Fragile Families
Opportunity Insights Data
on County Intergenerational Mobility
Restricted Use Data Appendage

Baseline, Year 1, Year 3, Year 5, and
Year 9 Follow-Up Waves

January 2019

Bendheim-Thoman
Center for Research on Child Wellbeing
Wallace Hall
Princeton University
Princeton, NJ 08544
http://crcw.princeton.edu

Columbia Population Research Center
1255 Amsterdam Avenue, Room 715
Columbia University
New York, NY 10027
http://cupop.columbia.edu/

Prepared by staff in the Bendheim-Thoman Center for Research on Child Wellbeing (CRCW), Princeton University. For more information about Fragile Families, please visit our web site at http://fragilefamilies.princeton.edu/ or email ffdata@princeton.edu

DATA APPENDAGE OVERVIEW

The Fragile Families Opportunity Insights Data on County Intergenerational Mobility Data Appendage contains county level data on intergenerational mobility and its theorized correlates. County data from Opportunity Insights was merged to the county residence of the focal child's mother and father at Baseline, Year 1, Year 3, Year 5, and Year 9. All information on counties was obtained from data made publically available by Raj Chetty and Nathaniel Hendren (2017) at Opportunity Insights. Chetty and Hendren used de-identified tax records on more than five million children between 1996 and 2012 to estimate the effect of growing up in each U.S. county on children's future prospects for upward mobility. In this data appendage, we include Chetty and Hendren's preferred estimates of causal place effects by county for adult household income, individual income, and probability of marriage. These county measures are also disaggregated by family income (low- and high-income) and child sex (male and female). In addition, the data appendage includes the 48 county level covariates theorized by Chetty and Hendren to be correlated with intergenerational mobility.

For more detailed information about Chetty and Hendren's *The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates* paper and data see:

- https://opportunityinsights.org/paper/neighborhoodsii/

FILE LAYOUT

The file contains 4,898 observations (one per family) and is sorted by *idnum.*

VARIABLE NAMING CONVENTION

The variable names are constructed as follows:

Position	Character	Indicates
1	r	Restricted data
2	g	Geographical measure
3	m	Mother residence
	f	Father residence
4	1	First wave (Baseline)
	2	Second wave (Year 1)
	3	Third wave (Year 3)
	4	Fourth wave (Year 5)
	5	Fifth wave (Year 9)
5-9	opin_	Opportunity Insights source data
10	c	County-level geographical measure
11-	*	Description of variable

For example, variable **rgm1opin_cimp25_kr26** refers to a restricted geographical variable corresponding to the mothers' residence at Baseline, sourced from Opportunity Insights, at the county level, representing the percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the 25th percentile of the national household income distribution.

MISSING VALUES

Unlike other FFCWS data appendages, variables with county level data in this appendage do not include negative values to indicate missing values. The negative values for county level variables in this appendage correspond to real negative values. For example, a negative value for intergenerational mobility corresponds to a loss in income (see description of variables for more information). For Stata users, the missing data is coded using system missing variable notation.

Value	Label	Indicates
.n	Not in wave	Focal child's mother/father did not participate in FFCWS survey (usually coded as -9 Not in wave).
.c	County unknown	Focal child's mother/father participated in FFCWS survey but county residence at wave is unknown.
.m	Missing	Focal child's mother/father participated in FFCWS survey and county residence at wave is known, but data is not available for the county of residence.

COUNTY MEASURES OF INTERGENERATIONAL MOBILITY

All measures of county level intergenerational mobility come from Chetty and Hendren's Online Data Table II: Preferred Estimates of Causal Place Effects by County (2017). The authors describe these data as follows:

> *[The first 16 variables in] this table list the percentage gain (or loss) in income at age 26 from spending one more year of childhood in a given county relative to the national mean. We report estimates for children with parents at the 25th and 75th percentile of the national income distribution for all children, girls, boys. We also show the average of the girls and boys estimates (gender average), which is not identical to the estimate for all children because the statistical model does not impose the requirement that the overall effect has to be the average of the gender-specific effects. We report estimates for two different income definitions: individual (own) income and household income (including spousal income). [The next 2 variables] show the percentage point change in the probability of being married at age 26 from spending one more year of childhood in a given county. As an example, spending one more year of childhood in Lawrence County, AL increases household income by 0.23% and increases the probability of being married at age 26 by 0.39 percentage points relative to the national mean for a child with parents at the 25th percentile.*

Using this language, the following county level measures of intergenerational mobility available in this data appendage can be interpreted as follows.

