Understanding spatial variations in the impact of accessibility on land value using geographically weighted regression

World Symposium on Transport & Land Use Research
28-30 July 2011

Dr Hongbo Du and Professor Corinne Mulley
Chair in Public Transport | Institute of Transport and Logistics Studies
The University of Sydney
corinne.mulley@sydney.edu.au
http://sydney.edu.au/business/itls
Outline

› The impact of new transport infrastructure on accessibility
› Possible methodologies for measuring impacts
› Geographically Weighted Regression
   - methodology
› The Tyne and Wear case study
› The GWR model: results
› Conclusions

Source: www.pteg.net
Impact of public transport infrastructure on residential land value

› New infrastructure changes accessibility
› Changes value of unimproved land

BUT

› Changes in improved land observed

Source: photos.com
Taking account of improvements to land value

› House prices include
  - Land
  - Improvements to land

› Methodology needs to standardise for improvements to identify unimproved land value

Source: blog.evolutionrealty.com.au
Possible methodologies

› Comparison method
  - Fails to capture complexities which make up house prices

› Hedonic modelling
  - Common approach
  - Approaches to variations across space

› Spatial modelling
Geographically Weighted Regression (GWR)

- Spatial modelling approach
  - ‘global’ model equivalent to Ordinary Least Squares/hedonic modelling
  - ‘local’ model takes account of geographical relationships
- Outputs mapped
In OLS regression, the relationships being modelled are assumed to be constant across the whole study area.

\[ Y_i = \beta_0 + \sum_k \beta_k X_{ik} + \varepsilon_i \]

Geographically Weighted Regression (GWR)

\[ Y_i(u_i, v_i) = \beta_0 (u_i, v_i) + \sum_k \beta_k (u_i, v_i) X_{ik} + \varepsilon_i \]

In GWR, relationships are modelled using the location of the observation (X and Y) in comparison to its neighbours.

The neighbouring observations are given more weight than those further away.
GWR spatial considerations

$W_{ij}$ is the weight of data point $j$ at the regression point $i$.
$d_{ij}$ is the distance between regression point $i$ and data $j$.

Source: Fotheringham et al. (2002) *Geographically Weighted Regression*, pp. 44
The case study area: Tyne and Wear

Source: based on Tyne and Wear Accessibility Modelling and Edina Digimap
Model

- $P_i = f(C, T, N)$
  - $C$ is a vector of characteristics of properties
  - $T$ is a vector of transport accessibility
  - $N$ is a vector of the neighbourhood environment

Data

- House prices
  - Asking prices not transactional prices
  - Collected at six digit postcode level
  - Included internal features
- Transport accessibility
- Socio-economic data

Source: thewordguy.files.wordpress.com/2009/07/data.jpg
\[
\ln P_i = \alpha_0 \\
+ \alpha_1 \text{BEDROOM}_i \\
+ \alpha_2 \text{FLAT}^*\text{BED}_i + \alpha_3 \text{SEMI}^*\text{BED}_i + \alpha_4 \text{DETA}^*\text{BED}_i \\
+ \alpha_5 \text{SP}\_\text{AVE}_i + \alpha_6 \%\text{ETHNM}_i + \alpha_7 \%\text{HPROF}_i + \alpha_8 \%\text{UNEM}_i \\
+ \alpha_9 \text{CAR}\_\text{ACC}_i + \alpha_{10} \text{PT}\_\text{ACC}_i
\]

Where:

FLAT, SEMI (semi-detached), DETA (detached) are types of property

SP\_AVE is the point score for the local primary school

\%ETHNM is the % ethnic minority, \%HPROF is the % higher professionals and \%UNEM is the % unemployed

CAR\_ACC is accessibility by car and PT\_ACC is accessibility by public transport, measured in minutes
Model specification

\[ \ln P_i = \alpha_0 + \alpha_1 \text{BEDROOM}_i + \alpha_2 \text{FLATBED}_i + \alpha_3 \text{SEMIBED}_i + \alpha_4 \text{DETABED}_i + \alpha_5 \text{SP_AVE}_i + \alpha_6 \%\text{ETHNM}_i + \alpha_7 \%\text{HPROF}_i + \alpha_8 \%\text{UNEM}_i + \alpha_9 \text{CAR_ACC}_i + \alpha_{10} \text{PT_ACC}_i \]
\[
\ln P_i = \alpha_0 + \alpha_1 \text{BEDROOM}_i + \alpha_2 \text{FLATBED}_i + \alpha_3 \text{SEMIBED}_i + \alpha_4 \text{DETABED}_i + \alpha_5 \text{SP_AVE}_i + \alpha_6 \%\text{ETHNM}_i + \alpha_7 \%\text{HPROF}_i + \alpha_8 \%\text{UNEM}_i + \alpha_9 \text{CAR_ACC}_i + \alpha_{10} \text{PT_ACC}_i
\]
\[ \ln P_i = \alpha_0 + \alpha_1\text{BEDROOM}_i + \alpha_2\text{FLATBED}_i + \alpha_3\text{SEMIBED}_i + \alpha_4\text{DETABED}_i + \alpha_5\text{SP_AVE}_i + \alpha_6\%\text{ETHNM}_i + \alpha_7\%\text{HPROF}_i + \alpha_8\%\text{UNEM}_i + \alpha_9\text{CAR_ACC}_i + \alpha_{10}\text{PT_ACC}_i \]
# Global Regression results

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>Significance (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>11.965</td>
<td>0.000</td>
</tr>
<tr>
<td>No. of bedrooms</td>
<td>-0.074</td>
<td>0.000</td>
</tr>
<tr>
<td>Flat*bedrooms</td>
<td>0.070</td>
<td>0.000</td>
</tr>
<tr>
<td>Semi-detached*bedrooms</td>
<td>0.082</td>
<td>0.000</td>
</tr>
<tr>
<td>Detached*bedrooms</td>
<td>0.092</td>
<td>0.000</td>
</tr>
<tr>
<td>Average point score of primary school</td>
<td>-0.009</td>
<td>0.085</td>
</tr>
<tr>
<td>% ethnic minority</td>
<td>0.001</td>
<td>0.685</td>
</tr>
<tr>
<td>% higher professional</td>
<td>0.029</td>
<td>0.000</td>
</tr>
<tr>
<td>% unemployment</td>
<td>-0.074</td>
<td>0.000</td>
</tr>
<tr>
<td>Car accessibility (mins)</td>
<td>0.021</td>
<td>0.000</td>
</tr>
<tr>
<td>PT accessibility (mins)</td>
<td>-0.012</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Is it better than the global results?

- AIC measure and Monte Carlo test

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Significance level for spatial variability (p-value)</th>
<th>Significance level in global regression (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>No. of bedrooms</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Flat*bedrooms</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Semi-detached*bedrooms</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Detached*bedrooms</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Average point score of primary school</td>
<td>0.040</td>
<td>0.085</td>
</tr>
<tr>
<td>% ethnic minority</td>
<td>0.100</td>
<td>0.685</td>
</tr>
<tr>
<td>% higher professional</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>% unemployment</td>
<td>0.017</td>
<td>0.000</td>
</tr>
<tr>
<td>Car accessibility (mins)</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>PT accessibility (mins)</td>
<td>0.000</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Local model results: public transport accessibility
Local model results: car accessibility
Conclusions

› Methodological contribution of GWR to explaining impact of accessibility on land value

› Empirical results for Tyne and Wear
  - Distributional impacts

› Policy
  - Implications for land value capture policy as a means of funding new transport infrastructure

Source: free foto.com