Understanding Violence in Urban Areas: the Case of Rio de Janeiro

Jessie Bullock (joint work with Joana Monteiro and Luciano de Lima Gonçalves)
Department of Government & Institute for Quantitative Social Sciences
Harvard University
www.jessiebullock.com

Motivation
What makes certain neighborhoods more prone to violence than others? Can we predict when and where urban violent crime is more likely to occur?

Conventional explanations of urban violent crime highlight the following risk factors of where urban crime is more likely to occur:

- Areas with low education and low access to education
- Areas with high levels of social exclusion
- Areas with weak institutions, especially weak law enforcement institutions

However, these factors are sticky and change slowly. They do not explain why we see so much variation in aggregate inter-temporal levels of violent crime in a neighborhood predicted to be violent (bottom panel):

![Violent Crime in a Neighborhood](source: Instituto de Segurança Pública)

Operationalizing Violence

Crime and violence data usually suffers from bias due to underreporting or selective reporting in certain areas. To account for this, we use three measures of crime data:

- Geocoded official crime records from police reports
- Geocoded anonymous tips about crime from the 311 hotline
- Text data from a crime blog reporting on conflict between organized criminal factions

Expected Topic Proportions

<table>
<thead>
<tr>
<th>Topic 1: trafic, vil, nao, tcp, comun</th>
<th>Topic 2: trafic, luc, par, comun, com</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topic 3: pedr, trafic, ada, playboy, barr</td>
<td></td>
</tr>
<tr>
<td>Topic 4: trafic, milu, comun, michaels, nass</td>
<td></td>
</tr>
<tr>
<td>Topic 5: trafic, polic, operaca, captur, cv</td>
<td></td>
</tr>
<tr>
<td>Topic 6: complex, trafic, mai, chapels, nass</td>
<td></td>
</tr>
<tr>
<td>Topic 7: trafic, tcp, comun, cv, com</td>
<td></td>
</tr>
<tr>
<td>Topic 8: luc, par, comun, com</td>
<td></td>
</tr>
</tbody>
</table>

Next Steps

- Define types of interactions: invasion, firefight, loss of turf, etc
- Look at temporal exogenous shocks (exchange rate, prisoners)

Theory

We formulate a hypothesis for each “type” of criminal activity:

H1: Illicit market participation groups (drug traffickers) will be more violent in areas with a strategic business interest. This includes areas close to major transport centers in the city and areas close to local consumption markets.

H1a: This violence will be directed at rival groups or law enforcement.

H2: Illicit rent extraction groups (militias) will be more violent in areas isolated from the formal sector and where rule of law is weak. This implies areas with poor transport infrastructure and linkages to the rest of the city.

H2a: This violence will be directed at civilians.

Criminal Factions and Patterns in Violence

We construct time-series data of the dominant faction (militias or drug traffickers) in each urban slum using the following tools:

- Scraped crime blog data from January 2015 - present
- Transcript data from anonymous calls January 2015 - present
- Police interviews from January - December 2015

Text data speaks to the microdynamics of violence and helps us appropriately characterize who is fighting whom, and over what. A preliminary STM analysis of blog data shows three patterns:

- Common topics refer to police conducting operations in slums to capture traffickers (Topics 10 and 5)
- Data includes the geographic traffickers or militias occupy: a slum, complex, or community (Topics 9, 4, 7, 1, 6)
- Data includes details on a hub of trafficking activity and kingpin, the “Playboy” (Topic 3)