Preferences of Higher Educated Households for Location Characteristics and Housing Types

Jan Möhlmann

Based on joint work with
Jasper Dekkers, Mark van Duijn, Or Levkovich, Jan Rouwendal

VU University Amsterdam

Research strategy

- Estimating household preferences based on revealed preferences

- Differentiating between household types

- Using estimating results to predict effects of scenarios and policy
Structure of presentation

- Sorting model
- Data description
- Estimation results
- Scenario analysis
- Conclusions
Sorting model

- Input of the model: current housing supply and current household population

- Households choose a region and a housing type based on regional characteristics and household preferences

- Which household preferences will lead to the current equilibrium?
Sorting model

- Core is a multinomial logit model

- Number of alternatives: 472
  (118 regions x 4 housing types)
Sorting model

- Core is a multinomial logit model

- Number of alternatives: 472 (118 regions x 4 housing types)

- Utility of household $i$ in alternative $n$:
  \[ u_{in} = \alpha_i P_n + \beta_i X_n + \xi_n + \epsilon_{in} \]
Sorting model

- Core is a multinomial logit model

- Number of alternatives: 472
  (118 regions x 4 housing types)

- Utility of household $i$ in alternative $n$:
  \[ u_{in} = \alpha_i P_n + \beta_i X_n + \xi_n + \epsilon_{in} \]

- Probability that household $i$ chooses alternative $n$:
  \[ \pi_{in} = \frac{e^{\hat{u}_{in}}}{\sum e^{\hat{u}_{in}}} \]
Endogeneity problem

- Unobserved characteristics influence utility and household prices
  - Housing prices
  - Accessibility
  - Urban amenities
  - Nature
  - Unobserved characteristics

Utility
Endogeneity problem

- Unobserved characteristics influence utility and household prices

- Sorting model
- Data
- Results
- Scenario analysis
- Conclusions

- Housing prices
  - Accessibility
  - Urban amenities
  - Nature
  - Unobserved characteristics

Utility
Estimation strategy

- Solution: estimation in two steps

\[ u_{in} = \alpha_i P_n + \beta_i X_n + \xi_n + \varepsilon_{in} \]
Estimation strategy

- Solution: estimation in two steps

\[ u_{in} = \alpha_i P_n + \beta_i X_n + \xi_n + \epsilon_{in} \]

\[ \alpha_i = \alpha + \alpha_1 (edu_i - \overline{edu}) \]
\[ \beta_i = \beta + \beta_1 (edu_i - \overline{edu}) \]
Estimation strategy

- Solution: estimation in two steps

\[ u_{in} = \alpha_i P_n + \beta_i X_n + \xi_n + \epsilon_{in} \]

\[ \alpha_i = \alpha + \alpha_1 (edu_i - \bar{edu}) \]

\[ \beta_i = \beta + \beta_1 (edu_i - \bar{edu}) \]

\[ u_{in} = \alpha P_n + \beta X_n + \xi_n + \alpha_1 (edu_i - \bar{edu}) P_n + \beta_1 (edu_i - \bar{edu}) X_n + \epsilon_{in} \]
Estimation strategy

- Solution: estimation in two steps

\[ u_{in} = \alpha_i P_n + \beta_i X_n + \xi_n + \epsilon_{in} \]

\[ \alpha_i = \alpha + \alpha_1 (\text{edu}_i - \bar{\text{edu}}) \quad \beta_i = \beta + \beta_1 (\text{edu}_i - \bar{\text{edu}}) \]

\[ u_{in} = \alpha P_n + \beta X_n + \xi_n + \alpha_1 (\text{edu}_i - \bar{\text{edu}}) P_n + \beta_1 (\text{edu}_i - \bar{\text{edu}}) X_n + \epsilon_{in} \]

- Step 1: estimate \( \alpha_1 \) and \( \beta_1 \) and an alternative specific constant (asc = \( \alpha P_n + \beta X_n + \xi_n \))

- Step 2: explain the asc’s based on characteristics of alternatives using 2SLS
Structure of presentation

