

# Using Federal Administrative Data to Evaluate and Improve Economic Policy

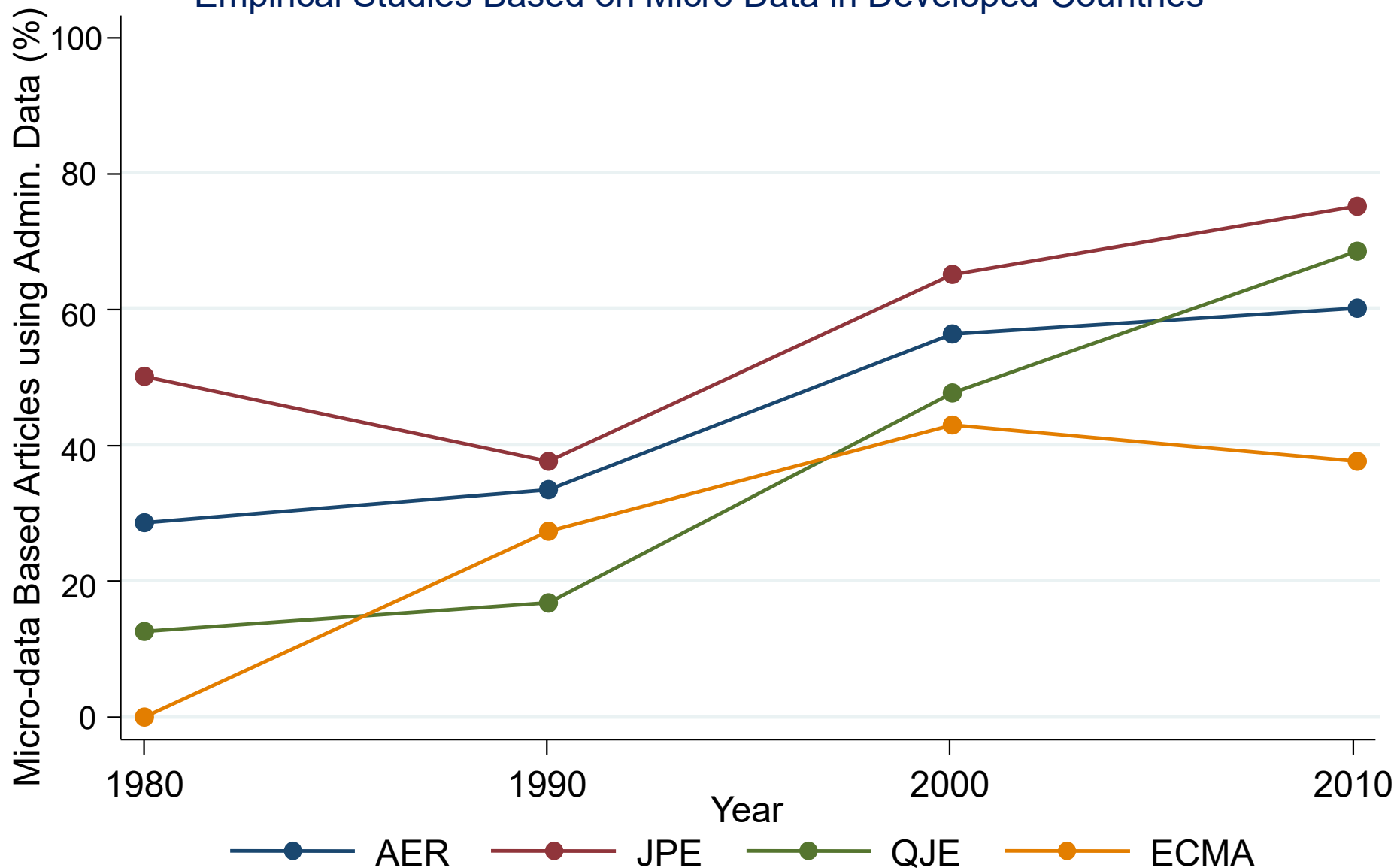
Raj Chetty

Stanford and Harvard

*The opinions expressed in this paper are those of the authors alone and do not necessarily reflect the views of the Internal Revenue Service or the U.S. Treasury Department. This work is a component of a larger project examining the effects of eliminating tax expenditures on the budget deficit and economic activity, approved under IRS contract TIRNO-12-P-00374.*

# Use of Administrative Data in Publications in Economics Journals, 1980-2010

Empirical Studies Based on Micro Data in Developed Countries



Note: "Administrative" datasets refer to any dataset that was collected without directly surveying individuals (e.g., scanner data, stock prices, school district records, social security records). Sample excludes studies whose primary data source is from developing countries.

# Why the Shift Toward Administrative Data?

- Administrative data has great value for several reasons:
  1. Comprehensive, high quality data → gold-standard descriptive statistics
  2. Large samples → quasi-experimental methods of causal inference
  3. Longitudinal tracking without attrition → long-term evaluations
  4. Ability to link other datasets → rich set of outcomes

# Administrative Data in the United States

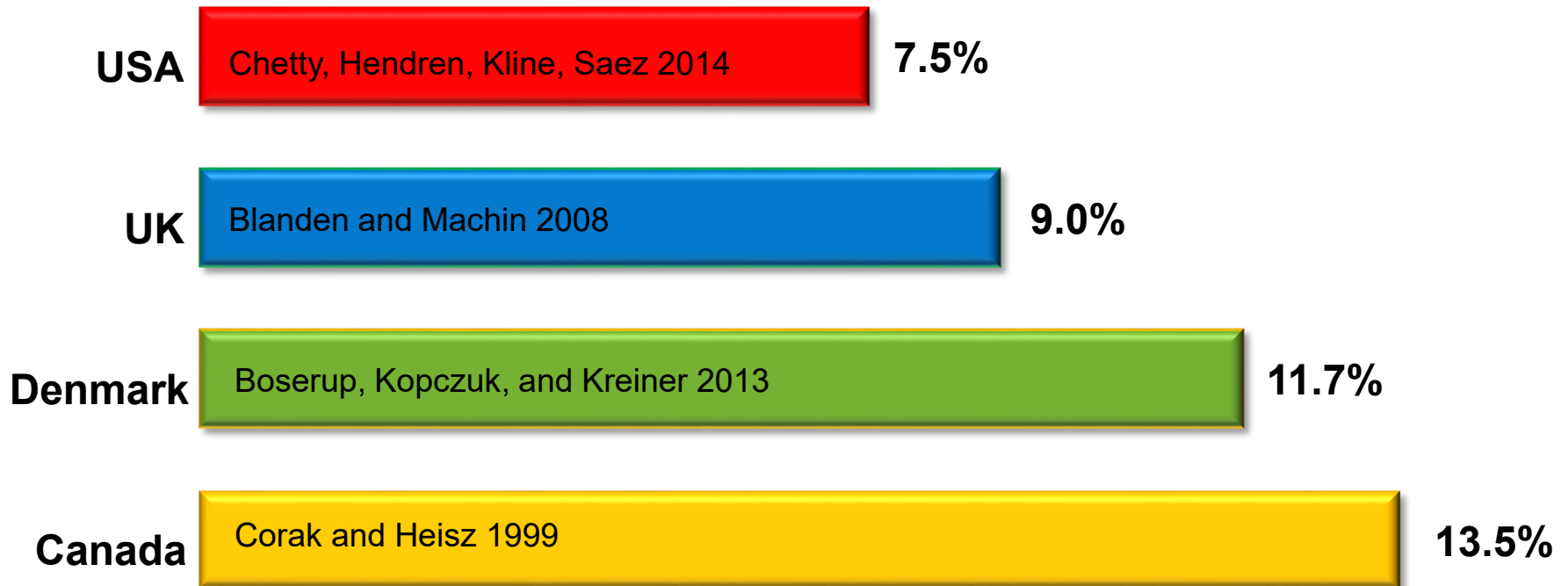
- Researchers are shifting away from studying the U.S. because admin. data are more accessible in Europe
  - U.S. government agencies have made important strides to counter this trend in recent years, but more work is needed
- Goal of this talk: illustrate the benefits of administrative data and downstream impacts on policy in the U.S.
  - Discuss recent studies in our research group that analyze how to improve equality of opportunity in America
  - Part of a larger project studying tax expenditures in the U.S.

