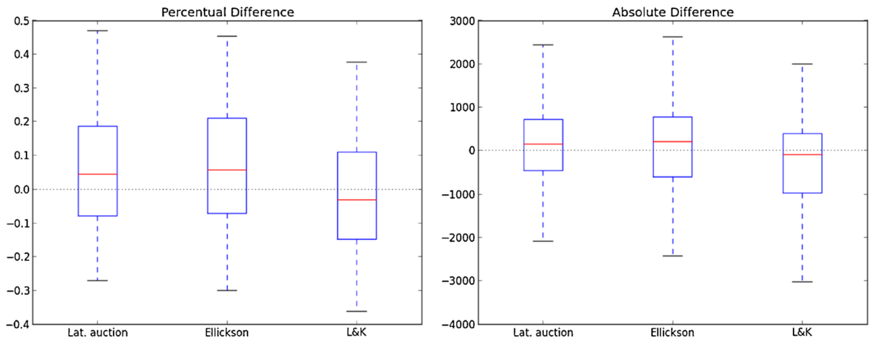

Fig. 5 Error in forecast: number of people with university degree by commune

6 Conclusions

An estimation method for bid-rent functions that accounts for observed locations and coarse price indicators (like zonal averages) is proposed. Results show that including a measurement equation for the expected auction price and the observed average prices in the log-likelihood maximization process allows to obtain better estimates of the bid function parameters. The proposed model is able to forecast, with a reasonable error, the location choice distribution of agents in the city while simultaneously adjusting the expected maximum bid to the average prices. Because of this, the Latent Auction model outperforms Lerman and Kern's model, since the later adjusts well the bid-rent level but deviates significantly from the observed spatial distribution of agents. Moreover, when applied in forecasting, Lerman and Kern's approach is not able to follow the average price trend.

The proposed model has the advantage of not requiring detailed data about real estate goods and prices. In fact, for the case study, only average values were available for both dwelling attributes and prices. This makes the method easier to implement when data is scarce or of aggregated nature.

The differences observed between forecast and observed prices is expected and explained by the aggregated nature of the price indicator. While prices are modeled at the zone level and for four types of dwelling, the indicator is an average value at the commune level for two types of dwellings. A more disaggregated indicator should allow for a better estimation and, consequently, a better fit. Further research should investigate the relevance of choice set formation phenomena (identification of the active bidders in each auction) and the use of more sophisticated (non-linear) structural relationships between the latent auction price and the observed price indicators.

Acknowledgments Research in this article has been funded by the European Commission's Seventh Framework Programme and the Complex Engineering Systems Institute (ICM: P-05-004-F, CONICYT: FBO16). The authors would like to thank the SustainCity team (www.sustaincity.org) for their contribution with data collection and processing.

References

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

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

Bayer P, Ferreira F, McMillan R (2007) A unified framework for measuring preferences for schools and neighborhoods. J Polit Econ 115(4):588–638

Bayer P, Timmins C (2005) On the equilibrium properties of locational sorting models. J Urban Econ 57(3):462–477

Ben-Akiva ME, Lerman SR (1985) Discrete choice analysis: theory and application to travel demand. MIT Press, Cambridge

Bhat CR, Guo JY (2007) A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. Transp Res B Methodol 41(5):506–526

Bierlaire M (2003) Biogeme: a free package for the estimation of discrete choice models. In: Proceedings of the Swiss transport research conference. Ascona, Switzerland

