The Usefulness of Models

Schelling's Segregation

John Eric Humphries

University of Chicago, Department of Economics

THE UNIVERSITY OF CHICAGO

Lets Be Serious...

... **most economic models look nothing like reality**

So why do we use them?

They:

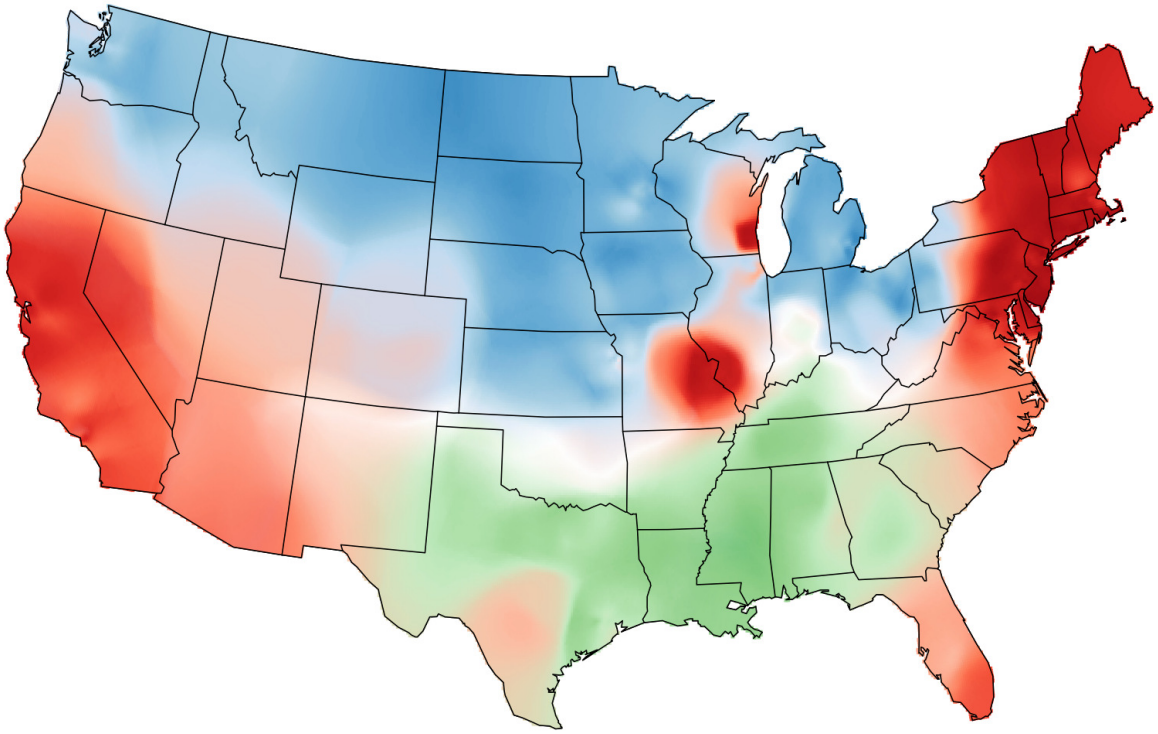
- oversimplify reality
- don't account for things we know matter
- most do not even use data...

THE UNIVERSITY OF CHICAGO

Yet...

- Models may describe aggregate behavior surprisingly well.
- Models provide insights into data.
- Theory from models can aid econometric estimation.
- Models create testable hypotheses for data.
- Competing models can be compared and tested against data in an iterative process (models and data!).