# The American Dream?

- Probability that a child born to parents in the bottom fifth of the income distribution reaches the top fifth:

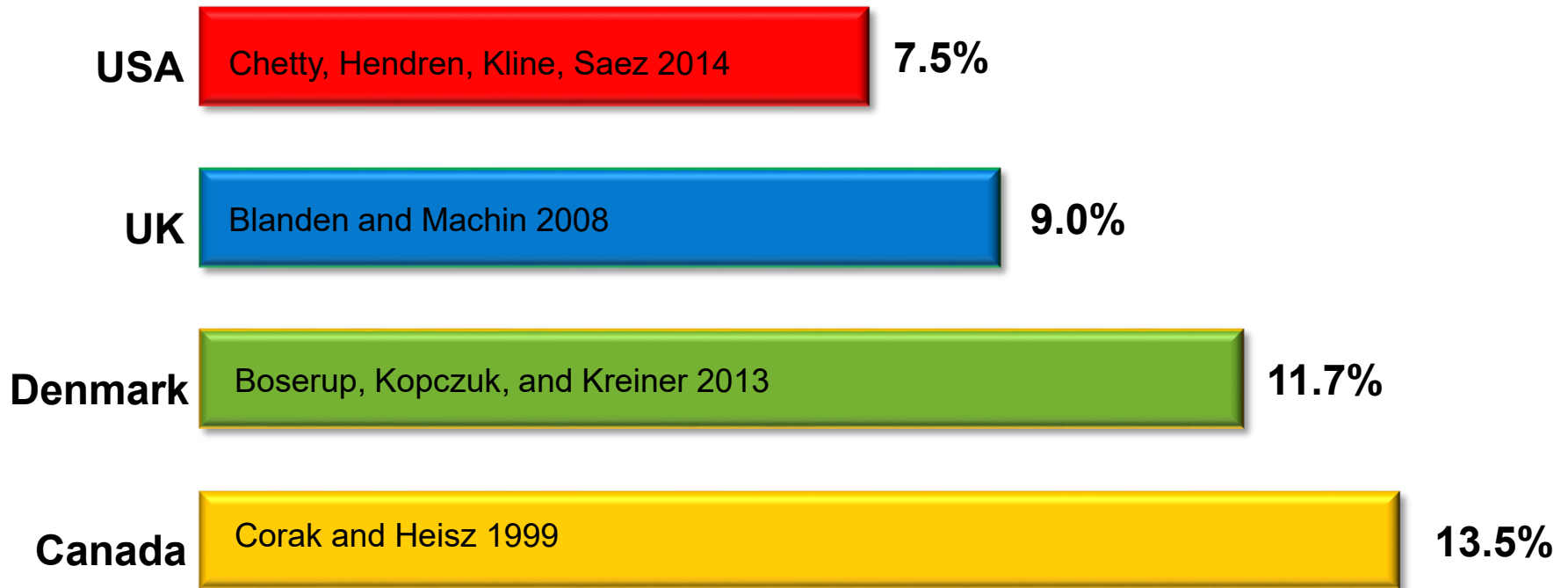
# The American Dream?

- Probability that a child born to parents in the bottom fifth of the income distribution reaches the top fifth:



# The American Dream?

- Probability that a child born to parents in the bottom fifth of the income distribution reaches the top fifth:



→ Chances of achieving the “American Dream” are almost two times higher in Canada than in the U.S.

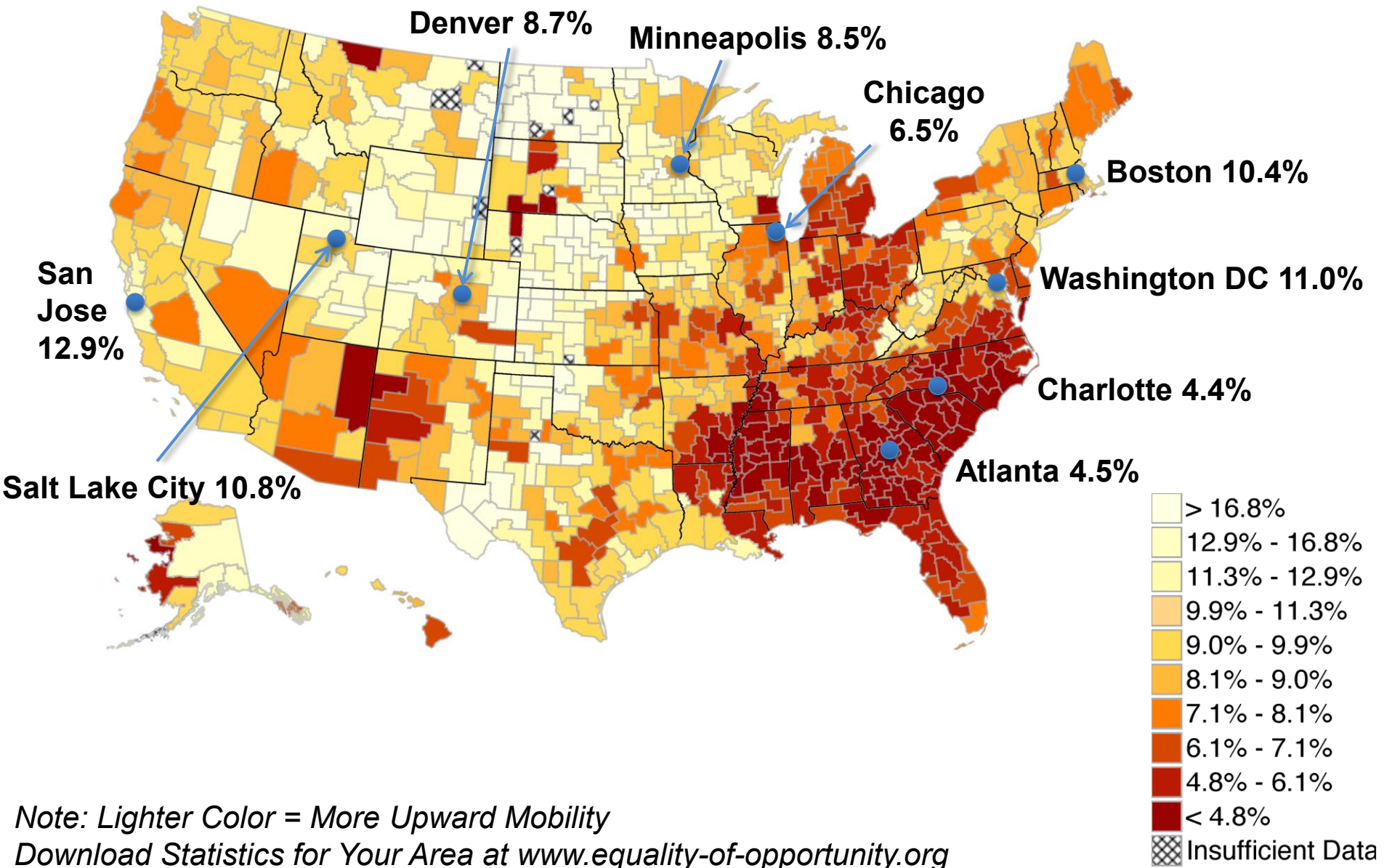
# Differences in Opportunity Within the U.S.

- Research on mobility has traditionally focused on differences across countries
- But social mobility varies even more *within* the U.S.
- We calculate upward mobility for every metro and rural area in the U.S.
  - Use anonymous data on earnings draw from tax records on 40 million children born between 1980-1993