County effect on household income

County effect on household income: Child from low-income family

Variable: rg*opin_cimp25_kr26

Description: The percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the 25th percentile of the national household income distribution.

County effect on household income: Child from high-income family

Variable: rg*opin_cimp75_kr26

Description: The percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the 75th percentile of the national household income distribution.

County effect on household income: Female child from low-income family

Variable: rg*opin_cimp25_kr26_f

Description: The percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a female child with parents at the 25th percentile of the national household income distribution.

County effect on household income: Female child from high-income family

Variable: rg*opin_cimp75_kr26_f

Description: The percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a female child with parents at the 75th percentile of the national household income distribution.

County effect on household income: Male child from low-income family

Variable: rg*opin_cimp25_kr26_m

Description: The percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a male child with parents at the 25th percentile of the national household income distribution.

County effect on household income: Male child from high-income family

Variable: rg*opin_cimp75_kr26_m

Description: The percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a male child with parents at the 75th percentile of the national household income distribution.

County effect on household income: Child from low-income family (gender average)

Variable: rg*opin_cimp25_kr26_avg

Description: The percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the 25th percentile of the national household income distribution, calculated as the average of the gender-specific effects.

County effect on household income: Child from high-income family (gender average)

Variable: rg*opin_cimp75_kr26_avg

Description: The percentage gain (or loss) in household income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the <u>75th</u> percentile of the national household income distribution, calculated as the <u>average of the gender-specific effects</u>.

County effect on individual income

County effect on individual income: Child from low-income family

Variable: rg*opin_cimp25_kir26

Description: The percentage gain (or loss) in individual income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the <u>25th</u> percentile of the national household income distribution.

County effect on individual income: Child from high-income family

Variable: rg*opin_cimp75_kir26

Description: The percentage gain (or loss) in individual income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the <u>75th</u> percentile of the national household income distribution.

County effect on individual income: Female child from low-income family

Variable: rg*opin_cimp25_kir26_f

Description: The percentage gain (or loss) in individual income at age 26 from spending one more year of childhood in the given county relative to the national mean for a <u>female</u> child with parents at the <u>25th</u> percentile of the national household income distribution.

County effect on individual income: Female child from high-income family

Variable: rg*opin_cimp75_kir26_f

Description: The percentage gain (or loss) in individual income at age 26 from spending one more year of childhood in the given county relative to the national mean for a <u>female</u> child with parents at the <u>75th</u> percentile of the national household income distribution.

County effect on individual income: Male child from low-income family

Variable: rg*opin_cimp25_kir26_m

Description: The percentage gain (or loss) in individual income at age 26 from spending one more year of childhood in the given county relative to the national mean for a <u>male</u> child with parents at the <u>25th</u> percentile of the national household income distribution.

County effect on individual income: Male child from high-income family

Variable: rg*opin_cimp75_kir26_m

Description: The percentage gain (or loss) in individual income at age 26 from spending one more year of childhood in the given county relative to the national mean for a <u>male</u> child with parents at the <u>75th</u> percentile of the national household income distribution.

County effect on individual income: Child from low-income family (gender average)

Variable: rg*opin_cimp25_kir26_avg

Description: The percentage gain (or loss) in individual income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the 25th percentile of the national household income distribution, calculated as the average of the gender-specific effects.

County effect on individual income: Child from high-income family (gender average)

Variable: rg*opin_cimp75_kir26_avg

Description: The percentage gain (or loss) in individual income at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the 75th percentile of the national household income distribution, calculated as the average of the gender-specific effects.

County effect on marriage

County effect on marriage: Child from low-income family

Variable: rg*opin_cimp25_km26

Description: The percentage point gain (or loss) in the probability of marriage at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the 25th percentile of the national household income distribution.

