Addressing Unobserved Neighborhood Quality in Hedonic Models Using Google Street View Imagery and Urban Perception Data

Zander Gordan

Andrew Young School of Policy Studies
Georgia State University

agordan1@student.gsu.edu

August 13, 2018

Unobserved Neighborhood Quality in Housing Hedonics

$$p(x, g) = \alpha \cdot h(x) + \beta \cdot g + \epsilon$$

p = price
x = housing characteristics
g = neighborhood characteristics (schools, pollution, etc.)

Concern: $Cov(g, \epsilon) \neq 0$

Aesthetic Character of Neighborhoods

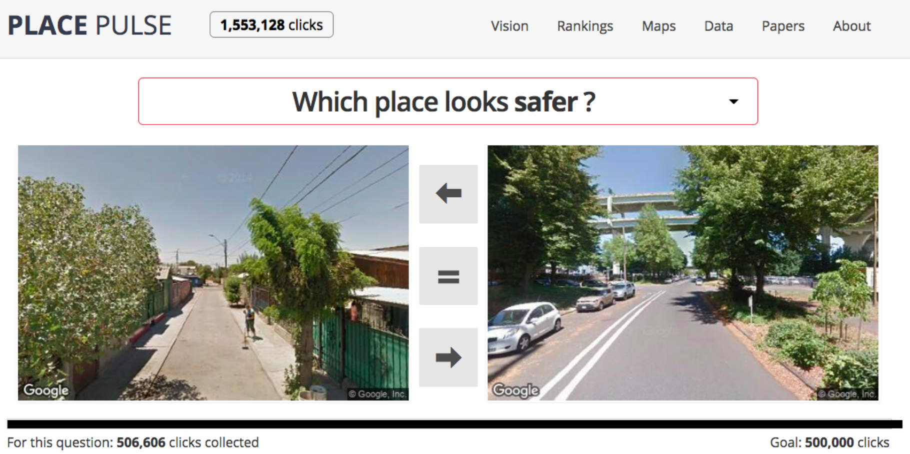
Aesthetic Character of Neighborhoods

Google Street View Imagery

- Extensive coverage
  - Every MSA in the US
  - increasing global coverage
- Data collection began in 2007, many cities now have multiple years of observation

But how do we use this information?

Place Pulse Survey



Place Pulse Survey

- 56 cities from 28 countries across 6 continents
- Over 100,000 images
- Over 1.5 million comparisons

Enough data to *generalize*

Binary Comparisons → Rankings

Treat each comparison as a 'competition' between images, where performance levels are drawn from a normal distribution centered on each contestant's rank. Iteratively update rankings with the Microsoft 'TrueSkill' algorithm.

$$\mu_{winner} \leftarrow \mu_{winner} + \frac{\sigma^2_{winner}}{c} \cdot v\left(\frac{(\mu_{winner} - \mu_{loser})}{c}, \frac{\varepsilon}{c}\right)$$

$$\mu_{loser} \leftarrow \mu_{loser} - \frac{\sigma^2_{loser}}{c} \cdot v\left(\frac{(\mu_{winner} - \mu_{loser})}{c}, \frac{\varepsilon}{c}\right)$$

$$\sigma^2_{winner} \leftarrow \sigma^2_{winner} \cdot \left[1 - \frac{\sigma^2_{winner}}{c^2} \cdot w\left(\frac{(\mu_{winner} - \mu_{loser})}{c}, \frac{\varepsilon}{c}\right)\right]$$

$$\sigma^2_{loser} \leftarrow \sigma^2_{loser} \cdot \left[1 - \frac{\sigma^2_{loser}}{c^2} \cdot w\left(\frac{(\mu_{winner} - \mu_{loser})}{c}, \frac{\varepsilon}{c}\right)\right]$$

$$c^2 = 2\beta^2 + \sigma^2_{winner} + \sigma^2_{loser}$$

Deep Learning The City