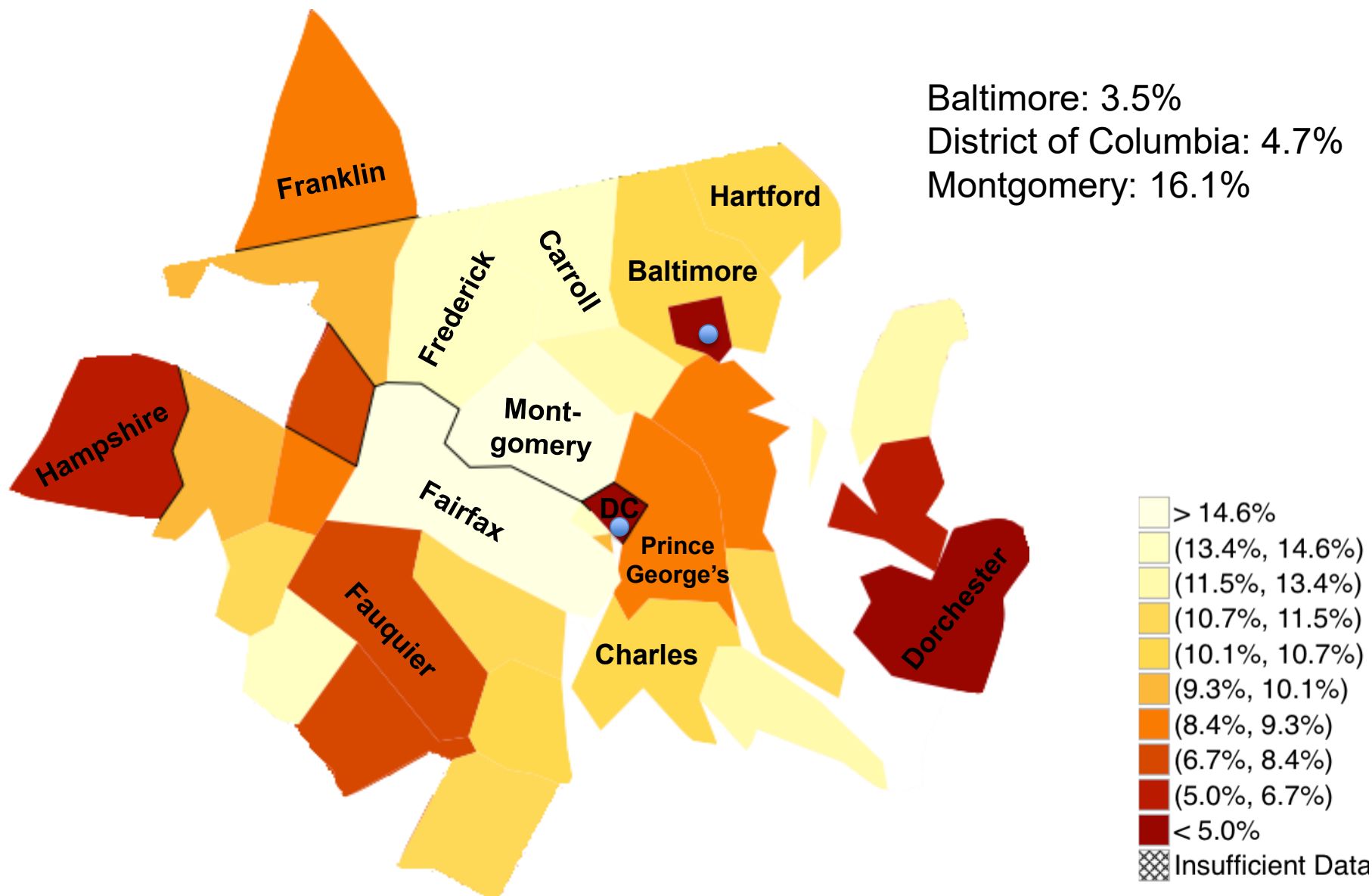
# The Geography of Upward Mobility in the United States

Chances of Reaching the Top Fifth Starting from the Bottom Fifth by Metro Area



# The Geography of Upward Mobility in the Washington Metro Area

Odds of Reaching the Top Fifth Starting from the Bottom Fifth by County



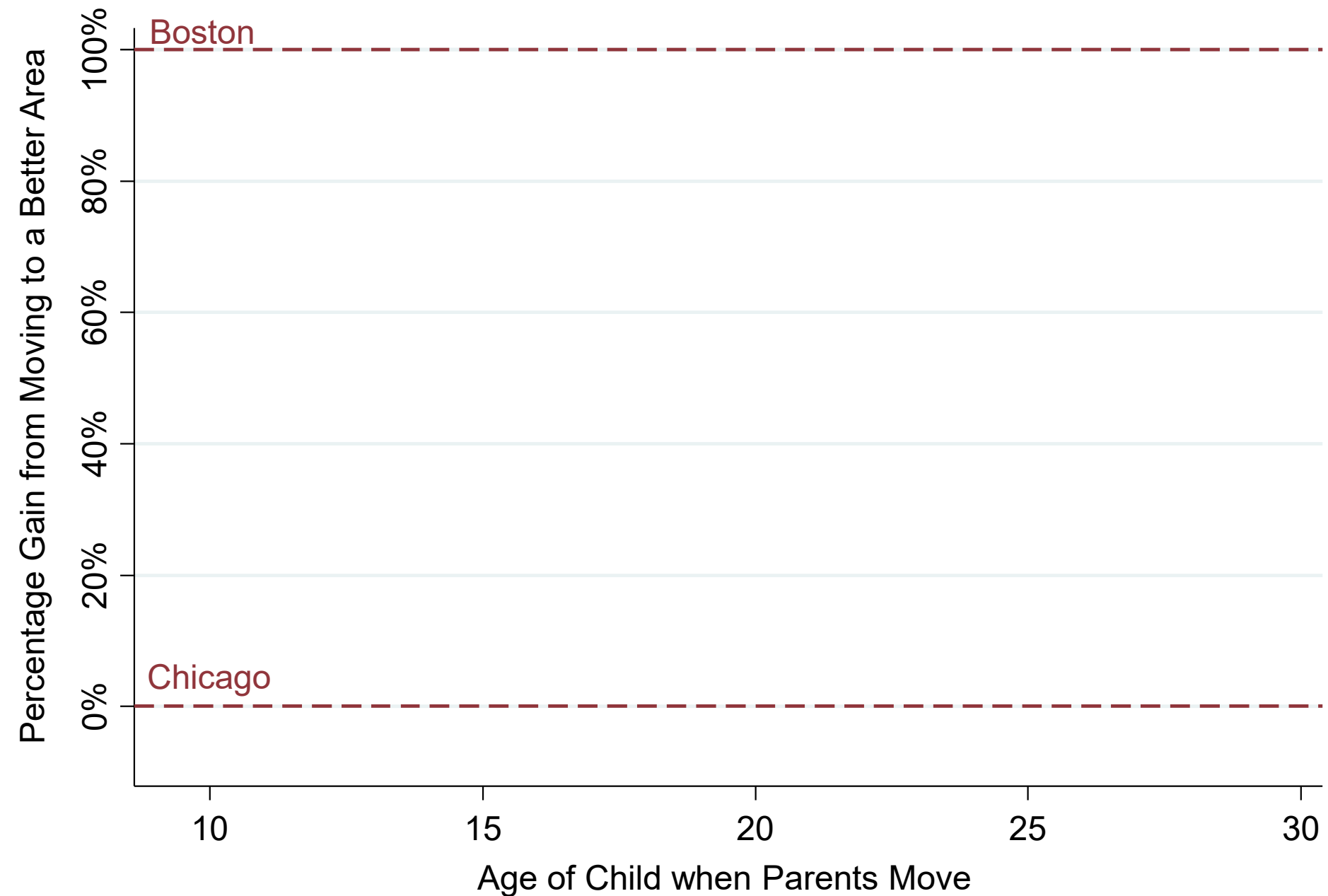
# Causal Effects of Place vs. Sorting

- Two very different explanations for variation in children's outcomes across areas:
  1. Heterogeneity: different people live in different places
  2. Neighborhood effects: places have a *causal* effect on upward mobility for a given person