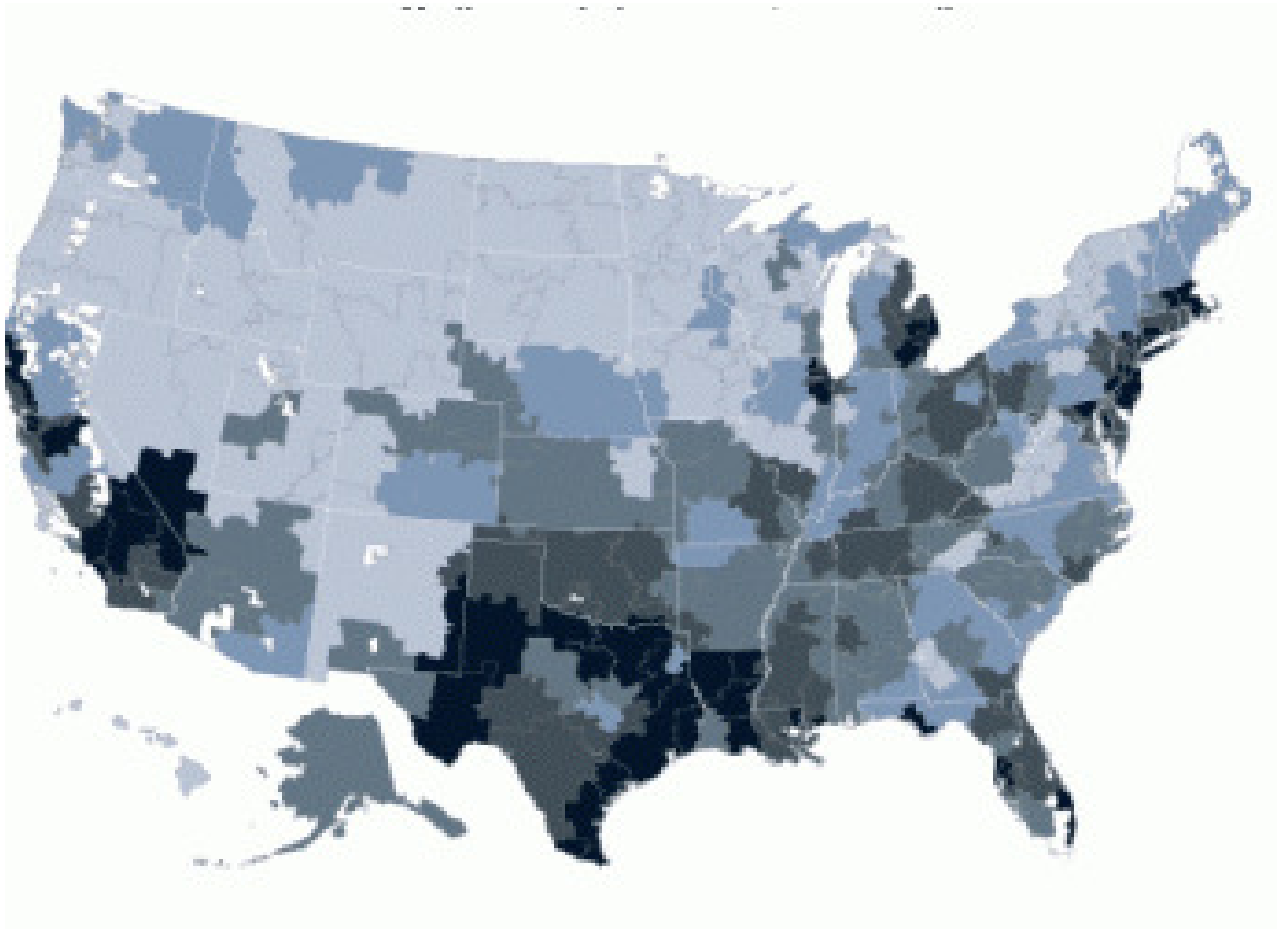
But lets consider some data without theory:



