Past and Present Differences in Opportunity by Neighborhood

Liana Fox
US Census Bureau

Catherine Massey
Welch Consulting and University of Michigan

Jonathan Rothbaum
US Census Bureau

Disclaimer: This work was made possible by the US 2050 project, supported by the Peter G. Peterson Foundation and the Ford Foundation. The statements made and views expressed are solely the responsibility of the authors. This report is released to inform interested parties of ongoing research and to encourage discussion. Any views expressed on statistical, methodological, technical, or operational issues are those of the authors and not necessarily those of the US Census Bureau.

Abstract

Recent research finds that childhood neighborhoods affect adult economic outcomes, especially for children of low-income parents. However, understanding why one neighborhood results in better outcomes for low-income children than another is extremely challenging using estimates from only one point in time. Because places are shaped by both contemporary and historical factors, it is important to understand geographic differences in opportunity both today and in the past. Using 1940 Census data linked to 1040 tax returns, we examine geographic differences in child outcomes experienced by cohorts born roughly 50 years apart – revealing how intergenerational persistence of status has changed over time both at the national level and at smaller geographic levels. In studying these changes, we hope to shed light on the causes of intergenerational mobility and inequality of opportunity.

Motivation

- Substantial heterogeneity in mobility by neighborhood exists for US cohorts of born between 1980 and 1986 (Chetty et al. 2016, Chetty et al. 2018)
- Regional differences in opportunity may change over time as neighborhoods are shaped by both contemporary and historical factors
- Place-based opportunity may trend differently for minorities relative to whites
- Difficult to evaluate place-based policies that affect children with only cross section of child outcomes by place

3

Research Questions

- For older cohorts, what is the association (and causal relationship) between place and their adult outcomes?
- How has place-based mobility changed since the early 20th century?
- Has place-based mobility evolved differently for different groups (blacks vs. whites, for example)?

4

Data

- 1940 Census
- 2000 Census
- 2001-2015 American Community Survey
- 1974, 1979, 1984, 1989, 1994, 1995, 1998-2015 Internal Revenue Service (IRS) 1040 Tax Returns
  - Location and income information for filers
- Social Security Administration Numident File

5

Data Linkage – Person Identification Validation System (PVS)

- PVS assigns 9 digit, unique identifiers called Protected Identification Keys (PIKs) via probabilistic matching techniques to surveys and decennial data
- PIKs are used to facilitate removing duplicates and record linkage
- Once 'PIKed,' data can be linked to any other data processed through PVS

6

Data Linkage Example 1922-1940 Cohort

1940 Census

PIK	First Name	Last Name	Relation	Occupation	Wages	Age	Birth State
111-11-111	Abraham	Simpson	Head	Security Guard	4,000	26	New York
222-22-222	Mona	Simpson	Spouse		0	25	New York
333-33-333	Herb	Simpson	Child			7	New York
444-44-444	Homer	Simpson	Child			1	Springfield

1974 IRS 1040

PIK	First Name	Last Name	Relation	Adjusted Gross Income
444-44-444	Homer	Simpson	Head	2,000
555-55-555	Marge	Simpson	Spouse	

Numident

PIK	First Name	Last Name	Birth City	Birth State
444-44-444	Homer	Simpson	Springfield	Springfield

2000 Longform Census

PIK	First Name	Last Name	Relation	Education
444-44-444	Homer	Simpson	Head	12 Years
555-55-555	Marge	Simpson	Spouse	16 Years

Association between Parent and Child Status Over Time

Regression of Child Status on Parent Status

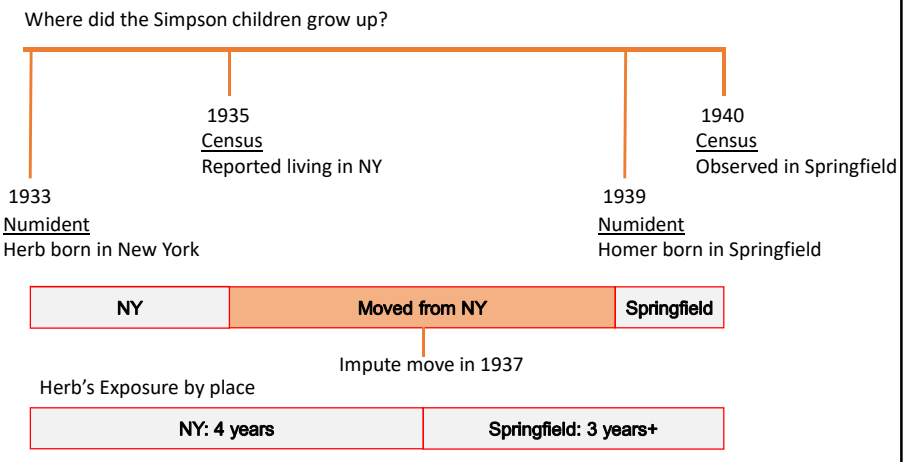
	Education (Years of Schooling)		Earnings/Income Rank	
	1940 Cohort	LF and ACS Cohort	1940 Cohort	LF and ACS Cohort
Slope	0.3372	0.4017	0.2351	0.3289
	(0.001)	(0.002)	(0.000)	(0.001)
Intercept	9.163	8.736	38.69	33.55
	(0.005)	(0.002)	(0.013)	(0.030)
Observations	2,130,000	218,000	19,110,000	5,180,000
R^2	0.16	0.15	0.06	0.11

Method

- Calculate outcomes for children born in each commuting zone following work by Chetty and Hendren (2015)
  - Rank-rank slope and intercept and expected outcome (for below- and above-median children) by location
- Replicate Chetty and Hendren's (2018) causal estimation of neighborhoods on intergenerational gaps
  - Limit sample to children who moved once during childhood
  - Regress income rank of children who move on origin characteristics, destination characteristics, and parental income rank all interacted with age-at-move fixed effects

9

Location in Historical Sample



10

Causal Estimate

	Birth-Move Estimates		Chetty and Hendren Estimate
	1940 Cohort	2000 Longform/ ACS Cohort*	
Causal Effect (Per Year)	0.016 (0.003)	0.030 (0.011)	0.040 (0.002)
Observations (1-Time movers in sample)	483,000	195,400	1,553,000

Challenges in 1922-1940 Cohort

- 1940 measures of status (asked of full-count census!)
  - Earnings
    - Wage and salary
    - Self-employment – a dummy variable for >= $50
      - Relatively common – weighted by child: 9% of mothers, 33% of fathers
      - Substantial income missing in data
    - Expect some attenuation of parent-child relationship as parent status is measured with error (life-cycle bias, measurement error in earnings, transitory shocks, etc.)
  - Education
    - Years of schooling
- Child linkage – 70 percent of children are linked
  - Non-random – observables such as race are associated with linkage probability

12

Solutions

- Inverse Probability Weights
  - Several Weights
    - No weights (full-count census represents all children, weight of 1 to each)
    - IPW – regress dummy for group on family and parent characteristics as well as geographic summary variables at county, CZ, and state level
      - Child PIK IPW (Dummy = is child PIKed)
      - Parent Earnings IPW (Dummy = parent earnings > 0)
      - Both (Dummy = parent earnings > 0 | is child PIKed = 1)
  - Calculate any summary stat/regression coefficient for all samples
  - Evaluate impact of different weights on results, especially for local mobility statistics

13

Next Steps

- Concuct analysis separately for blacks and whites
- Evaluate CZ changes that are associated with mobility changes
- Predict causal effects by place in 1940 (as in Chetty and Hendren, 2018)

14