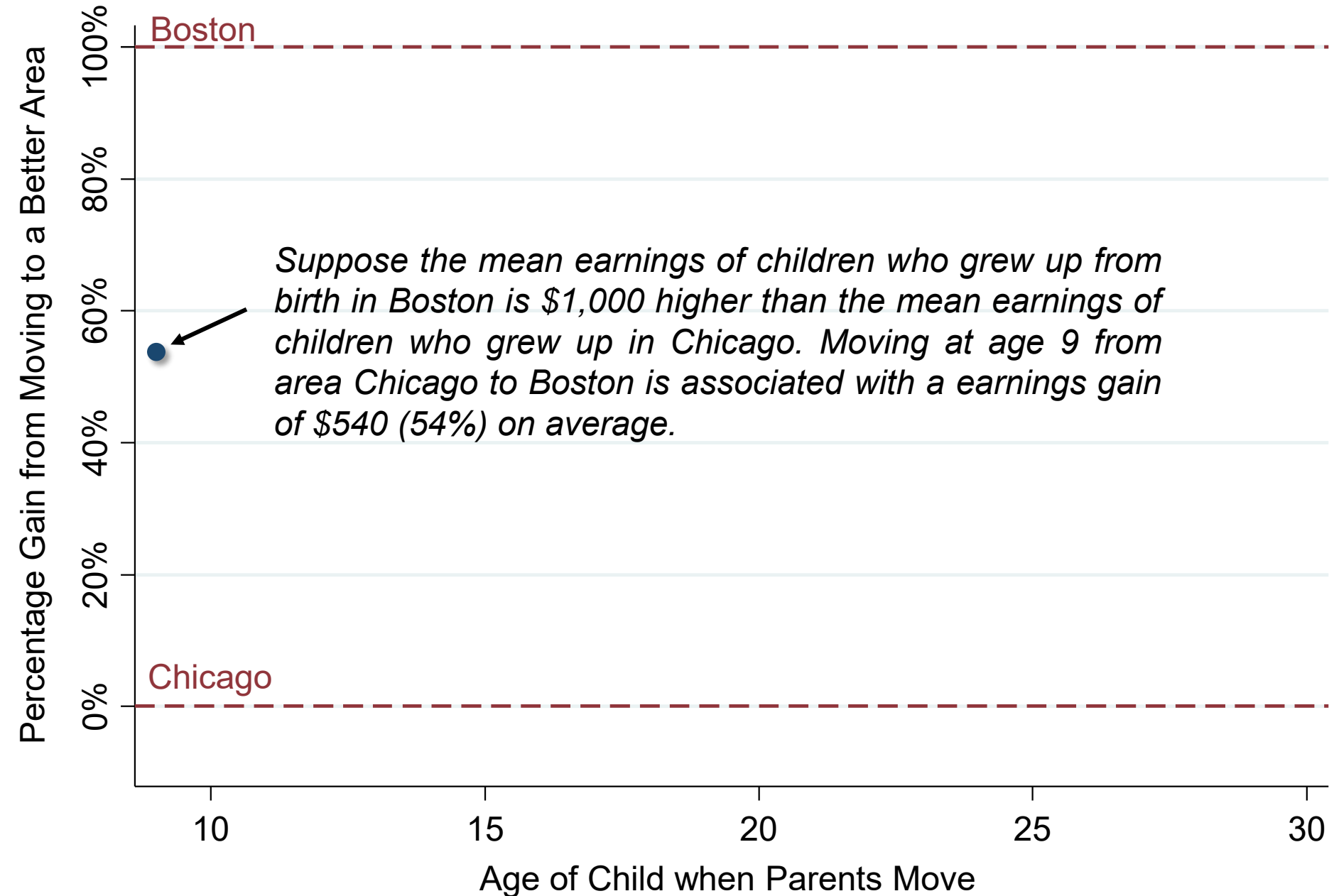
# Causal Effects of Place vs. Sorting

- Ideal experiment: randomly assign children to neighborhoods and compare outcomes in adulthood
- We approximate this experiment using a quasi-experimental design [Chetty and Hendren 2015]
  - Study families who move across areas with children of different ages in observational data

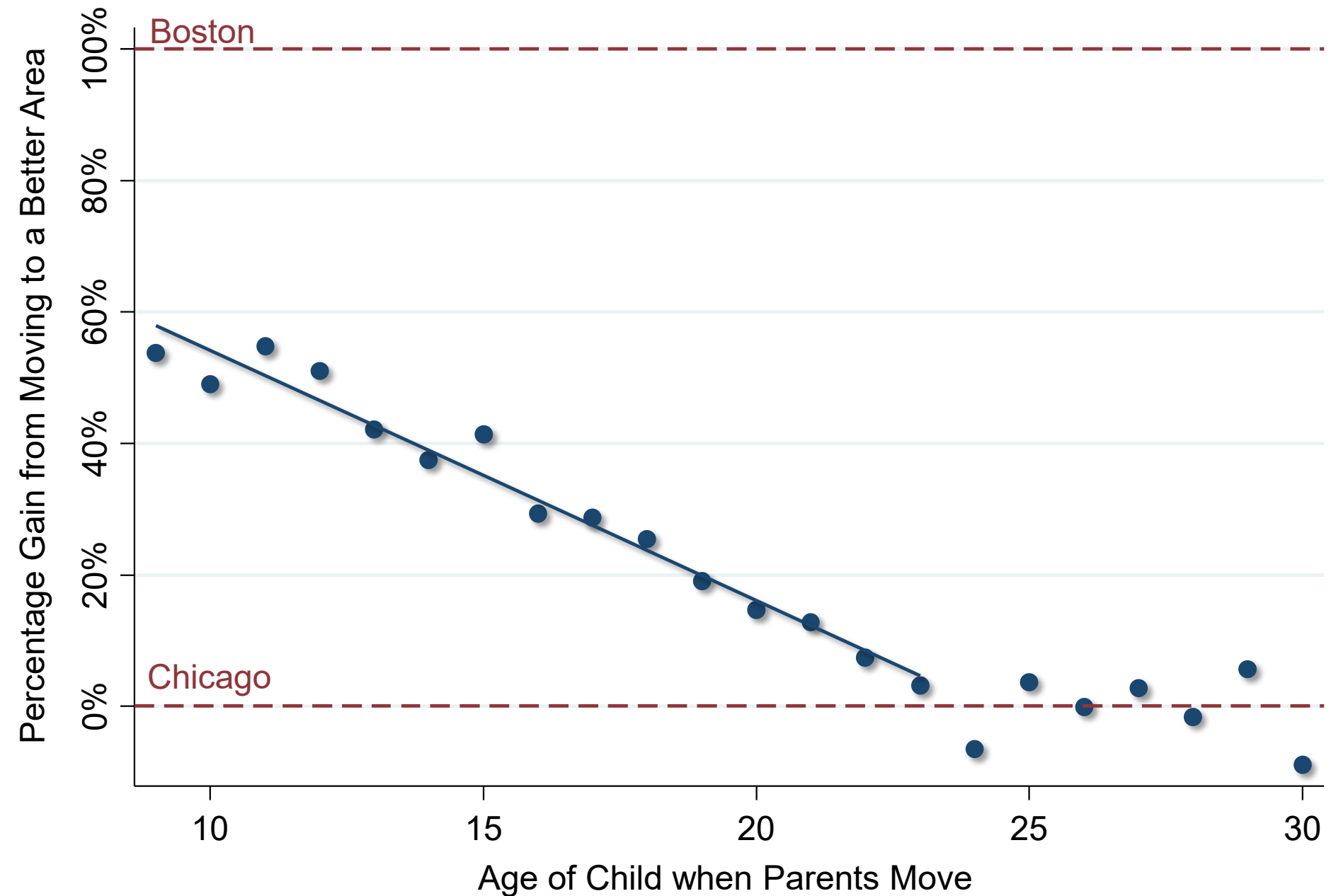
# Effects of Moving to a Different Neighborhood on a Child's Income in Adulthood by Age at Move



## Effects of Moving to a Different Neighborhood on a Child's Income in Adulthood by Age at Move



# Effects of Moving to a Different Neighborhood on a Child's Income in Adulthood by Age at Move



# Identifying Causal Exposure Effects

- Key assumption underlying quasi-experimental design: selection effect does not vary with child's age at move
  
- This assumption could be violated through two channels:
  1. Parents who move to good areas when their children are young may invest more in their children in other ways
  2. Moving may be correlated with other factors (e.g. change in parent income) that affect children directly



# Identifying Causal Exposure Effects

- Address these concerns using two approaches:
  1. Sibling comparisons: replicate baseline analysis with family fixed effects
    - When a family moves to a better area, we find that younger sibling does better than older sibling on avg., in proportion to age gap
  2. Placebo tests exploiting heterogeneity across subgroups
    - Some areas produce better outcomes for boys than girls
    - When a family moves to an area that produces better outcomes for boys but not girls, son's outcomes improve but daughter's do not

# Two Policy Approaches to Improving Upward Mobility

- Importance of place for mobility suggests two types of policy approaches:
  1. Help people move to better areas
  2. Invest in places with low levels of opportunity to replicate successes of areas with high upward mobility

# Policy Approach 1: Moving to Opportunity

- One way to help low-income families move to better neighborhoods: housing vouchers
- HUD Moving to Opportunity Experiment: gave such vouchers using a randomized lottery [Ludwig et al. 2013]
  - 4,600 families in Boston, New York, LA, Chicago, and Baltimore in mid 1990's

# Most Common MTO Residential Locations in New York



# MTO Experiment: Exposure Effects?

- Prior research on MTO (including HUD's Final Impacts Evaluation) found no economic gains from moving
  - But that work focused on adults and older youth at point of move
- We analyze long-term impacts of MTO on children who moved when young by linking MTO data to tax data

# Moving to Opportunity Experiment

- Children who moved to low-poverty areas when young (e.g., below age 13) do much better as adults:
  - 30% higher earnings = \$100,000 gain over life in present value
  - 27% more likely to attend college
  - 30% less likely to become single parents
- But moving had little effect on the outcomes of children who were already teenagers
- Moving also had no effect on parents' earnings
- Reinforces conclusion that *childhood exposure* is a key determinant of upward mobility

## Policy Approach 2: Improving Neighborhoods

- Limits to scalability of policies that move people
  - Also need policies that improve existing neighborhoods
- Challenging to identify causal effects of local policies
  - But we can characterize the features of areas that generate good outcomes