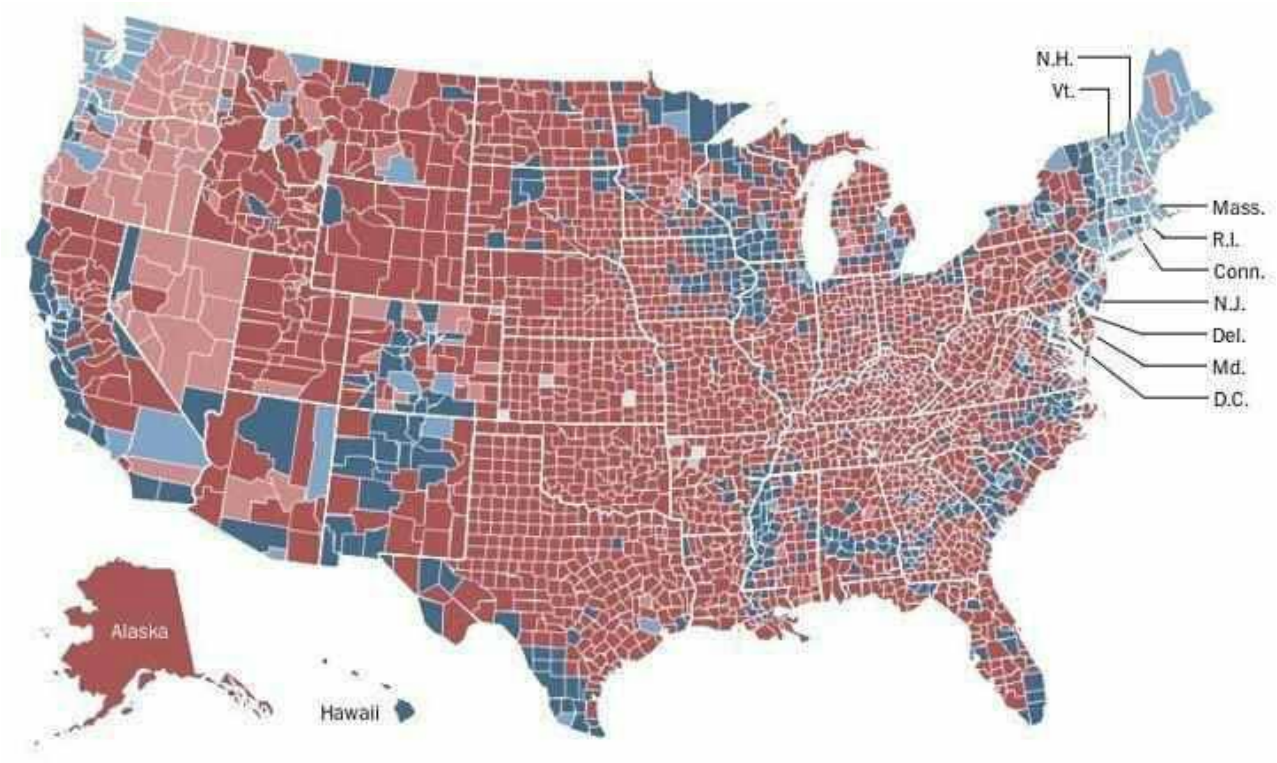
Map by Joshua Katz, Department of Statistics, NC State University
Based on survey data from Bert Vaux, Department of Linguistics, University of Cambridge

THE UNIVERSITY OF CHICAGO

But lets consider some data without theory:



But lets consider some data without theory:



But lets consider some data without theory:

- What is driving the clustering in these three images?
- sorting? homophily? peer effects?
- hard to say without some theory

- First map is a map of usage of the words “pop”, “soda”, and “coke” (http://laughingsquid.com/soda-pop-or-coke-maps-of-regional-dialect-variation-in-the-united-states/)

- Second is a map of medicare reimbursement per enrollee (http://dd.dynamicdiagrams.com/2010/06/visual-bias-at-work/)

- Third is a map of which counties went Republican and Democrat in the 2012 presidential election. (http://www.huffingtonpost.com/tim-young/logic-from-a-republican-w_b_2093378.html)