County effect on marriage: Child from high-income family

Variable: rg*opin_cimp25_km26

Description: The percentage point gain (or loss) in the probability of marriage at age 26 from spending one more year of childhood in the given county relative to the national mean for a child with parents at the 75th percentile of the national household income distribution.

COUNTY MEASURES OF OTHER CONTEXTS

All other county level measures come from Chetty and Hendren's Online Data Table IV: County-Level Causal Place Effects and Covariates. The authors categorize and describe these data as follows:

Segregation and Poverty

Fraction Black

Description: Number of individuals who are black alone divided by total population

Source: 2000 Census SF1 100% Data Table P008

Poverty Rate

Description: Fraction of population below the poverty rate

Source: 2000 Census SF3 Sample Data Table P087

Racial Segregation

Variable: Multi-group Theil Index calculated at the census-tract level over four groups: White alone, Black alone, Hispanic, and Other

Source: 2000 Census SF1 100% Data Table P008

Income Segregation

Variable: Rank-Order index estimated at the census-tract level using equation (13) in Reardon (2011); the δ vector is given in Appendix A4 of Reardon's paper. $H(p_k)$ is computed for each of the income brackets given in the 2000 census.

Source: 2000 Census SF3 Sample Data Table P052

Segregation of Poverty (p75)

Description: $H(p25)$ estimated following Reardon (2011); $H(p)$ is computed for 16 income groups defined by the 2000 census; $H(p25)$ is computed using a fourth-order polynomial of the weighted linear regression in equation (12) of Reardon (2011)

Source: 2000 Census SF3 Sample Data Table P052

Segregation of Affluence (>p75)

Description: Same definition as segregation of poverty, but using p75 instead of p2

Source: 2000 Census SF3 Sample Data Table P052

Fraction with Commute < 15 Minutes

Description: Number of workers that commute less than 15 minutes to work divided by total number of workers. Sample restricts to workers that are 16 or older and not working at home.

Source: 2000 Census SF3 Sample Data Table P031

Logarithm of Population Density

Description: Logarithm of the Population Density where the Population Density is defined as the Population divided by the Land Area in square miles.

Source: 2000 Census Gazetteer Files

Income Distribution

Household Income per Capita

Description: Aggregate household income in the 2000 census divided by the number of people aged 16-64

Source: 2000 Census SF3 Sample Data Table P054

Gini Coefficient

Description: Gini coefficient computed using parents of children in the core sample, with income topcoded at $100 million in 2012 dollars

Source: Tax Records, Core Sample of Chetty et al. (2014)

Top 1% Income Share

Description: The fraction of income within a CZ going to the top 1% defined within the CZ, computed using parents of children in the core sample

Source: Tax Records, Core Sample of Chetty et al. (2014)

Gini Bottom 99%

Description: Gini coefficient minus top 1% income share

Source: Tax Records, Core Sample of Chetty et al. (2014)

Fraction Middle Class (between p25 and p75)

Description: Fraction of parents (in the core sample) whose income falls between the 25th and 75th percentile of the national parent income distribution

Source: Tax Records, Core Sample of Chetty et al. (2014)

Tax

Local Tax Rate

Description: Total tax revenue per capita divided by mean household income per capita for working age adults (in 1990)

Source: 1992 Census of Government county-level summaries

Local Tax Rate Per Capita

Description: Total tax revenue per capita

Source: 1992 Census of Government county-level summaries

Local Government Expenditures Per Capita

Description: Total local government expenditures per capita

Source: 1992 Census of Government county-level summaries

K-12 Education

School Expenditure per Student

Description: Average expenditures per student in public schools

Source: NCES CCD 1996-1997 Financial Survey

Student Teacher Ratio

Description: Average student-teacher ratio in public schools

Source: NCES CCD 1996-1997 Financial Survey

Test Score Percentile (Income adjusted)

Description: Residual from a regression of mean math and English standardized test scores on household income per capita in 2000

Source: George Bush Global Report Card

High School Dropout Rate (Income adjusted)

Description: Residual from a regression of high school dropout rates on household income per capita in 2000