# **What are the Characteristics of High-Mobility Areas?**

## **Five Strongest Correlates of Upward Mobility**

### **1. Segregation**

- Racial and income segregation associated with less mobility
- Long commute times (sprawl) associated with less mobility



# **What are the Characteristics of High-Mobility Areas?**

## **Five Strongest Correlates of Upward Mobility**

1. Segregation
2. Income Inequality
  - Places with smaller middle class have much less mobility

# What are the Characteristics of High-Mobility Areas?

## Five Strongest Correlates of Upward Mobility

1. Segregation
2. Income Inequality
3. Family Structure
  - Areas with more single parents have much lower mobility
  - Strong correlation even for kids whose *own* parents are married

# **What are the Characteristics of High-Mobility Areas?**

## **Five Strongest Correlates of Upward Mobility**

1. Segregation
2. Income Inequality
3. Family Structure
4. Social Capital
  - “It takes a village to raise a child”
  - Putnam (1995): “Bowling Alone”

# **What are the Characteristics of High-Mobility Areas?**

## **Five Strongest Correlates of Upward Mobility**

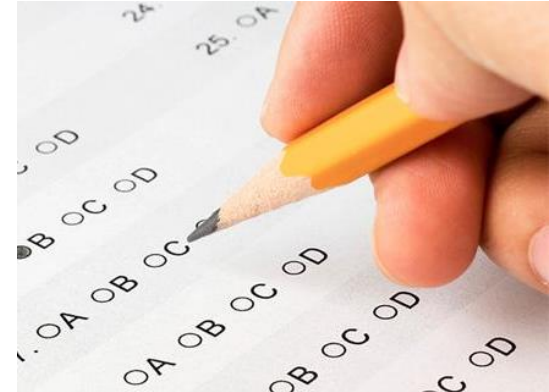
1. Segregation
2. Income Inequality
3. Family Structure
4. Social Capital
5. School Quality
  - Greater expenditure, smaller classes, higher test scores correlated with more mobility
  - Clear evidence of causal effects here

# Using Administrative Data to Study Teachers' Impacts

## School district records

2.5 million children

18 million test scores



## Tax records

Earnings, College

Attendance, Teen Birth



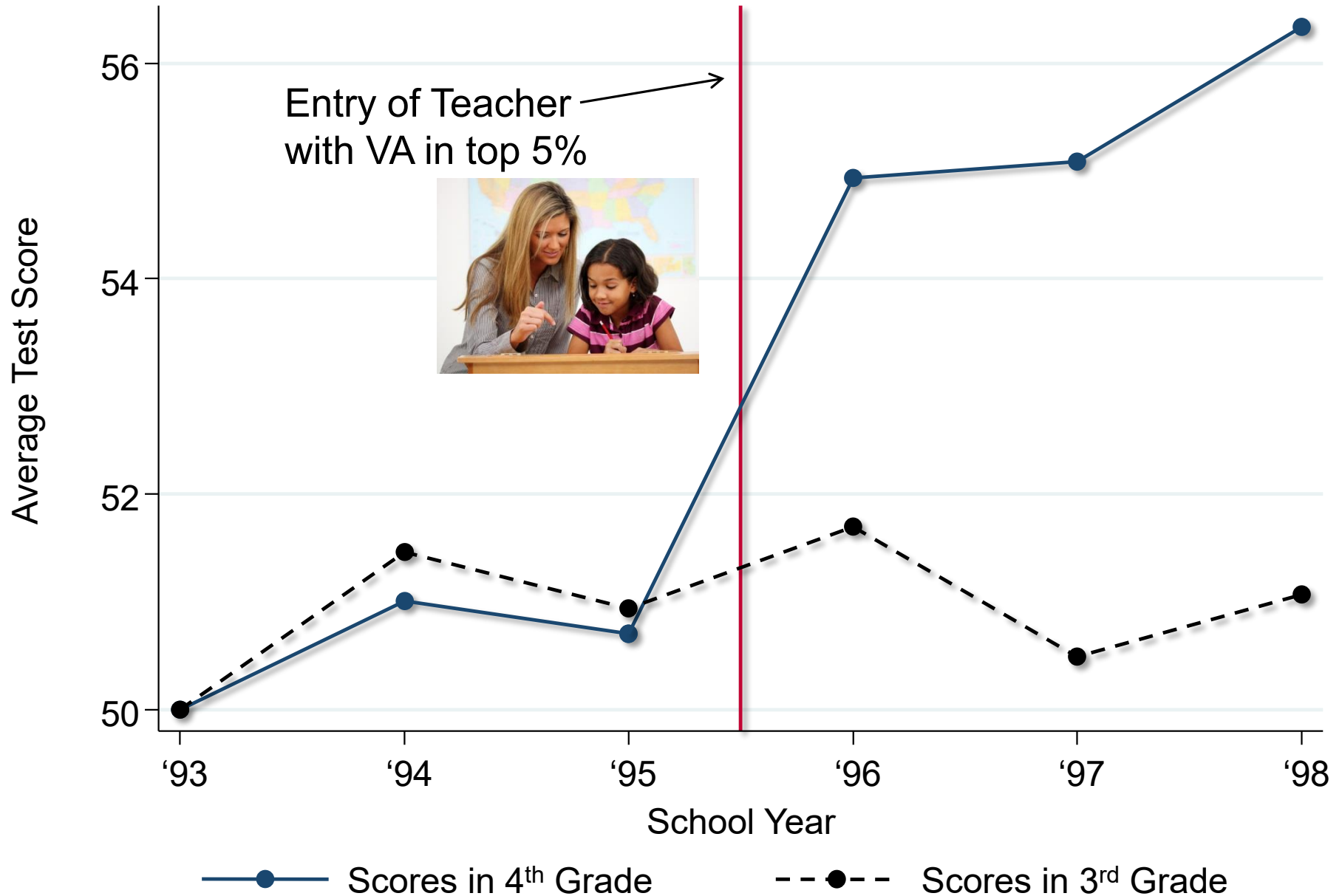
# Measuring Teacher Quality: Test-Score Based Metrics

One prominent measure  
of teacher quality:  
teacher *value-added*

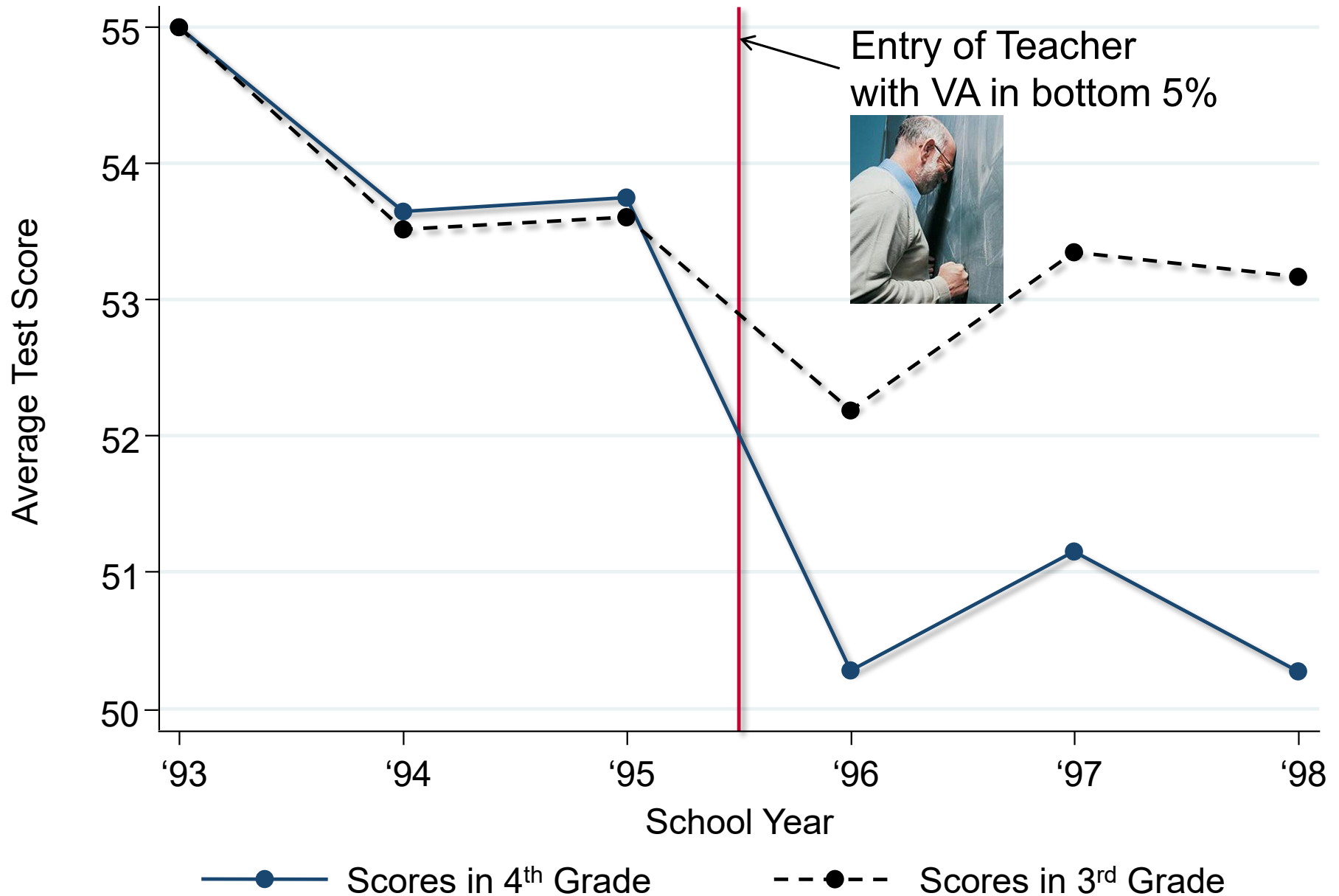
How much does a  
teacher raise her/his  
students' test scores  
on average?



