

Big Data and Economics

Neighborhoods and Upward Mobility

Kyle Coombs
Bates College | EC/DCS 368

Table of contents

1. Prologue
2. Geographical Variation in Upward Mobility
3. Characteristics of High-Mobility Areas
4. Spatial Correlation and Decay

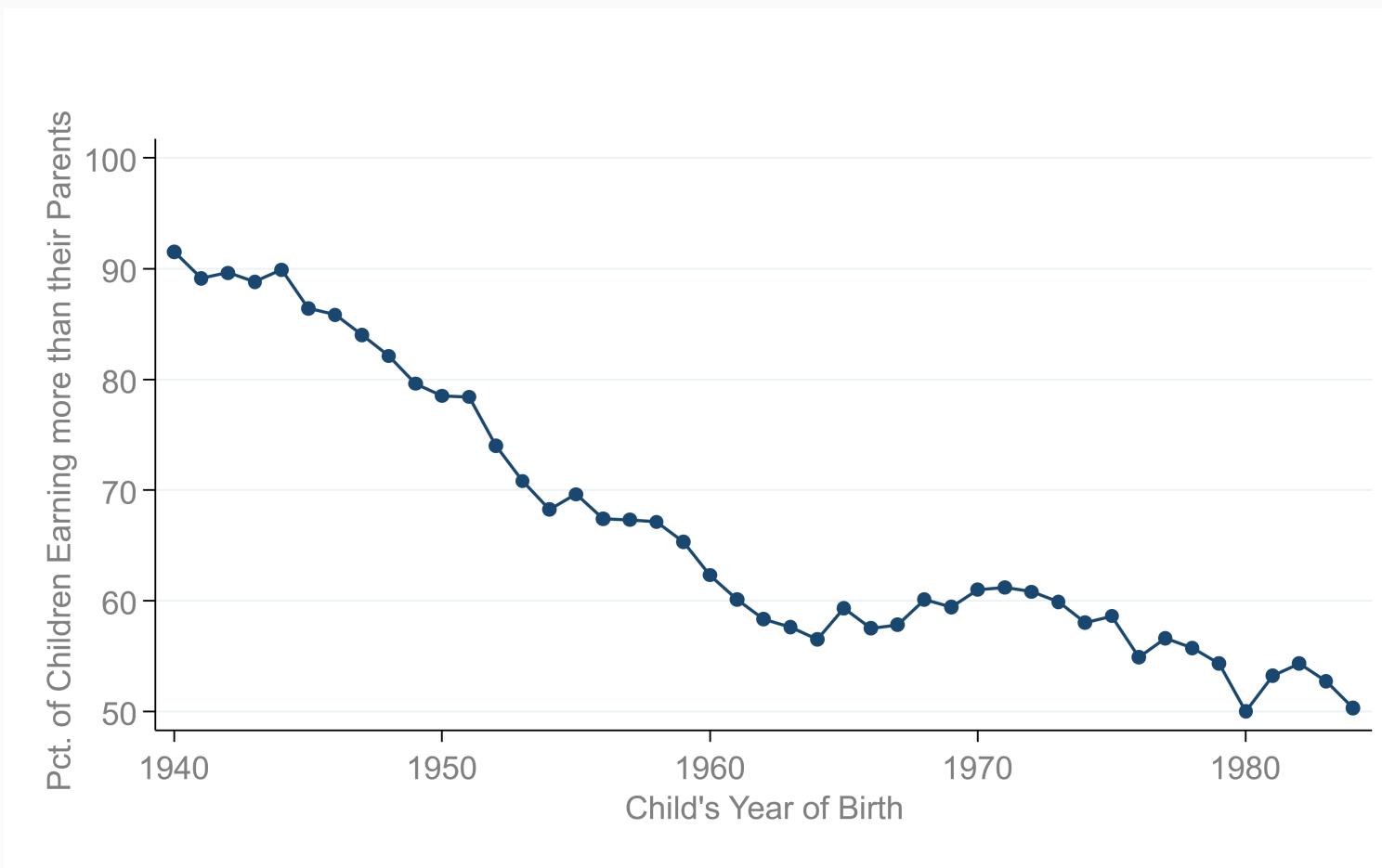
Prologue

Housekeeping

- Problem Set 2 is due on Friday
- Your data description is due after break

Prologue

- Today's lecture is a little different than the last few
- We're talking about an application of big data to a big question: why do some people move up the income ladder and others don't?
- This is a big question in economics and public policy
- Chetty answered it using big data and spatial analysis
 - By big I mean: essentially all tax returns in the USA from 1989-2015
- He released summaries of the data publicly in 2018 as the Opportunity Atlas
- These show tons of descriptive measures of income mobility at various levels of geography: state, county, and Census Tract
- Problem Set 3 will involve using the Census Tract data to learn about income mobility in Lewiston



Source: Chetty et al. (2014)

Why is the "American Dream" Fading?