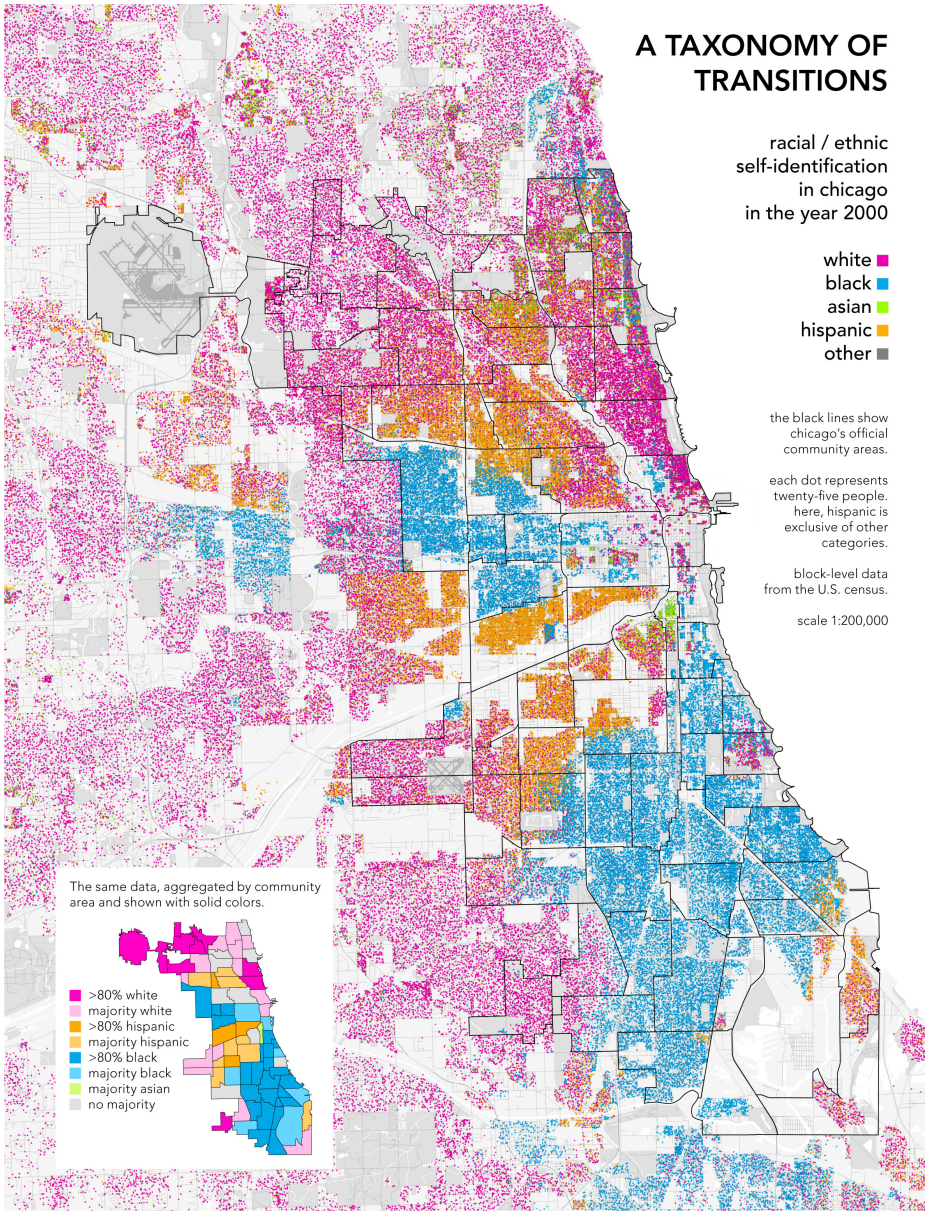
Could be driven by...

- Homophily?
- Peer Effects

How could we tell different models apart?