## A Quasi-Experiment: Entry of High Value-Added Teacher

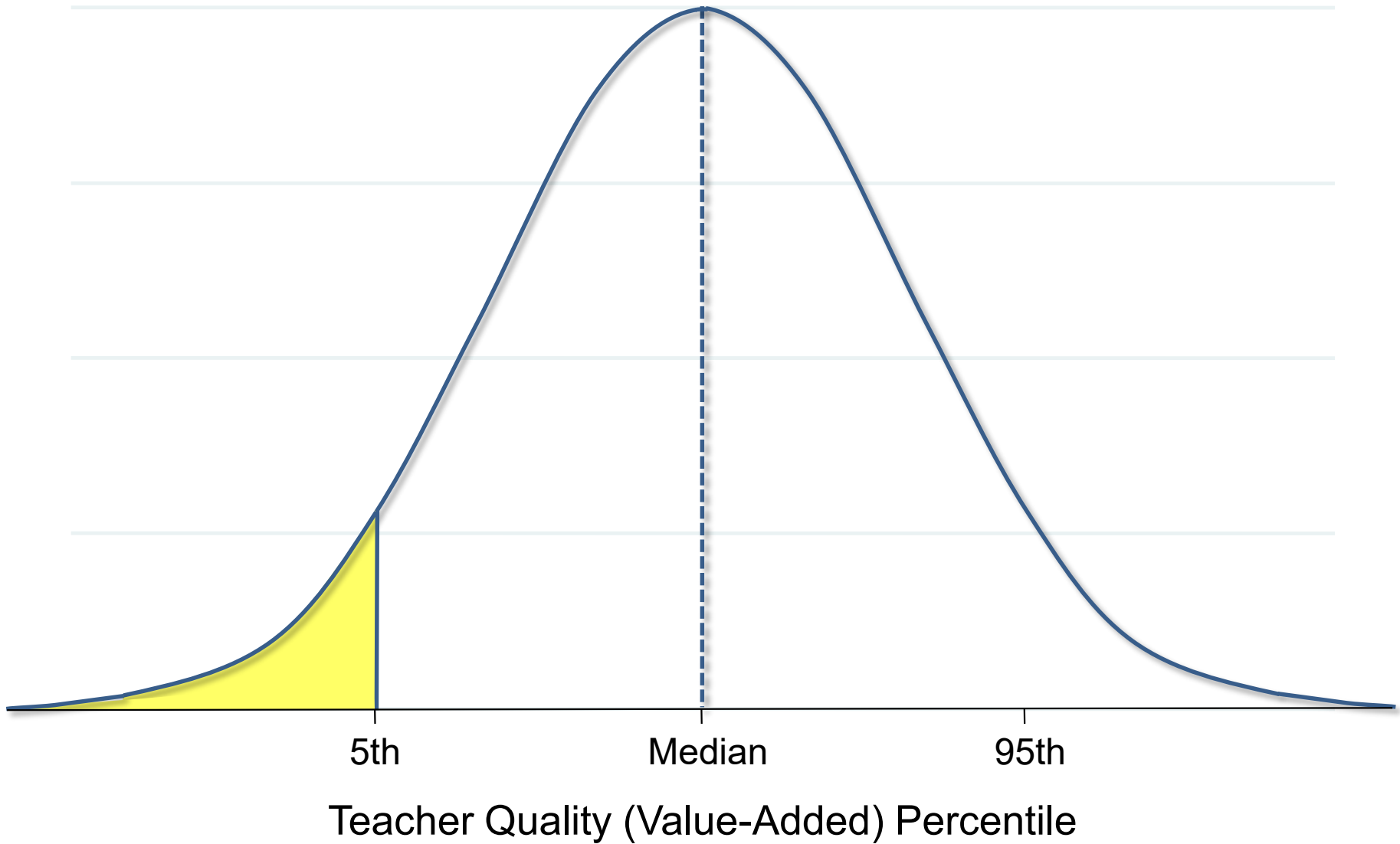


## A Quasi-Experiment: Entry of Low Value-Added Teacher





# The Value of Improving Teacher Quality



# The Value of Improving Teacher Quality

**+\$50,000** lifetime earnings per child  
**= \$1.4 million** per classroom of 28 students  
**= \$250,000** in present value at 5% int. rate