Bierlaire M, Fetiarison M (2009) Estimation of discrete choice models: extending biogeme. In: Proceedings of the 9th swiss transport research conference. Ascona, Switzerland
Chattopadhyay S (1998) An empirical investigation into the performance of Ellickson's random bidding model, with an application to air quality valuation. J Urban Econ 43(2):292–314
Clark WAV, Deurloo MC, Dieleman FM (1997) Entry to home-ownership in Germany: some comparisons with the United States. Urban Stud 34(1):7–19
Clark WAV, Deurloo MC, Dieleman FM (2006) Residential mobility and neighborhood outcomes. Hous Stud 21(3):323–342
Clark WAV, Onaka JL (1983) Life cycle and housing adjustment as explanations of residential mobility. Urban Stud 20:47–57
Cropper ML, Deck LB, McConnell KE (1988) On the choice of funtional form for hedonic price functions. Rev Econ Stat 70(4):668–675
DiPasquale D, Wheaton W (1996) Urban economics and real-estate markets. Prentice Hall, Englewood Cliffs, New Jersey
Dubin JA, McFadden DL (1984) An econometric analysis of residential electric appliance holdings and consumption. Econometrica 52(2):345–362
Ellickson B (1981) An alternative test of the hedonic theory of housing markets. J Urban Econ 9(1):56–79
Epple D (1987) Hedonic prices and implicit markets: estimating demand and supply functions for differentiated products. J Polit Econ 95(1):59–80
Fujita M (1989) Urban economic theory: land use and city size. Cambridge University Press, Cambridge
Fujita M, Krugman P, Venables A (1999) The spatial economy: cities, regions and international trade. MIT Press, Cambridge
Gabriel S, Rosenthal S (1989) Household location and race: estimates of a multinomial logit model. Rev Econ Stat 71(2):240–249
Gin A, Sonstelie J (1992) The streetcar and residential location in nineteenth century philadelphia. J Urban Econ 32(1):92–107
Glaeser EL (2008) Cities, agglomeration, and spatial equilibrium. Oxford University Press, Oxford
Gross DJ (1988) Estimating willingness to pay for housing characteristics: an application of the Ellickson bid-rent model. J Urban Econ 24(1):95–112
Gross DJ, Sirmans C, Benjamin JD (1990) An empirical evaluation of the probabilistic bid-rent model: The case of homogenous households. Reg Sci Urban Econ 20(1):103–110
Guevara CA, Ben-Akiva M (2006) Endogeneity in residential location choice models. Transp Res Rec J Transp Res Board 1977:60–66
Horner MW (2004) Spatial dimensions of urban commuting: a review of major issues and their implications for future geographic research. Prof Geogr 56:170–173
Jara-Díaz SR, Martínez FJ (1999) On the specification of indirect utility and willingness to pay for discrete residential location models. J Reg Sci 39(4):675–688
Lee L-F (1982) Some approaches to the correction of selectivity bias. Rev Econ Stud 49(3):355–372
Lerman SR, Kern CR (1983) Hedonic theory, bid rents, and willingness-to-pay: some extensions of Ellickson's results. J Urban Econ 13(3):358–363
Levy JP (2003) Population patterns and household trajectories in the residential milieu: the example of the Seine-Saint-Denis. Population-E 58(3):323–358
Lo AY, Jim CY (2010) Willingness of residents to pay and motives for conservation of urban green spaces in the compact city of Hong Kong. Urban For Urban Green 9(2):113–120
Martínez F (1996) Mussa: land use model for Santiago city. Transp Res Rec J Transp Res Board 1552(1):126–134
Martinez FJ (1995) Access: the transport - land use economic link. Transp Res B Methodol 29:457–470
McFadden D (1978) Modeling the choice of residential location. In: Karlqvist A (ed) Spatial interaction theory and residential location. North-Holland, Amsterdam, pp 75–96
McMillen DP (1997) Multiple regime bid-rent function estimation. J Urban Econ 41(2):301–319
Miyamoto A, Kitazume K (1989) A land-use model based on random utility/rent-bidding analysis (rurban). Transport policy management and technology - towards 2001 selected proceedings of the fifth world conference on transport research, Yokohama
Muto S (2006) Estimation of the bid rent function with the usage decision model. J Urban Econ 60(1):33–49
Rieser M, Nagel K, Beuck U, Balmer M, RÃŒmenapp J (2007) Agent-oriented coupling of activity-based demand generation with multiagent traffic simulation. Transp Res Rec 2021:10–17

Rosen S (1974) Hedonic prices and implicit markets: product differentiation in pure competition. J Polit Econ 82(1):34–55
Salvini P, Miller EJ (2005) Ilute: an operational prototype of a comprehensive microsimulation model of urban systems. Netw Spatial Econ 5:217–234
Tiebout CM (1956) A pure theory of local expenditures. J Polit Econ 64(5):416–424
Waddell P (1993) Exogenous workplace choice in residential location models: is the assumption valid? Geogr Anal 25:65–82
Waddell P, Borning A, Noth M, Freier N, Becke M, Ulfarsson G (2003) Microsimulation of urban development and location choices: design and implementation of urbansim. Netw Spatial Econ 3(1):43–67
Walker J, Ben-Akiva M (2002) Generalized random utility model. Math Soc Sci 43(3):303–343
Walker JL, Li J (2007) Latent lifestyle preferences and household location decisions. J Geogr Syst 9:77–101
Wegener M (2008) The irpud model: overview, Technical report, Spiekermann & Wegener, Urban Reg Res. http://irpud.raumplanung.unidortmund.de/irpud/pro/mod/mod.htm