Modeling residential location choice and real estate prices with a bid-auction approach

Ricardo Hurtubia
Francisco Martínez
Michel Bierlaire

ERSA
August 31, 2011
Barcelona
1. Motivation
2. Bid approach for location choice
3. Proposed method
4. Brussels case study
5. Discussion
6. Further research
Motivation

- Evolution of land use (location choice) models:
  - Aggregated $\rightarrow$ Disaggregated
  - Equilibrium $\rightarrow$ Dynamic microsimulation

- Market clearing / location distribution:
  - Bid-auction
  - Choice

- Bid approach: endogenous price determination. Usually implemented in equilibrium models (e.g. MUSSA)

- Choice approach: easier to implement in a microsimulation context (e.g. UrbanSim). Requires hedonic rents/prices
Motivation

- Bid-auction approach applied to microsimulation
  - Price formation problem
    - Consistency with observed prices
    - Reaction to market conditions
  - Dynamics (pseudo-equilibrium)
  - Active bidders in the auction (choice set)
Bid approach for location choice

- Assumptions:
  - Real estate goods (locations) are traded in auctions
  - Agents bid their willingness to pay for each location ($B_{hi}$)
  - The best bidder is selected and occupies the location
  - The amount/value of the best bid determines the rent/price
Bid approach for location choice

- Probability of agent $h$ being the best bidder for location $i$:

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

- Expected maximum bid (rent):

$$r_i = \frac{1}{\mu} \ln \left( \sum_{g \in H} \exp(\mu B_{gi}) \right)$$
Bid approach for location choice

- Problems:
  - Requires equilibrium between supply and demand (or at least demand > supply)
  - In the case of supply surplus it not clear which locations are not selected
  - Logsum ($r_i$) doesn’t necessarily reproduce observed prices or rents
Proposed framework

- Bid based location choice model
- Assumptions:
  - Goods (locations) traded in auctions, period-wise
  - Agents bid their willingness to pay for each location
  - Agents adjust the level of their bids as a reaction to market conditions (represented by observed prices)
  - Agents are myopic regarding the outcome of future and present auctions
Proposed framework

- Bid function:

\[ B^t_{hi} = b^t_h + b^t_{hi} \]

Bid adjustment (utility level) \quad Willingness to pay for attributes

\[ b^t_{hi} = f(z^{t-1}_i, x^t_h, \beta) \quad \text{estimated via max log-likelihood, assuming } b^t_h = 0 \]
Proposed framework

- Bid adjustment:
  - Bidding households attempt to ensure winning, on average, at least one auction:

  \[
  \sum_i P_{h/i}^t = \sum_i \frac{\exp(b_h^t + b_{hi}^t)}{\sum_{g \in H} \exp(B_{gi}^t)} = 1 \quad \forall h
  \]

  But… households do not observe bids of other households in the same period. They can only observe transaction prices in previous periods

  \[
  * \mu = 1
  \]

  \(H : \text{full choiceset}\)
Proposed framework

• Bid adjustment:

\[
\sum_{i \in S} P_{h/i}^t = \sum_{i \in S} \frac{\exp(b_h^t + b_{hi}^t)}{\sum_{g \in H} \exp(B_{gi}^{t-1})} = 1
\]

\[
\sum_{g \in H} \exp(B_{gi}^{t-1}) = \exp(r_{i}^{t-1})
\]

\(S\): full choice set of dwellings/locations
Proposed framework

- In each period:
  \[
  b_h^t = -\ln\left(\sum_{i \in S} \exp(b_{hi}^t - r_i^{t-1})\right)
  \]

- In the base year (calibration year):
  \[
  b_h^0 = -\ln\left(\sum_{i \in S} \exp(b_{hi}^0 - r_i^*)\right)
  \]

Observed prices at the base year
Proposed framework

- If the prices are the outcome of an auction, the maximum bid and maximum consumer surplus probabilities generate the same aggregated location distribution*

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

*Bid-choice equivalence (Martínez, 1992)
Price dynamics

- Simulation of a supply surplus scenario with synthetic data

Supply shock (increase)
Brussels case study

Data collected for the SustainCity project:
- Census 2001 (aggregated data by zone)
- Household survey 1999 (disaggregated data, ~1300 obs)
- 1985-2008 average transaction prices by commune and dwelling type

- 1267997 households, 1274701 dwellings
- 157 communes
- 4975 zones
- 4 types of dwelling
  - Detached houses
  - Semi-detached houses
  - Attached houses
Brussels case study