- Why are children's chances of climbing the income ladder falling in the USA?
 - What can be done to reverse this trend?
- Need to go beyond macroeconomic data to answer this question. Why?
 - Too many changes happening over time and across space to separate out the causal factors.
 - Also: only a handful of data points (classic macro problem)

Enter the Opportunity Atlas

- Created in 2018, the Opportunity Atlas offers one measure of how income mobility differs by location in the USA
 - If some areas have more mobility than others, can we learn why and apply those lessons elsewhere?
- Data sources:
 - Anonymized Census data (2000, 2010 ACS) covering U.S. population
 - Federal income tax returns from 1989-2015.
- Method: Link parents based on dependent claiming on tax returns
- Target sample: Children born between 1978-1983 (U.S. citizens and authorized immigrants who arrived as children)

There's bound to be a mess with this much data, so they create an analysis sample

- **Analysis sample:** 20.5 million children, 96% coverage of target sample

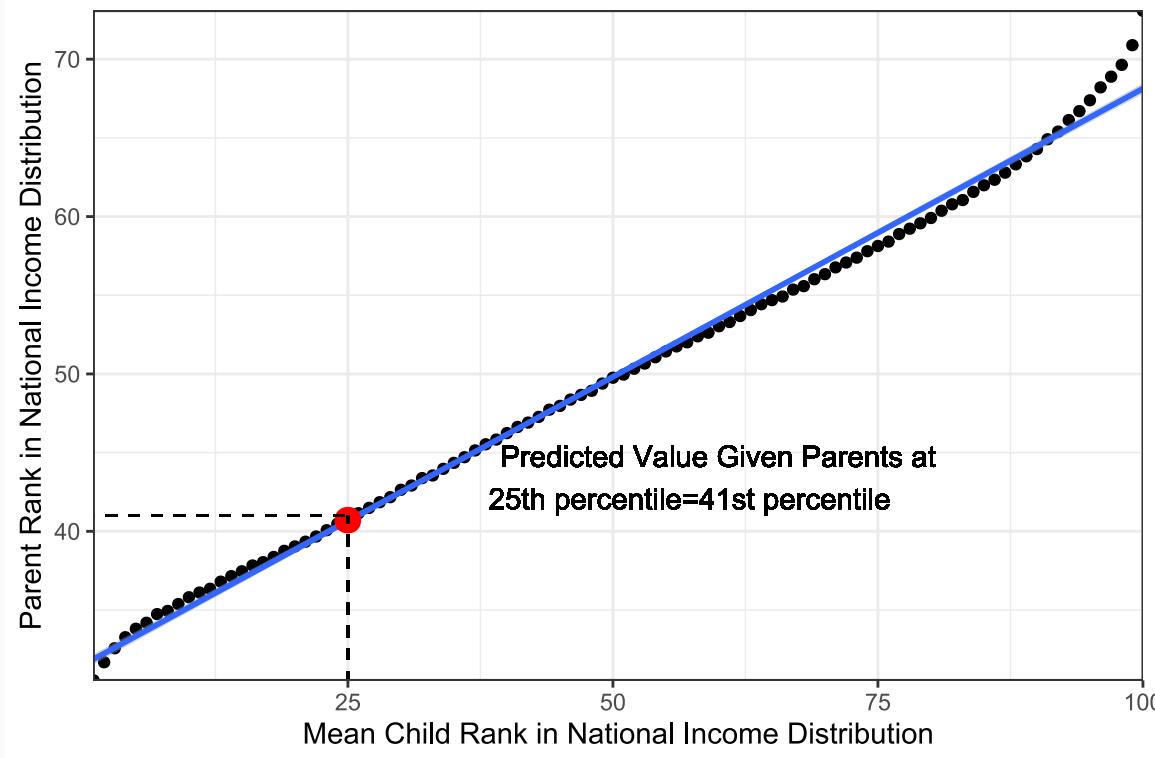
Toolkit to use these data

- Data cleaning and wrangling
- Data visualization
- Spatial analysis (this week)
- Regression analysis (after break)

