

# Subjective Neighborhood Identification and Analysis

Cory McCartan,<sup>1</sup> Jacob R. Brown,<sup>2</sup> Kosuke Imai<sup>1,2</sup>

1. Department of Statistics, Harvard University  
2. Department of Government, Harvard University

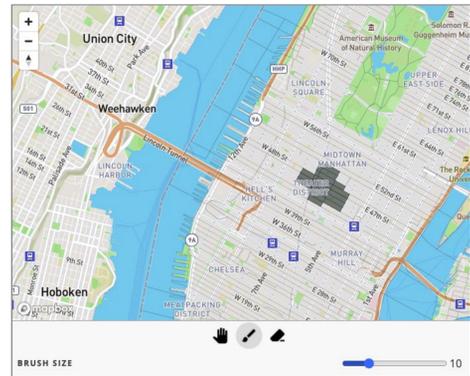
## Background

- Many social science studies about neighborhoods:
  - Effects of segregation on inter-group conflict, social trust, and socio-economic outcomes [1, 2, 3]
  - Behaviors spreading through geographic networks [4]
- Limited methods to measure how and why people define their neighborhood
  - Objective measures (distance, administrative boundaries)
  - Subjective definitions have real-world effects [3]

# 471 voters in 3 cities drew us their neighborhoods.

# We developed a model to analyze them.

**1** We built a **custom survey tool** that allows respondents to easily **draw their neighborhood** on a map.



[TRY IT OUT]

### The TOOL

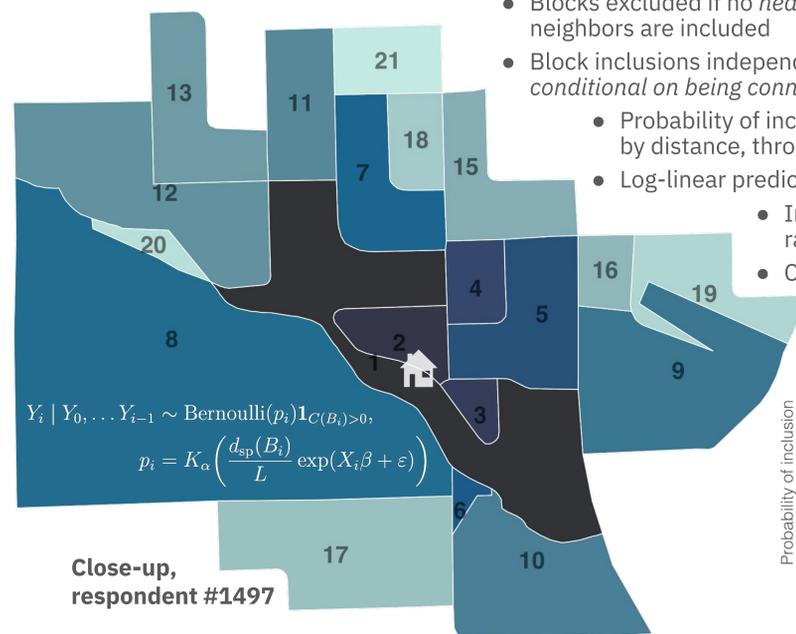
- Respondent types in home address and map zooms to local area
- Respondent can “paint” and “erase” parts of the neighborhood on the map
- Editor enforces contiguity requirement

### The SURVEY

- E-mailed to sample of voter file in NYC, Phoenix, and Miami metropolitan areas
- Collect demographics, political views
- This poster analyzes only the initial wave of 471 complete responses

**2** We fit a **hierarchical Bayesian model** incorporating demographic information and local geographic features.

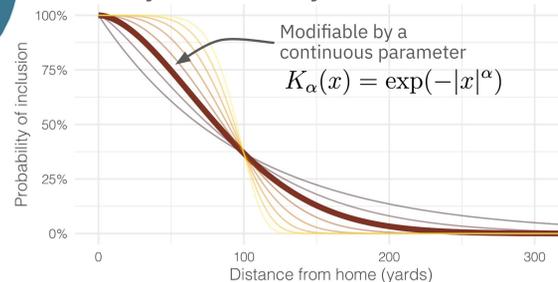
- Model the probability that each Census block is included in the neighborhood
- Visit blocks one at a time, working outwards from respondent’s home (as below)



Close-up, respondent #1497

- Blocks excluded if no *nearer* neighbors are included
- Block inclusions independent, *conditional on being connected*
  - Probability of inclusion driven by distance, through kernel function
  - Log-linear predictor for covariates
- Individual-level random effects
- Can be reduced to GLMM with cloglog link

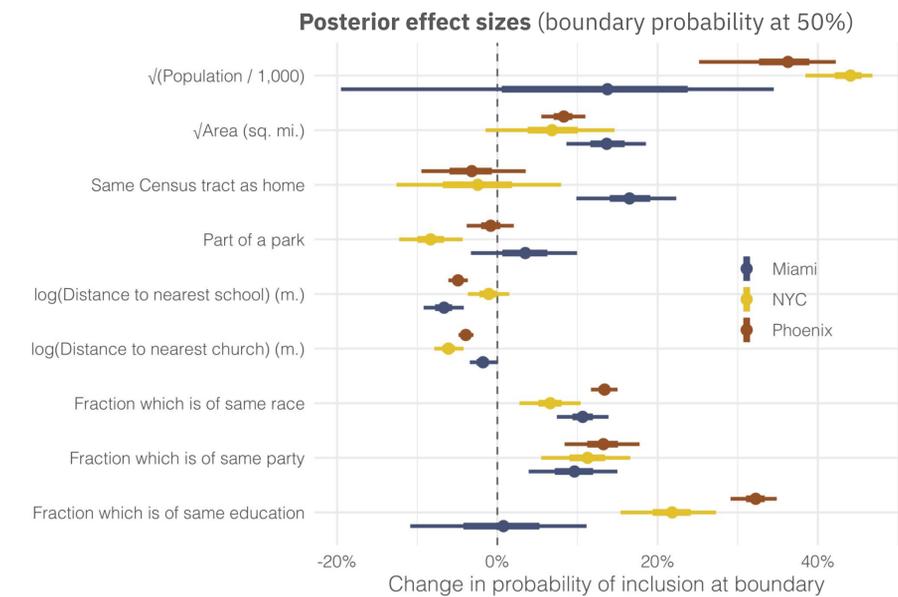
### Probability of inclusion by distance



Pct. white 0% 25% 50% 75% 100%

Shown above is the neighborhood and surrounding area of respondent #1497, selected for display due to their location near a strong racial boundary.

**3** Coefficient estimates show the importance of local features, and the **consistent influence of demographics**.



**4** We can **simulate from the model** to understand how subjective perceptions of neighborhood are shaped.

### Posterior distribution of neighborhood demographics

Respondent #1497: baseline versus full model



Difference in inclusion probability -10.0% -5.0% 0.0% 5.0%

Baseline model removes demographic covariates.

1. Massey, D. S. & Denton, N. A. (1993), *American Apartheid: Segregation and the Making of the Underclass*, Harvard University Press, Cambridge, MA.  
2. Dinesen, P. T. & Sønderskov, K. M. (2015), 'Ethnic diversity and social trust evidence from the micro-context', *American Sociological Review* 80(3), 550–573.  
3. Legewie, J. & Schaeffer, M. (2016), 'Contested boundaries: Explaining where ethnoracial diversity provokes neighborhood conflict', *American Journal of Sociology* 122(1), 125–161.  
4. Huckfeldt, R. & Sprague, J. (1987), 'Networks in context: The social flow of political information', *American Political Science Review* 81(4), 1197–1216.