Source: NCES CCD 2000-2001

College

Number of Colleges per Capita

Description: Number of Title IV, degree offering institutions per capita

Source: IPEDS 2000

College Tuition

Description: Mean in-state tuition and fees for first-time, full-time undergraduates

Source: IPEDS 2000

College Graduation Rate (Income Adjusted)

Description: Residual from a regression of graduation rate (the share of undergraduate students that complete their degree in 150% of normal time) on household income per capita in 2000

Source: IPEDS 2009

Local Labor Market

Labor Force Participation

Description: Share of people at least 16 years old that are in the labor force

Source: 2000 Census SF3 Sample Data Table P043

Share Working in Manufacturing

Description: Share of employed persons 16 and older working in manufacturing.

Source: 2000 Census SF3 Sample Data Table P049

Growth in Chinese Imports

Description: Percentage growth in imports from China per worker between 1990 and 2000

Source: Autor, Dorn, and Hanson (2013)

Teenage (14-16) Labor Force Participation

Description: Fraction of children in birth cohorts 1985-1987 who received a W2 (i.e. had positive wage earnings) in any of the tax years when they were age 14-16

Source: Tax Records, Extended Sample

Migration

Migration Inflow Rate

Description: Migration into the county from other county (divided by county population from 2000 Census)

Source: IRS Statistics of Income 2004-2005

Migration Outflow Rate

Description: Migration out of the county from other counties (divided by county population from 2000 Census)

Source: IRS Statistics of Income 2004-2005

Fraction Foreign Born

Description: Share of county residents born outside the United States

Source: 2000 Census SF3 Sample Data Table P021

Social Capital

Social Capital Index

Description: Standardized index combining measures of voter turnout rates, the fraction of people who return their census forms, and measures of participation in community organizations.

Source: Rupasingha and Goetz (2008)

Fraction Religious

Description: Share of religious adherents

Source: Association of Religion Data Archives

Violent Crime Rate

Description: Number of arrests for serious violent crimes per capita

Source: Uniform Crime Reports

Total Crime Rate

Description: Number of arrests for all crimes per capita

Source: Uniform Crime Reports

Family Structure

Fraction of Children with Single Mothers

Description: Number of single female households with children divided by total number of households with children

Source: 2000 Census SF3 Sample Data Table P015

Fraction of Adults Divorced

Description: Fraction of people 15 or older who are divorced

Source: 2000 Census SF3 Sample Data Table P018

Fraction of Adults Married

Description: Fraction of people 15 or older who are married and not separated

Source: 2000 Census SF3 Sample Data Table P018

Prices

Median Monthly Rent for Above-Median Income Families

Description: Median "Contract Rent" (monthly) for the universe of renter-occupied housing units paying cash rent at the tract level. Aggregated to the county/CZ level by taking the mean, weighting by the population with above-median incomes for families with children < 18 present (income bins 10 and above)

Source: 2000 Census SF3a (NHGIS SF3a, code: GBO) [Rent]; 2000 Census SF3a (NHGIS SF3a, code: GYI) [for weights]

Median Monthly Rent for Below-Median Income Families

Description: Median "Contract Rent" (monthly) for the universe of renter-occupied housing units paying cash rent at the tract level. Aggregated to the county level by taking the mean, weighting by the population with below-median incomes for families with children < 18 present (income bins 9 and below)

Source: 2000 Census SF3a (NHGIS SF3a, code: GBO) [Rent]; 2000 Census SF3a (NHGIS SF3a, code: GYI) [for weights]

Median House Price for Above-Median Income Families

Description: Median value of housing units at the tract level. Aggregated to the county/CZ level by taking the mean, weighting by the population with above-median incomes for families with children < 18 present (income bins 10 and above)

Source: 2000 Census SF3a (NHGIS SF3a, code: GB7); 2000 Census SF3a (NHGIS SF3a, code: GYI) [for weights]

Median House Price for Below-Median Income Families

Description: Median value of housing units at the tract level. Aggregated to the county/CZ level by taking the mean, weighting by the population with below-median incomes for families with children < 18 present (income bins 9 and below)

Source: 2000 Census SF3a (NHGIS SF3a, code: GB7); 2000 Census SF3a (NHGIS SF3a, code: GYI) [for weights]