Boundaries in Residential Segregation

Guanpeng Dong\textsuperscript{a} & Gwilym Pryce\textsuperscript{b}

\textsuperscript{a} Department of Geography and Planning, University of Liverpool
Email: guanpeng.dong@liverpool.ac.uk

\textsuperscript{b} Sheffield Methods Institute, University of Sheffield
Email: g.pryce@sheffield.ac.uk
Plan of Presentation

1. Introduction
2. Methods
3. Results
4. Conclusion
• Terminology:
  – Socially differentiated borders = “boundaries”
  – Arbitrary borders = “borders”
• Minimum border length hypothesis
  – Steep boundary may imply that no-one wants to live near the frontier ⇒ underlying hostility

• Schelling
  – Segregation can occur when no-one wants it

• Segregation itself can affect behaviour
  – Allport’s contact hypothesis:
    – Prejudice more likely to occur, linger and grow if a lack of interaction between groups.
Not aware of empirical research that specifically looks at boundary effects

Some studies find that:

– “Crime, particularly violent crime, is higher in cities and metropolitan areas where Blacks and Latinos are segregated in different neighborhoods from Whites.” (Krivo et al. 2015)

  • (e.g., Feldmeyer, 2010; Krivo et al., 2009; Peterson and Krivo, 2010a)
• Studies tend to ignore the spatial nature of segregation.
• Proximity to social boundaries may be particularly important
  – E.g. Belfast “peace walls”
Social boundary within arbitrary border

Social boundary overlapping with border

Border, but not social boundary
2. Methods

Identifying Boundaries

- Two Step approach to identifying boundaries:
  - **Step 1:** Identify significant step changes in the spatial distribution of ethnicity (e.g. non-white population).
    - using a locally adaptive spatial conditional autoregressive model (Lee and Mitchell 2013), indicating significant differences in minority proportions on the two sides of a boundary.
  - **Step 2:** Set a threshold to ensure that the boundary is not just statistically significant from zero, but also substantively different from zero.
A Binomial Locally Adaptive Spatial Conditional Autoregressive Model

- The study region (each city) is partitioned into $n$ non-overlapping areal units (e.g. LSOAs or DZs), denoted as $A = \{A_1, \ldots, A_n\}$
- $Y_k$ denote the number of people who are in the ethnicity minority (e.g. non-white) in area $A_k$
- $N_k$ denote the total number of people in $A_k$
- Constructed a Bayesian locally adaptive spatial conditional autoregressive model
  - for a binomial dependent variable.
  - Lee and Mitchel (2013)
The basic model specification

\[ Y_k \sim \text{Binomial}(N_k, p_k), \quad k = 1, \ldots, n \]

\[ \ln(p_k/(1 - p_k)) = \beta_0 + u_k \]

\[ u_k \mid u_{-k}, W, \lambda, \tau^2 \sim N \left( \frac{\lambda \sum_{l=1}^{n} w_{kl} u_l}{1 - \lambda + \lambda \sum_{l=1}^{n} w_{kl}}, \frac{1}{\tau^2 (1 - \lambda + \lambda \sum_{l=1}^{n} w_{kl})} \right) \]

\[ \beta_0 \sim N(0, b); \quad \tau^2 \sim \text{Gamma}(e, f); \quad \logit(\lambda) \sim N(0, 100). \]

- \( u \) is a vector of random effects conceptualised as a conditional autoregressive model (CAR), capturing the spatial correlations in the distribution of \( p_k \) and the potential over-dispersion effect.
- \( W \) is a binary neighbourhood structure of spatial weights matrix with \( w_{jk} = 1 \) if units \( A_j \) and \( A_k \) share a common geographical boundary and \( w_{jk} = 0 \) otherwise.
Model estimation

• Model parameters to estimate
  – Random effects $u$, other hyper-parameters and $\beta_0$
  – The spatial weights matrix $W$, a new feature of this model

• An iterative estimation procedure
  – The estimation of $(\Theta \mid W)$ treating $W$ as given where $\Theta$ denotes all other unknown quantities
  – The estimation of $(W \mid \Theta)$ using a deterministic procedure.
    • Set $w_{kj} = 0$ if the marginal 95% posterior credible intervals of $u_k$ and $u_j$ do not overlap
    • Set $w_{kj} = 1$ if the marginal 95% posterior credible intervals of $u_k$ and $u_j$ do overlap
  – Iterate the two steps until a termination condition for the hyper-parameter matrix $W$ was met
    • The sequence of estimated $W$ is such that $W^{(t+1)} = W^{(t)}$

• Models implemented using R-INLA, details see Lee and Mitchel (2013)
• **Step 2:** Set a threshold to ensure that the boundary is not just statistically significant from zero, but also substantively different from zero
2. Methods
Impact on Crime

- Are these boundaries linked to crime rates?

- Street-level crime data (https://data.police.uk/) in the South Yorkshire Police force from December 2010 to December 2012
  - aggregated to the LSOA (lower super output areas) units in Sheffield LA.
  - LSOA, a spatial unit with an average population of about 1500

- We look at two groups:
  - Frontier/Social boundary-paired LSOAs
  - border-paired LSOAs
2. Methods
Impact on Crime

• Difference of crime rates between the two groups was compared and tested by using a permutation procedure.

\[
\frac{C_F}{N_F \times P_F} - \frac{C_B}{N_B \times P_B}
\]

– \(C_F\) and \(P_F\) represent the counts of crimes and the total population of paired LSOAs on the opposite sides of social frontiers identified above.
– \(N_F\) and \(N_B\) represent the number of the frontiers and borders
– So the simple text accounts for population distribution and the scale of social frontiers and borders
– 1000 permutations to give an inference on the difference statistic
3. Results

Social boundaries

- Fig. 1 The distribution of the proportion (%) of non-white population in Sheffield in 2001 (left) and 2011 (right)
3. Results

Social boundaries

- Fig. 1 The social boundary/frontiers of ethnicity distribution in Sheffield in 2001 (left) and 2011 (right)
3. Results

Social boundaries

• Fig.2 The social boundary/frontiers of COB in Sheffield in 2001 (left) and 2011 (right)
3. Results
– Permutation Sig. Test

<table>
<thead>
<tr>
<th>Units: Counts / per 1000 persons</th>
<th>Ethnicity frontiers</th>
<th>Country of birth frontiers</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Differences as in Equation (2)</td>
<td>p-values</td>
</tr>
<tr>
<td>All crimes</td>
<td>1.428</td>
<td>0.002</td>
</tr>
<tr>
<td>Burglary crimes</td>
<td>0.096</td>
<td>0.002</td>
</tr>
<tr>
<td>Violent crimes</td>
<td>0.083</td>
<td>0.011</td>
</tr>
<tr>
<td>Vehicle crimes</td>
<td>0.096</td>
<td>0.001</td>
</tr>
<tr>
<td>Shoplifting crimes</td>
<td>0.054</td>
<td>0.046</td>
</tr>
</tbody>
</table>

Table 1. Comparing differences in crime rates between the two groups and the permutation test results.
3. Results

Social boundaries

![Graphs showing social boundary distributions for different crime types with p-values.](image-url)
4. Findings and future work

- Localised spatial modelling provides an methodological framework for identifying boundaries/frontiers in segregation and inequality studies

- These social frontiers seem to make a difference in the distribution of crimes with crimes happening more often near frontiers

- Future work
  - Apply to different cities especially including those with known racial tensions
  - Apply to other aspects of ethnicity & social difference including a multi-variate approach to boundaries such as Religion, Social Class, Deprivation
• Many thanks
• Any comments?