- **Bid function specification:**

  \[
  b_{hv_i} = \beta_{surf} \cdot surf_{vi} \cdot \ln(N_h) + \beta_{sup} \cdot Q_i^{sup} \cdot N_h^{sup} + \beta_{house} \cdot X_{vi}^{house} \cdot N_h + \\
  \beta_{trans} \cdot Y_i^{trans} \cdot \gamma_{cars=0} + \beta_{trans2} \cdot Y_i^{trans} \cdot \gamma_{cars>1} + \beta_{comm} \cdot Y_i^{comm} \cdot \ln(N_h) + \\
  \beta_{off} \cdot Y_i^{off} \cdot W_h + \beta_{green} \cdot Y_i^{green} \cdot W_h + \ln \phi_h
  \]

- \textit{surf}_{vi}: average surface of a residential unit in buildings type \( v \) in zone \( i \) (calculated from the census).
- \( N_h \): number of individuals in a household.
- \( W_h \): number of active individuals (workers) in a household.
- \( N_h^{sup} \): number of persons in the household who achieved a university degree as their maximum education level.
- \( Q_i^{sup} \): percentage of the population in zone \( i \) with a superior level education-degree.
- \( Y_i^{trans} \): measurement of the quality of public transport (accessibility).
- \( Y_i^{comm}, Y_i^{off}, Y_i^{green} \): measurement of the presence of commerce, offices and public green areas.
- \( \phi_{vi} \): correction factor for the household-sampling protocol.
Brussels case study

- Estimation results with PythonBiogeme

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Std_err</th>
<th>t-test</th>
<th>p-value</th>
<th>Robust Std_err</th>
<th>Robust t-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>B_surf_s</td>
<td>0.00832</td>
<td>0.00265</td>
<td>3.14</td>
<td>0.00</td>
<td>0.00274</td>
<td>3.04</td>
<td>0.00</td>
</tr>
<tr>
<td>B_superior</td>
<td>0.484</td>
<td>0.105</td>
<td>4.62</td>
<td>0.00</td>
<td>0.104</td>
<td>4.63</td>
<td>0.00</td>
</tr>
<tr>
<td>B_trans</td>
<td>0.344</td>
<td>0.138</td>
<td>2.50</td>
<td>0.01</td>
<td>0.144</td>
<td>2.39</td>
<td>0.02</td>
</tr>
<tr>
<td>B_trans2</td>
<td>-0.454</td>
<td>0.157</td>
<td>-2.89</td>
<td>0.00</td>
<td>0.159</td>
<td>-2.87</td>
<td>0.00</td>
</tr>
<tr>
<td>B_house</td>
<td>0.419</td>
<td>0.0622</td>
<td>6.74</td>
<td>0.00</td>
<td>0.0638</td>
<td>6.57</td>
<td>0.00</td>
</tr>
<tr>
<td>B_comm</td>
<td>-1.48</td>
<td>0.286</td>
<td>-5.17</td>
<td>0.00</td>
<td>0.293</td>
<td>-5.05</td>
<td>0.00</td>
</tr>
<tr>
<td>B_green</td>
<td>-0.336</td>
<td>0.0736</td>
<td>-4.57</td>
<td>0.00</td>
<td>0.0771</td>
<td>-4.36</td>
<td>0.00</td>
</tr>
<tr>
<td>B_prof</td>
<td>-0.179</td>
<td>0.0906</td>
<td>-1.98</td>
<td>0.05</td>
<td>0.0933</td>
<td>-1.92</td>
<td>0.05</td>
</tr>
</tbody>
</table>

- Likelihood ratio test against null model 219.4
Brussels case study

- Number of people by commune
Brussels case study

- Number of people with university degree by commune
Brussels case study

- Logsums for each location
Brussels case study

- Logsum for each location after adjustment of $b_h$
Discussion

- Framework allows for supply or demand surplus
- Changes in (aggregate) market conditions are captured in the price
- Adjustment of $b_h$ produces maximum expected bids close to observed prices
- Scale of prices
  - Arbitrary? (positive or negative $b_h$)
  - Estimation of $\mu$?
  - Should bid’s be also adjusted location-wise ($b_i$)?
Further research

- Active bidders (choice set generation)
  - Price is affected by who is “competing” for the location
  - Choice set generation or importance sampling?
  - Relevance of the scale of the logsum

- Location assignment
  - Monte Carlo simulation following max bid probabilities?
  - Simultaneous location assignment?
Thanks
Choice approach for location choice

• Assumptions:
  – Each agent selects the location that provides maximum utility
  – Agents are price takers
  – Prices (usually) defined as function of the location attributes
Choice approach for location choice

• Assumption: consumer surplus is a proxy of utility:

\[ V_{hi} = B_{hi} - r_i \]

• Probability of location \( i \) providing maximum utility to agent \( h \):

\[ P_{i/h} = \frac{\exp(\mu(B_{hi} - r_i))}{\sum_j \exp(\mu(B_{hj} - r_j))} \]
Choice approach for location choice

- **Problems:**
  - Price-taker assumption (not good for quasi-unique goods)
  - Market conditions usually not captured by hedonic rents

- **Advantages:**
  - If prices are the outcome of an auction, the location distribution is the same for the bid and choice approaches