Parent and Children Incomes in Tax Data

- Parent household incomes: average income reported on Form 1040 tax return from 1994-2000
- Children incomes measured from tax returns in 2014-15 (ages 31-37)
- But income levels differ over time! How do we compare them?
 - Use percentile ranks in the *national* distribution
 - Rank children relative to others born in same year and parents relative to other parents
- **Income percentile:** The fraction of the national income distribution that a person's income exceeds
- Take average income percentile of children by parental income percentile

Average Child Income Percentile by



Source: The Opportunity Atlas

Geographic Variation in Upward Mobility

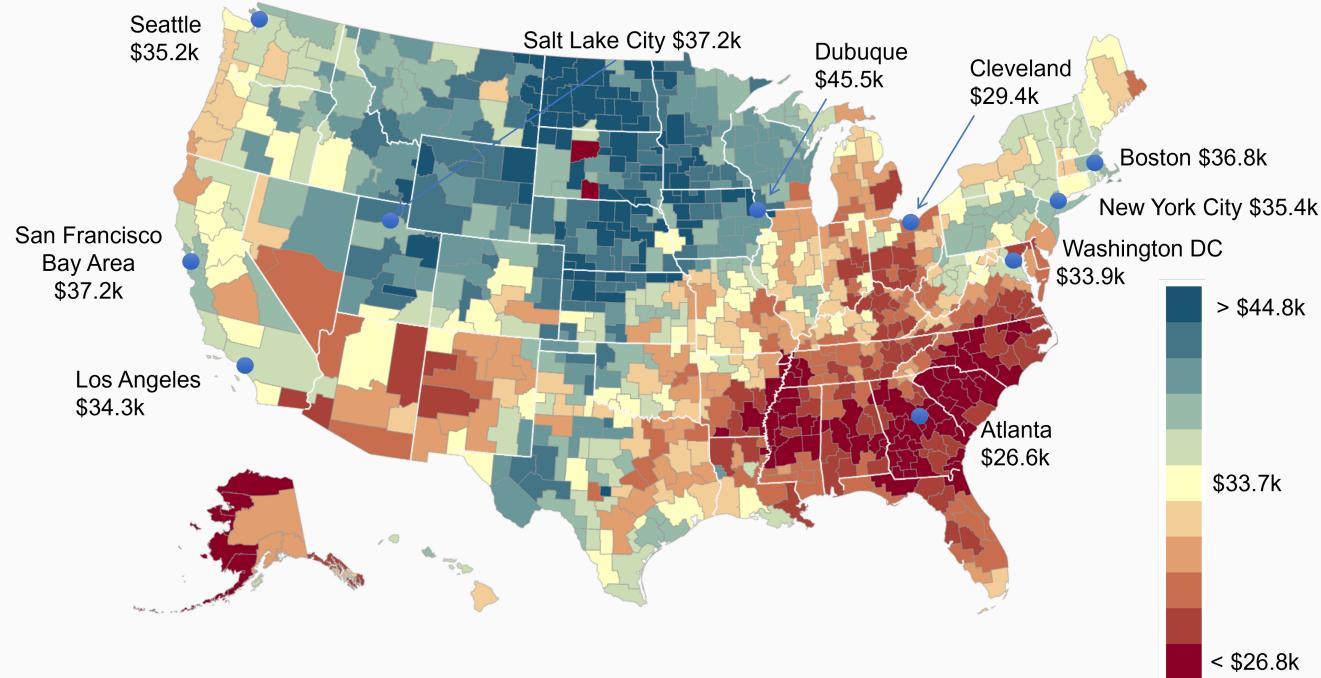
What is mobility for a given area?

- Run this same regression of income ranks by Census tract, county, or commuting zone in the USA¹
 - Census tracts are small geographic areas that contain 1,200-8,000 people
- For simplicity, Chetty et al. (2018) report the average income percentile of children whose parents were at the 25th percentile of the national income distribution
- This is a single measure of upward mobility that is easy to understand and compare across areas
 - It is not the only measure, but it is a good one
- **Big data tip:** Sensibly summary statistics make big data more useful
 - The right statistic depends on the question you're asking

¹ Technical detail: Weight each child by fraction of childhood (up to 23) in a given area to account for movement across areas during childhood

The Geography of Upward Mobility in the United States

Average Household Income for Children with Parents Earning \$27,000 (25th percentile)



Note: Blue = More Upward Mobility, Red = Less Upward Mobility Source: [The Opportunity Atlas](#)