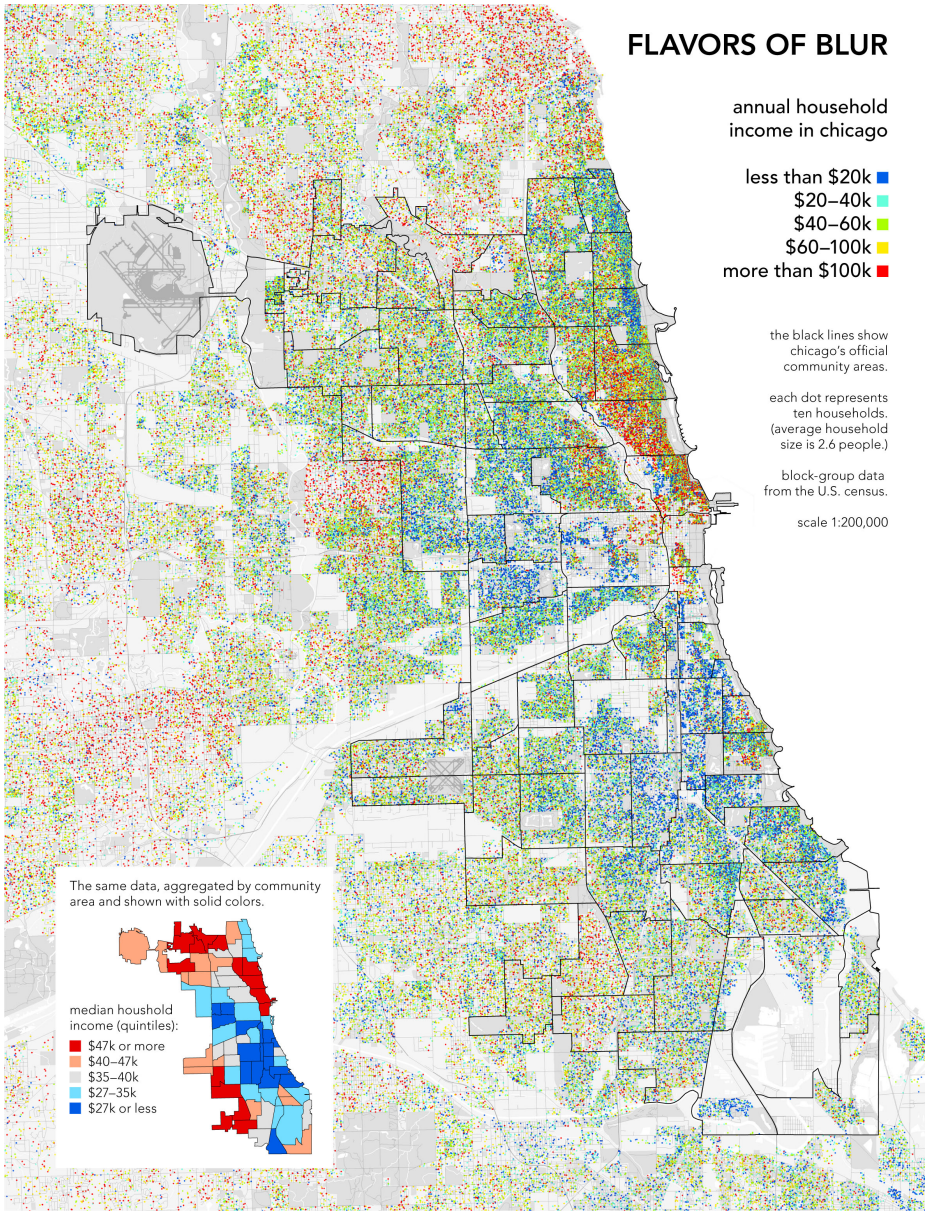
What kind of data would we need to test them?

Segregation By Race (Bill Rankin 2009)



- Maybe racial segregation is driven by income differences and property prices?

Segregation By Income (Bill Rankin 2009)



Metrics of Segregation

- How should we measure segregation?

What does segregation mean?

- "Evenness"
- "Exposure"
- "Connectedness"

Typically have to work with some level of aggregation.

James and Taeuber (1985) discuss ideal properties of "evenness" metrics:

- Transfer principle: index should be affected by reallocating minorities from a district which is slightly above average minority to one which is mostly minorities
- Compositional invariance: relative size of the minority population does not affect the index
- Size invariance: total population does not affect index
- Organizational Equivalence: aggregating geographic areas with same minority compositions does not affect the index

Very few metrics satisfy all four criteria!

Only Atkinson and Gini measure satisfy all four criteria.

Dissimilarity Index

Using the notation from the handout I link to in pset 2. Let there be n areas and let P be the city-wide proportion of the population that is minority while p_i is the proportion of the population in area i that is minority. T is total population in the city while t_i is the total population in area i. Then the dissimilarity index is given by:

$$\frac{\sum_{i=1}^{n} [t_i \mid (p_i - P) \mid]}{2TP(1-P)}$$

Gini Index (not much different than dissimilarity index)

Letting t_j similarly be population in area j and p_j the proportion of minorities in area j, we can write the Gini Index as:

$$\frac{\sum_{i=1}^{n}\sum_{j=1}^{n}\left[t_i t_j \mid (p_i - p_j) \mid\right]}{2T^2P(1-P)}$$

Sources

This lecture drew heavily from the coures "Model Thinking" from the University of Michigan and on Coursera
https://www.coursera.org/course/modelthinking