5th

Median

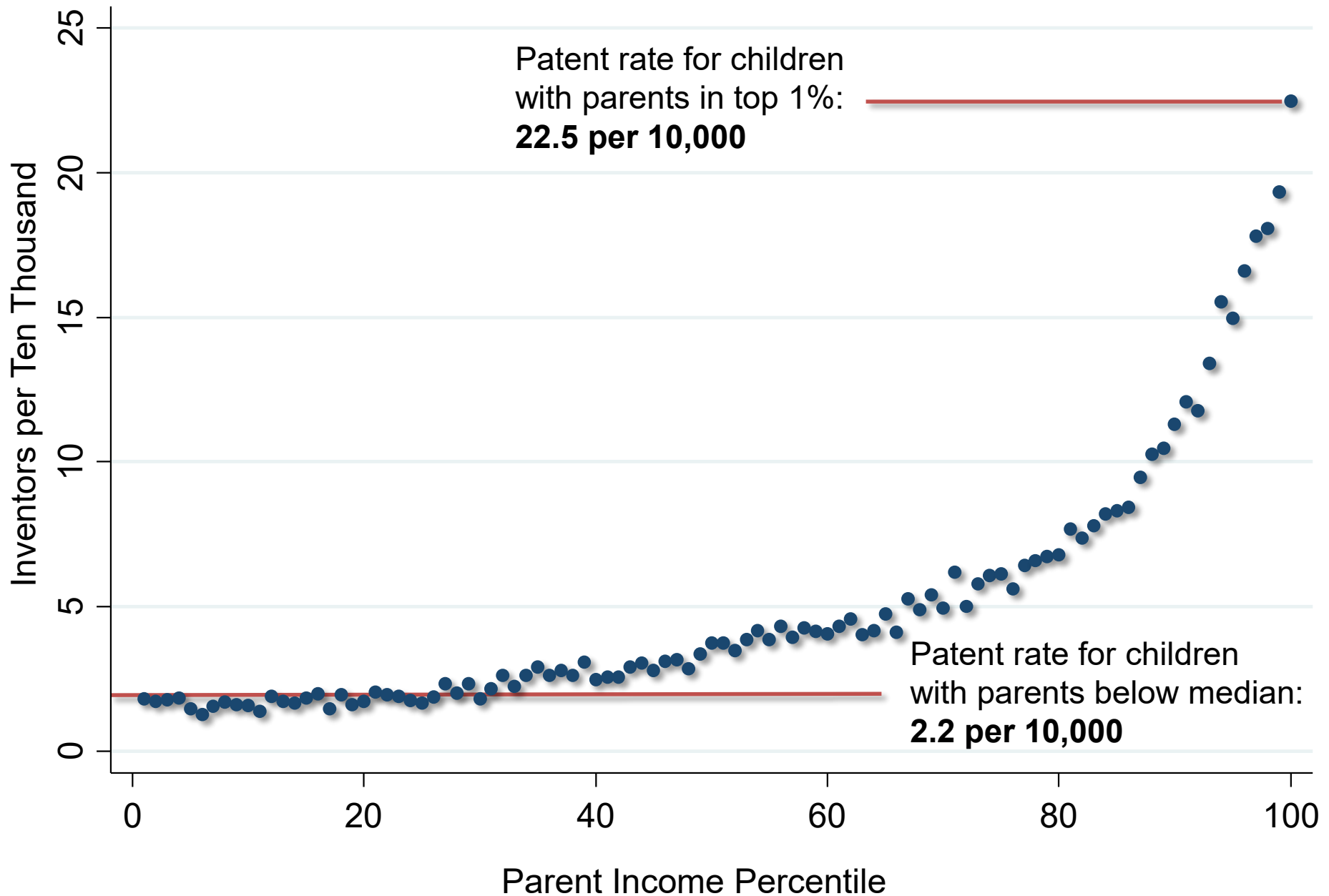
95th

Teacher Quality (Value-Added) Percentile

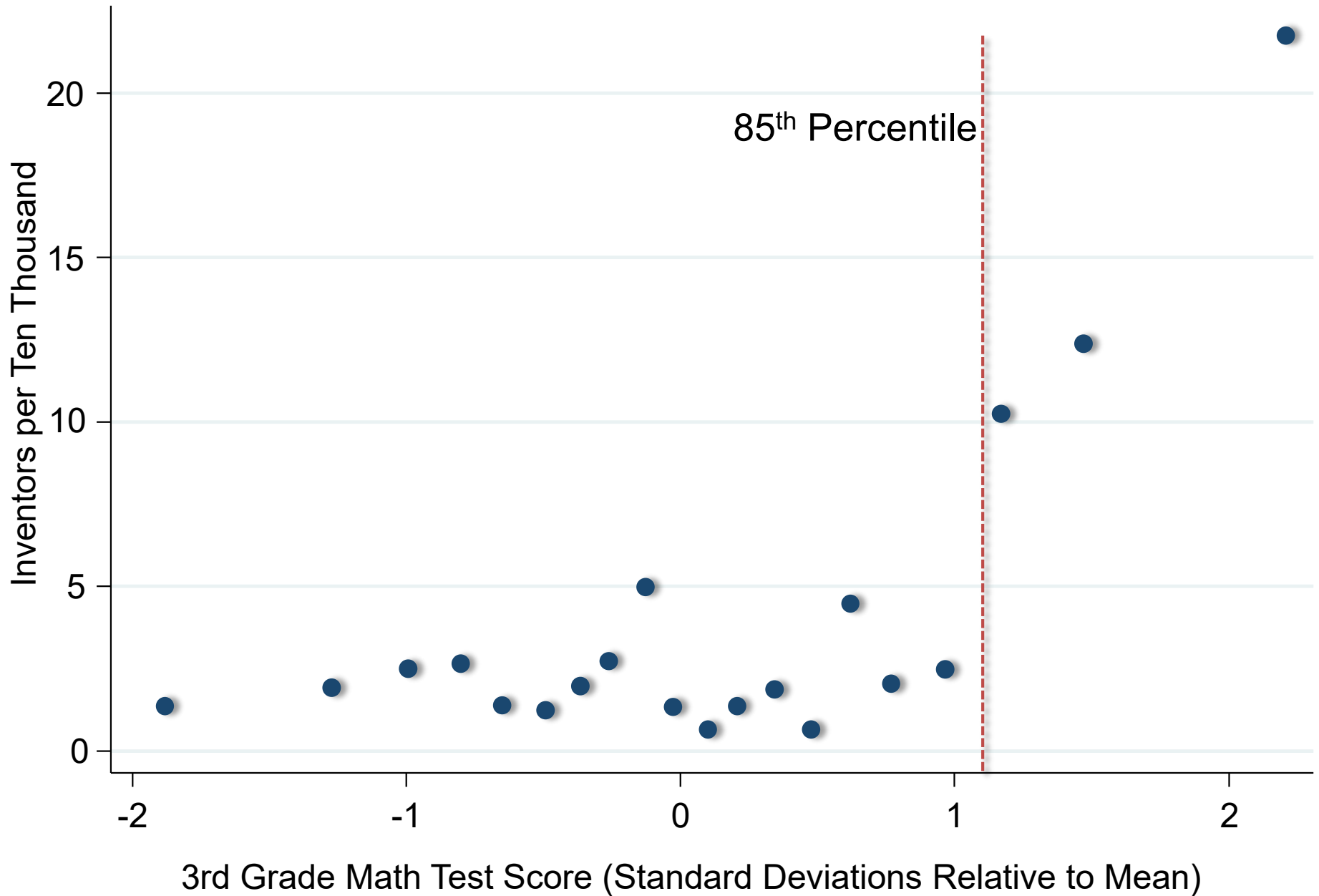
# Equality of Opportunity and Economic Growth

- Traditional argument for greater social mobility is based on principles of justice
- But improving opportunities for upward mobility can also increase size of the economic pie
  - One child's success need not come at another's expense
- To illustrate, focus on innovation
  - Study the lives of 750,000 patent holders in the U.S. by linking universe of patent data to tax records

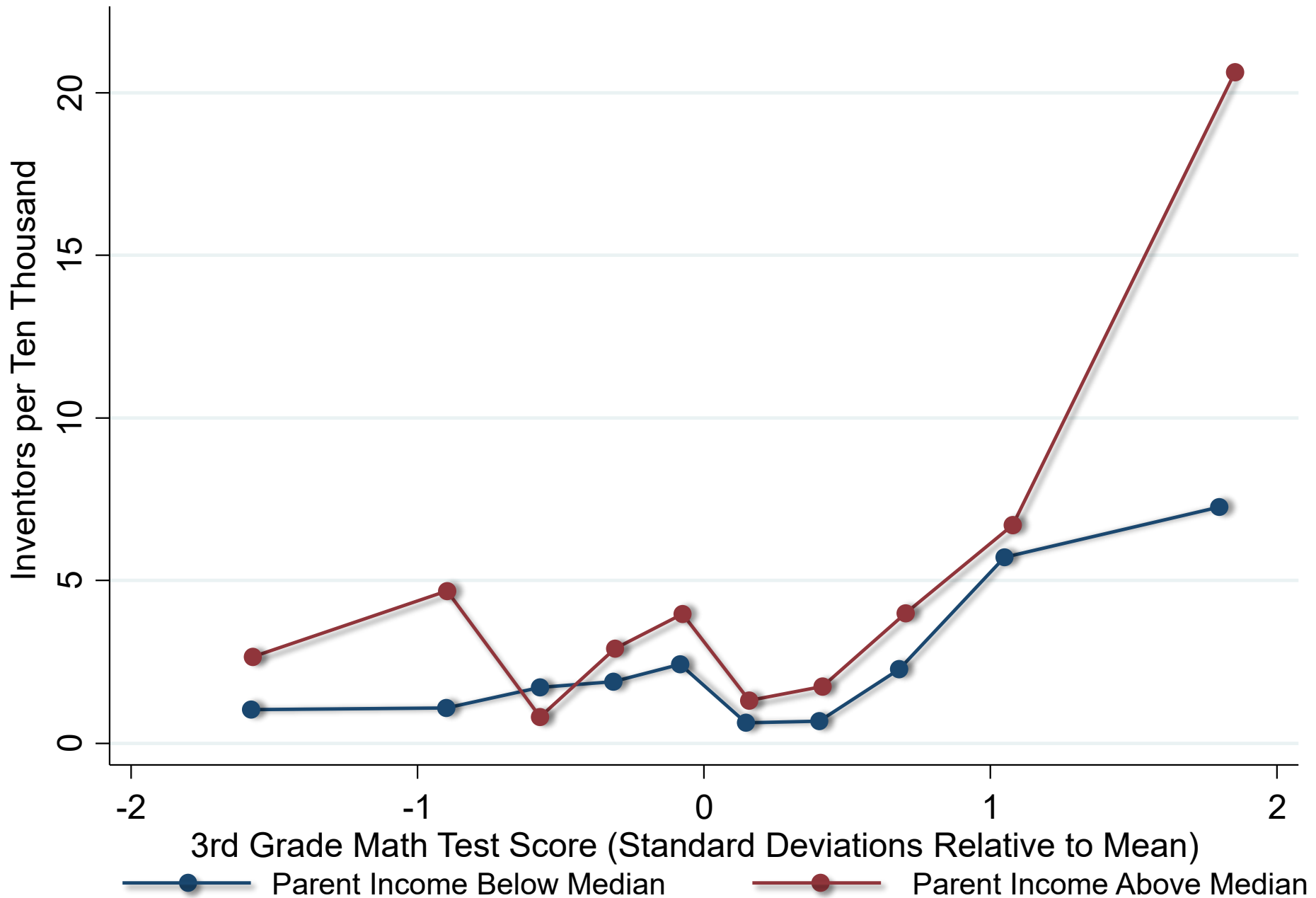
## Patent Rates vs. Parent Income Percentile