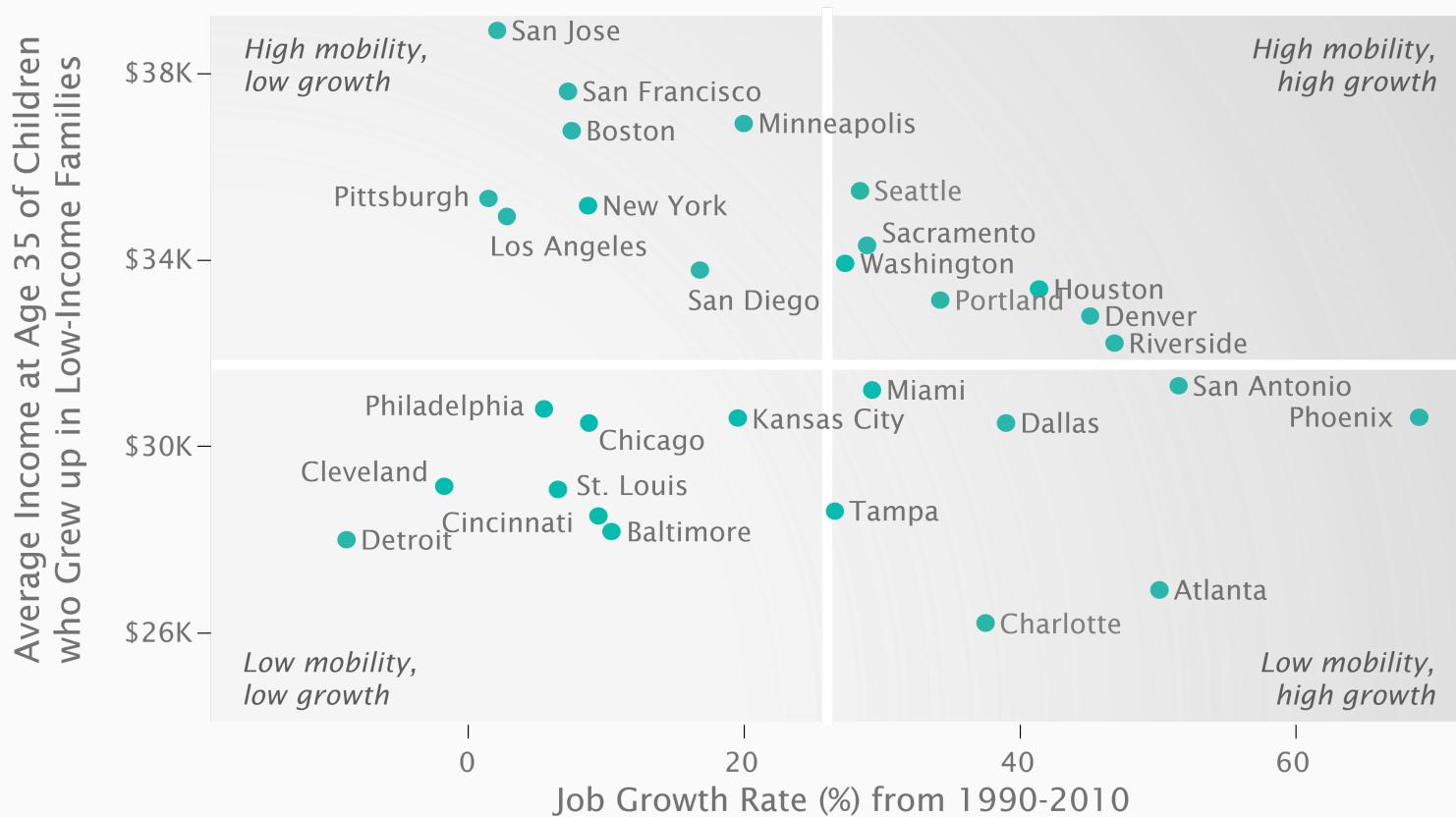
Characteristics of High-Mobility Areas

Why does upward mobility differ?

Armed with a summary measure of upward mobility, we can ask:

- Why do some areas have more upward mobility than others?
- Spatial and correlational analysis is a good place to start
- What are potential characteristics of high mobility areas?
 - Better jobs?
 - Better schools?
 - Institutional differences?
 - Culture?