- Sorting model
- Data description
- Estimation results
- Scenario analysis
- Conclusions
Data (households)

- Data are obtained from *Woon Onderzoek Nederland* (WoON) 2012

- **57,276 households**

- **Household characteristics**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Couple</td>
<td>0.63</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Children in household</td>
<td>0.35</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Higher education</td>
<td>0.30</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Age</td>
<td>51.7</td>
<td>17</td>
<td>100</td>
</tr>
</tbody>
</table>
Data (regions)

- 118 regions based on 415 adjacent municipalities
Data (regions)

- Every region provides four alternatives (rental houses and three types of owner-occupied houses)

Regional characteristics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance to nearest 100,000 jobs (in km)</td>
<td>12.6</td>
<td>3.6</td>
<td>32.8</td>
</tr>
<tr>
<td>Distance to intercity train station (in km)</td>
<td>7.5</td>
<td>1.5</td>
<td>27.8</td>
</tr>
<tr>
<td>Distance to highway onramp (in km)</td>
<td>4.1</td>
<td>1.0</td>
<td>20.3</td>
</tr>
<tr>
<td>Share of surface is nature (in %)</td>
<td>13.8</td>
<td>0.4</td>
<td>65.8</td>
</tr>
<tr>
<td>Size of historical city centre (in km²)</td>
<td>0.9</td>
<td>0</td>
<td>13.3</td>
</tr>
</tbody>
</table>

- Prices of owner-occupied houses differ by type
Price of a standard house is determined using a hedonic price analysis on transaction data.
Structure of presentation

- Sorting model
- Data description
- Estimation results
- Scenario analysis
- Conclusions
Willingness to pay

- Sorting model
- Data
- Results
- Scenario analysis
- Conclusions

Willingness to pay

Eur

jobs (km)   train station (km)   highway (km)   nature (%)   city centre (km2)

0  1000  2000  3000  4000  5000

Euro

Results - Scenario analysis

- Conclusions
Willingness to pay

- Sorting model
- Data
- Results
- Scenario analysis
- Conclusions

Apartment: reference type
Terraced housing: - 500
Detached housing: 39.000
WTP by household type

- For 1 km higher proximity to nearest 100,000 jobs
WTP by household type

- For detached housing (relative to apartments)
Structure of presentation

- Sorting model
- Data description
- Estimation results
- Scenario analysis
- Conclusions
Scenario analysis

- Estimated parameters for household preferences allow us to sort a given population of households over the alternatives.

- Scenario input:
  - Distribution of household types (e.g. education, age)
  - Regional characteristics (e.g. distance to jobs, nature)
  - Housing supply (distribution between regions and composition of housing types within regions)

- Scenario output:
  - Housing prices
  - Composition of household types for each region
Global economy 2030 scenario

- Example: housing supply in 2030 based on Ruimte Scanner XL
- Global Economy scenario

- Assumption: number of houses is equal to number of households

- Household demographics and regional characteristics remain constant
Global economy 2030 scenario

Difference between GE and BASE (in %) - detached

Price change of detached housing
Global economy 2030 scenario

Change in share of higher educated households
Structure of presentation

- Sorting model
- Data description
- Estimation results
- Scenario analysis
- Conclusions
Conclusions

- Sorting model uses revealed preferences to determine willingness to pay for regional characteristics

- Can distinguish between household types

- We find a positive willingness to pay for proximity to jobs, availability of nature and urban amenities, and for detached housing

- Estimation results can be used to predict the effects of scenarios and policy on housing prices and regional household composition
Alternative models

- Estimating the sorting model with different characteristics of households and regions
  - foreign knowledge workers and students
  - field of education or profession

- Different level of aggregation (e.g. neighbourhoods instead of municipalities)

- Estimating costs of moving (using distance to previous region)
Preferences of Higher Educated Households for Location Characteristics and Housing Types

Jan Möhlmann

Based on joint work with
Jasper Dekkers, Mark van Duijn, Or Levkovich, Jan Rouwendal

VU University Amsterdam