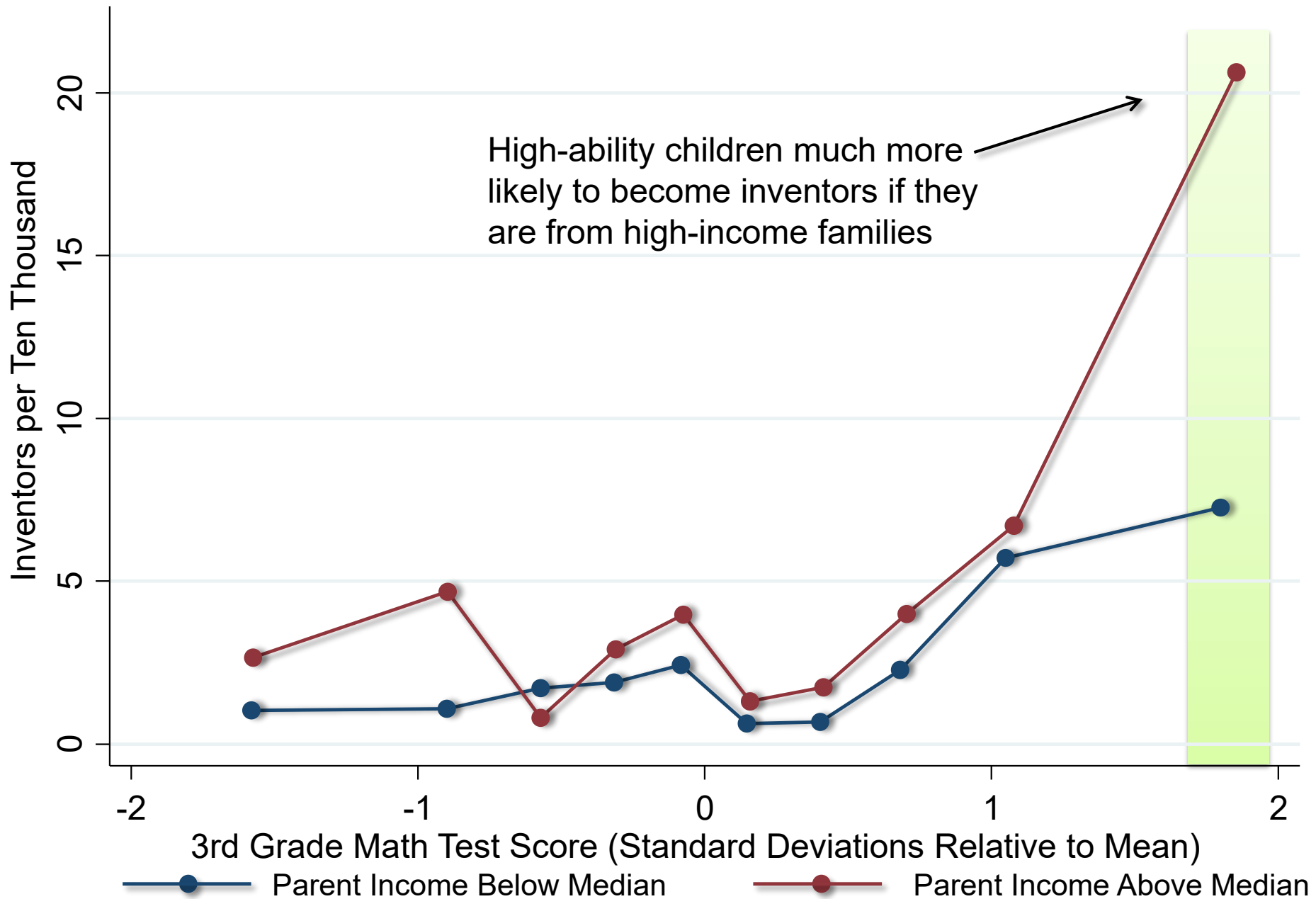
## Patent Rates vs. 3<sup>rd</sup> Grade Test Scores



# Patent Rates vs. 3<sup>rd</sup> Grade Test Scores for Children with Low vs. High Income Parents



# Patent Rates vs. 3<sup>rd</sup> Grade Test Scores for Children with Low vs. High Income Parents



# Policy Lessons

1. Improve childhood environments and primary education
  - Not just spending more money: US already spends more than other developed countries with better outcomes
  - Instead, focus on key inputs such as attracting and retaining talented teachers (e.g., Finland)



# Policy Impacts



“We know a good teacher can increase the lifetime income of a classroom by over \$250,000.... Every person in this chamber can point to a teacher who changed the trajectory of their lives”

- Barack Obama, *State of the Union*, 2012



“A recent study by Harvard and Columbia economists found that students with effective teachers are less likely to become pregnant, more likely to go to college and more likely to get higher-paying jobs....Ineffective teachers are hurting our students’ futures – we can’t allow that.”

- Michael Bloomberg, *State of the City*, 2012

# Policy Impacts

## *Vergara v. California* | Legal Claims

---

### *Civil Rights*

- Under longstanding California Supreme Court precedents, Plaintiffs have a fundamental right to equal educational opportunity.

---

### *Harm to Students*

- **Teacher quality is the key determinant of educational effectiveness and has a profound impact on students' lifetime achievement.**
- **The problem is worse for students who attend schools that serve predominantly minority and lower-income populations** because those schools are staffed by a disproportionate share of grossly ineffective teachers.
- In some school districts, **students of color are two to three times more likely to have bottom-quartile teachers than their white and Asian peers.**

# Policy Lessons

1. Improve childhood environments and primary education
2. Tackle social mobility at a local, not national level
  - Focus on specific cities such as Baltimore and Chicago, and on specific neighborhoods within those cities
  - Help families with young children move to high opportunity areas using housing vouchers or tax credits
    - Working with HUD to develop ways to make Section 8 voucher program more effective in achieving this goal
  - Invest in improving neighborhoods with low mobility

# Policy Lessons

1. Improve childhood environments and primary education
2. Tackle social mobility at a local, not national level
3. Harness administrative data to develop a scientific evidence base for economic and social policy
  - Identify which neighborhoods are in greatest need of improvement and which policies work
  - County-level data on mobility publicly available at [www.equality-of-opportunity.org](http://www.equality-of-opportunity.org)

# Download County-Level Data on Social Mobility in the U.S.

[www.equality-of-opportunity.org/data](http://www.equality-of-opportunity.org/data)

[HOME](#)[EXECUTIVE  
SUMMARY](#)[PAPER](#)[CITY  
RANKINGS](#)[INTERACTIVE  
MAP](#)[DOWNLOAD  
DATA](#)[FAQ'S](#)[RESEARCH  
TEAM](#)[IN THE NEWS](#)

## Downloadable Data on Intergenerational Mobility

### Data Description

Preferred Mobility Measures by Commuting Zone	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 1: National 100 by 100 Transition Matrix	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 2: Marginal Income Distributions by Centile	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 3: Intergenerational Mobility Statistics and Selected Covariates by County	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 4: Intergenerational Mobility Statistics by Metropolitan Statistical Area	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 5: Intergenerational Mobility Statistics by Commuting Zone	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 6: Quintile-Quintile Transition Matrices by Commuting Zone	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 7: Income Distributions by Commuting Zone	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 8: Commuting Zone Characteristics	<a href="#">Stata file</a>	<a href="#">Excel file</a>
Online Data Table 9: Commuting Zone Characteristics Definitions and Data Sources		<a href="#">Excel file</a>
Geographic Crosswalks (Tolbert and Sizer 1996, Autor and Dorn 2009 & 2013)	<a href="#">Zip file</a>	
Replication Stata Code and Datasets	<a href="#">Zip file</a>	

[Downloadable Map of Absolute Upward Mobility](#)

Version 2.0, released January 17, 2014. For Version 1.0 (released on July 22, 2013), click [here](#). Version 2.0 reports statistics using the 1980-82 birth cohorts (rather than 1980-81) and includes new data such as mobility statistics by county and MSA, new CZ-level covariates, and marginal income distributions for parents and children.

For more information on the data, please email [info@equality-of-opportunity.org](mailto:info@equality-of-opportunity.org)