Upward Mobility vs. Job Growth



Actual correlates

1. Segregation: Greater racial and income segregation associated with lower levels of mobility
2. Income Inequality: Places with smaller middle class have less mobility
3. School Quality: Higher expenditure, smaller classes, higher test scores correlated with more mobility
4. Family Structure:
 - Areas with more single parents have lower mobility
 - Strong correlation even for kids whose own parents are married
5. Social Capital
 - It takes a village to raise a child
 - Chetty et al. (2023) leveraged Facebook Data to create the Social Capital Atlas

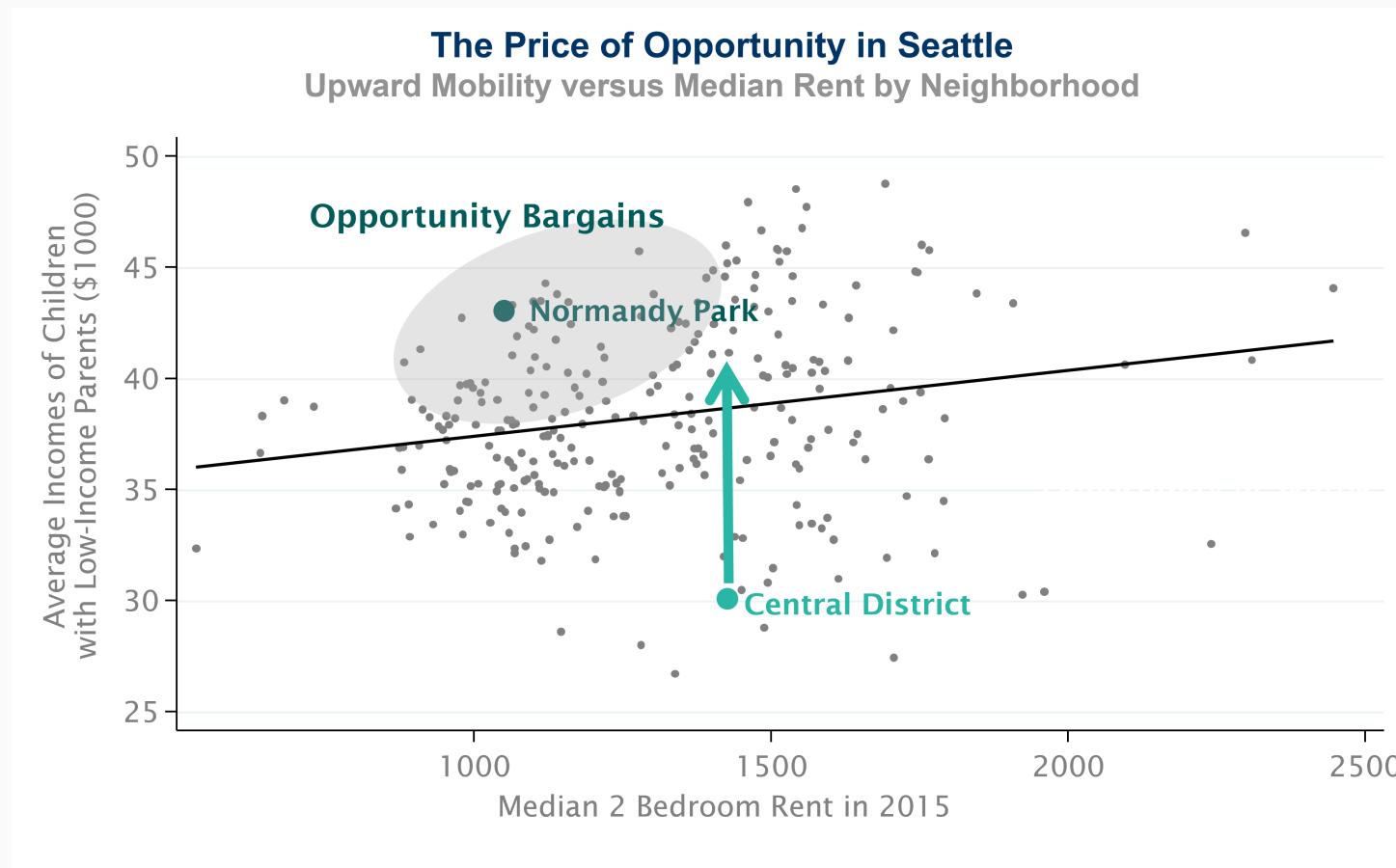
Spatial Correlation and Decay

Big question: why don't people move?

- If some areas have more mobility than others, why don't people move to those areas?
- Is it rent?

The Price of Opportunity in Seattle

Upward Mobility vs Median Rent by Neighborhood



Big question: why don't people move?

