Modelling urban house prices using open data

Shanaka Herath

Open Cities | Open Data workshop
2 October 2015
Outline

- House price modelling and data
- Traditional challenges of sourcing data
- Times are changing – open data initiatives of cities
- Two Australian case studies
- Final remarks
House price modelling and data

- Hedonic house price model

HPM estimates the values of houses, taking into account *implicit* values of utility-bearing (housing) characteristics:

\[ P = f(S, L, N, t) \]

- \( P \) is house prices;
- \( S \) is structural attributes;
- \( L \) is locational variables;
- \( N \) is neighbourhood characteristics; and
- \( t \) is an indicator of time.
House price modelling and data (contd.)

- Basic relationships – econometric framework

- Testing urban amenities theory
  - Whether a nearby bus stop increase house prices?
  - Whether a nearby public park increase house prices?
  - Whether proximity to a good school increase house prices?
  - Whether proximity to a major road increase or decrease house prices?
Traditional challenges of sourcing data

- **Types of data**
  - Structural attributes – part of transaction dataset
  - Location / neighbourhood variables – harder to obtain / compute

- **Custodians of transaction data - government agencies or commercial firms**
  - Longer processes of negotiation / higher costs
  - Confidentiality agreements limiting the use of data for specific projects.

- **Location / neighbourhood data**
  - Typically spatial data available through planning departments / city councils
Times are changing – open data initiatives of cities

- Many cities have recently adopted open data policies – urban data increasingly available to the public
- Though there are limitations in terms of privacy and anonymity – personal data strictly excluded
- These new datasets have created exciting new opportunities to undertake innovative house price analyses
  
  e.g. Vienna (https://open.wien.at/site) and NSW (http://data.nsw.gov.au)
Two Australian case studies

- The first case study\(^1\) evaluates the price premium associated with houses located in close proximity to public transport nodes
  - Study area: Sydney LGA
  - Dataset: house prices provided by RP Data; public transport nodes (i.e., bus stops) in Sydney, provided by LPI (http://www.lpi.nsw.gov.au)

- Bus stop zones (ped shed) 100m, 200m and 400m tested – a price premium if located within 200 metres

Two Australian case studies (contd.)

- In the second case study, price differences associated with houses located close to better schools, main roads, public parks and public transport are assessed
  - Study area: Sydney LGA
  - Dataset: Digital and administration data available through government websites.
    - House prices -> APM through AURIN;
    - Better schools -> NAPLAN scores and rankings (http://www.myschool.edu.au);
    - Main roads -> LPI (http://www.lpi.nsw.gov.au);
    - Public parks -> CoS (http://www.cityofsydney.nsw.gov.au/explore/facilities/parks);
    - Public transport nodes -> LPI
The map of POI data
Final remarks

- **Open data for citizens / communities**
  - Transparency and accountability
  - Facilitates dialogue between citizens and administration, improve participation and communication
  - Promotes democratic exchange of knowledge, innovation
    e.g. Assessing positive and negative economic externalities of amenities and dis-amenities in cities; strengthening the evidence base on the value of public amenities in cities.

- **TDX - opportunities are limitless**
**Acknowledgements:**
ECR Grant from Built Environment at UNSW

Shanaka Herath
*City Futures Research Centre*
Built Environment, UNSW AUSTRALIA
UNSW SYDNEY NSW 2052 AUSTRALIA
T: +61 (2) 9385 4931
E: shanaka.herath@unsw.edu.au
W: cityfutures.net.au