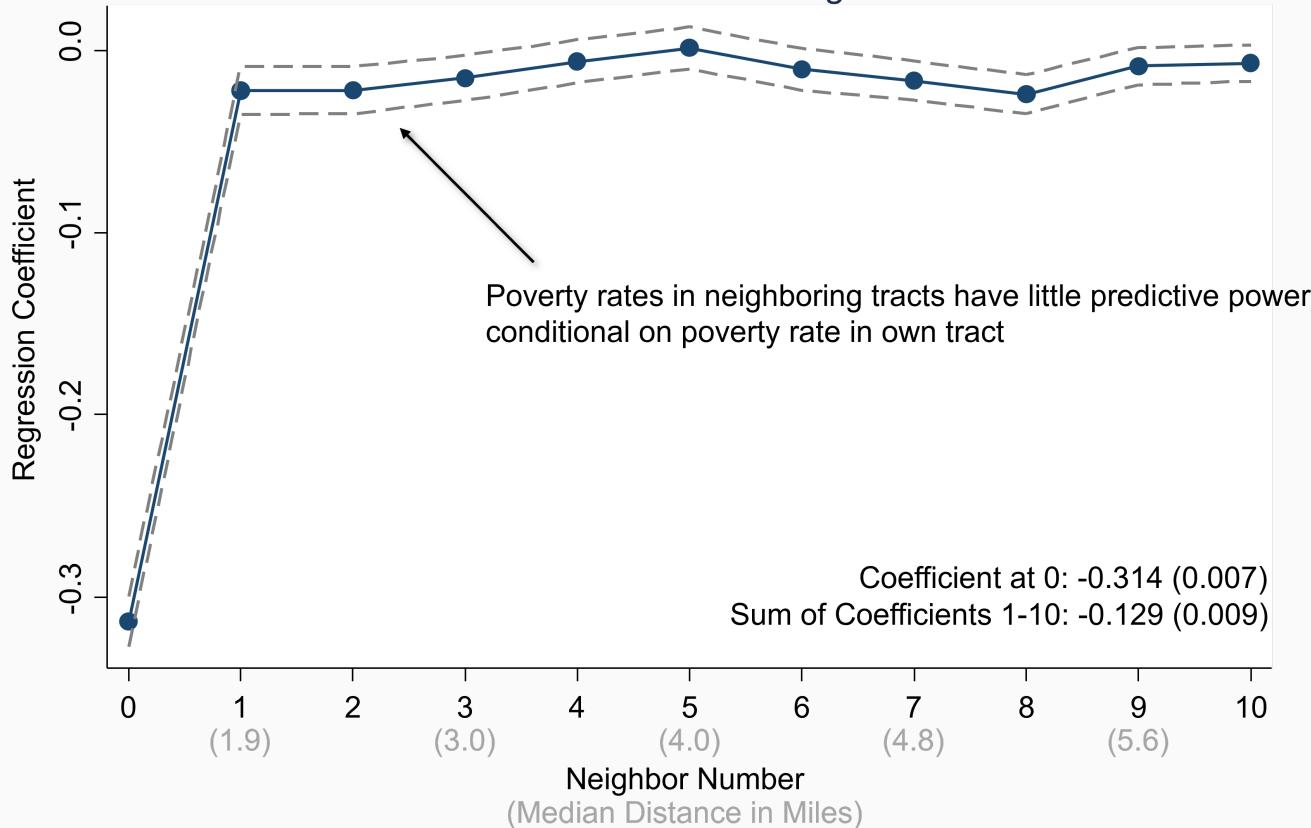
- Initial experiments indicate benefits exist from moving (we'll see later)
- If some areas have more mobility than others, why don't people move to those areas?
- Is it rent?
- Other costs of moving?
- Maybe they do not want to move as far?
- Overall, this is not a highly effective approach

Well what if we invest locally?

- What if we invest in the areas that have low mobility? (place-based approach)
- Would there be spillovers between locations?
 - It is tough to improve one neighborhood (e.g. a tract), let alone many at once
 - Do we have to improve them all at once to help people?
- The answer to this question changes the policy approach

Spatial decay suggests localized effects

Spatial Decay of Correlation between Upward Mobility and Tract-Level Poverty Rates
Estimates from Multivariable Regression



Overall Takeaways

- Correlation evidence is suggestive, but not causal
- Causality requires a more focused approach
- We will build this toolkit in the next few lectures

Next lecture: Regression and
Opportunity Atlas
