MARCH 2018

The Inheritance of Black Poverty: Technical Paper

Scott Winship
The Brookings Institution, Former Fellow

Richard V. Reeves
The Brookings Institution, Senior Fellow

Katherine Guyot
The Brookings Institution, Research Assistant

This report is available online at: https://www.brookings.edu/research/the-inheritance-of-black-poverty-its-all-about-the-men/

B | Economic Studies at BROOKINGS

The Brookings Economic Studies program analyzes current and emerging economic issues facing the United States and the world, focusing on ideas to achieve broad-based economic growth, a strong labor market, sound fiscal and monetary policy, and economic opportunity and social mobility. The research aims to increase understanding of how the economy works and what can be done to make it work better.

Data

Our main estimates, outlined and described in the summary paper, are based on the National Longitudinal Survey of Youth 1997 (NLSY97), which consists of participants who were 12 to 16 years old as of December 31, 1996. NLSY97 participants were most recently interviewed in 2015-16, when they were 30 to 36 years old. In this Technical Paper, we describe our approach and methodology. Our full results are available in data tables in a document labeled Results: Data. In addition to the NLSY97 calculations, we report additional findings from the National Longitudinal Survey of Youth 1979 (NLSY79), which consists of individuals who were 14 to 21 years old as of December 31, 1978. NLSY79 participants were interviewed annually through 1994 and are now interviewed biennially.

Race/ethnicity

To maintain consistency between the NLSY79 and NLSY97, our race/ethnicity variable codes individuals as black non-Hispanic, non-black non-Hispanic, Hispanic, or non-Hispanic of two or more races. The NLSY79 did not ask respondents for more detailed race/ethnicity information until the 2002 survey, when many of the original respondents were not interviewed. We use the 1978 pre-screening race/ethnicity variable because it contains more complete data for the sample, though this means that we are unable to distinguish those who are white non-Hispanic from other groups who are non-black non-Hispanic. The following table shows the proportion of individuals in our subsamples who fall into each race/ethnicity category:

	NLSY79	NLSY97
Black Non-Hispanic	27%	26%
Hispanic	18%	21%
Non-Black Non-Hispanic	55%	52%
Two or more	N/A*	1%

*This category does not exist in the NLSY79.

We often refer to non-black non-Hispanics as white, though it should be noted that a small number of those who are non-black non-Hispanic are also non-white. The NLSY97 contains more detailed information about race and ethnicity than the NLSY79, so we can see that of the 4,665 people in the NLSY97 who are coded as non-black non-Hispanic, 4,413 (or 95 percent) identify as white.

We do not estimate mobility rates for Hispanic respondents, though they are included in the overall income distributions. The NLSY is not ideal for examining the Hispanic population of the U.S. because it does not include individuals who immigrate to the U.S. after the first round of interviews. In Race and Economic Opportunity in the United States: An Intergenerational Perspective," Raj Chetty, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter find high rates of relative mobility for Hispanic individuals.

Age

We measure relative social mobility by comparing adult respondents' income quintiles to their parents' income quintiles when the youths were 14 to 16 years old. Our target population consists of youths who were ages 14 to 16 and living with their parents (biological, step, or foster) at the beginning of 1979 or 1997. For respondents in the NLSY79, we calculate ages using the date of birth reported in 1981, as is recommended by the Bureau of Labor Statistics due to discrepancies in

some respondents' reported birthdates.[1] We use the date reported in 1979 only if the 1981 value is missing.

Individuals in our target age group in the NLSY97 were last interviewed when they were 32 to 35,[2] so we initially define mobility according to the percentage of respondents who reach each quintile between the ages of 32 and 35. NLSY79 participants in our target age group were 33 to 35 during the 1998 follow-up, so we choose to compare NLSY97 respondents in 2015-16 to NSLY79 respondents in 1998. To minimize the effects of idiosyncratic income changes or missing values in either of those two years, we average respondents' income over three consecutive surveys to represent their incomes as adults. Incomes for NLSY79 respondents are measured in 1994, 1996, and 1998, and incomes for NLSY97 respondents are measured in 2011, 2013, and 2015-16.[3] This means that the mobility rates essentially reflect the percentage of respondents who reach each income quintile between the ages of 28 and 35. We also assess mobility later in NLSY79 participants' lives, when they are between the ages of 41 and 47, using the 2006, 2008, and 2010 follow-ups.

Income

We measure parental income using the average of parent-reported total net family income in 1979 and 1980 (NLSY79) and the average of parent-reported gross household income 1997 and 1998 (NLSY97). The NLSY97 asked respondents about household rather than family income until the 2004 follow-up, when it switched to family income. The NLSY79 always asks about family income rather than household income. In general, family income is limited to the incomes of household members who are related by blood, marriage, or adoption, whereas household income may include the incomes of unmarried partners or other unrelated individuals living in the household. However, in later years of the NLSY97, "marriage-like" relationships are treated as akin to marriage, so family income includes the incomes of unmarried partners. This distinction will affect our ability to compare estimates between the NLSY79 and the NLSY97 to the extent that the incomes of unmarried partners and other unrelated household members differentially impact certain groups' income ranks.

In an alternative version of our family income variable, we adjust for family size according to a square root equivalence scale (dividing income reported in year *t* by the square root of family size reported in year *t-1*). As mentioned above, a family unit generally includes individuals related by blood, marriage, or adoption, but may include unmarried partners in the NLSY97. When using household income, we adjust by the square root of household size. Family income for adult respondents in the NLSY97 is also adjusted by household size due to data availability, though family income should be adjusted by family size and household income by household size.

We draw a key distinction between family income and individual earnings. Our measure of individual earnings combines an individual's wage and salary income with his or her business and farm income. The NLSY97 includes military income in its measure of wage and salary income. We add

. . .

[1] https://www.nlsinfo.org/content/cohorts/nlsy79/topical-guide/household/age

[2] The latest follow-up for the NLSY97 was conducted between the fall of 2015 and the spring of 2016, so some of those who were 16 in 1997 had the chance to turn 35 by the time they participated in Round 17.

[3] If an individual's income is reported missing in one of these three years, we use the average of his or her income in the remaining two years. If only one year of income is reported, we simply use the non-missing value to represent income over those three years.

military income to wage and salary income for respondents in the NLSY79 to be consistent with the NLSY97. Importantly, individual earnings do not include spousal or partner earnings. Of course, a spouse's earnings may affect one's own earnings if the addition of a secondary earner allows one to work part-time or stop working altogether. To determine whether this or other unexplained reasons for having zero earnings are driving differences in mobility by race, we run additional earnings tabulations excluding those who have zero earnings in 1994-1998 (NLSY79) or 2011-2015 (NLSY97).

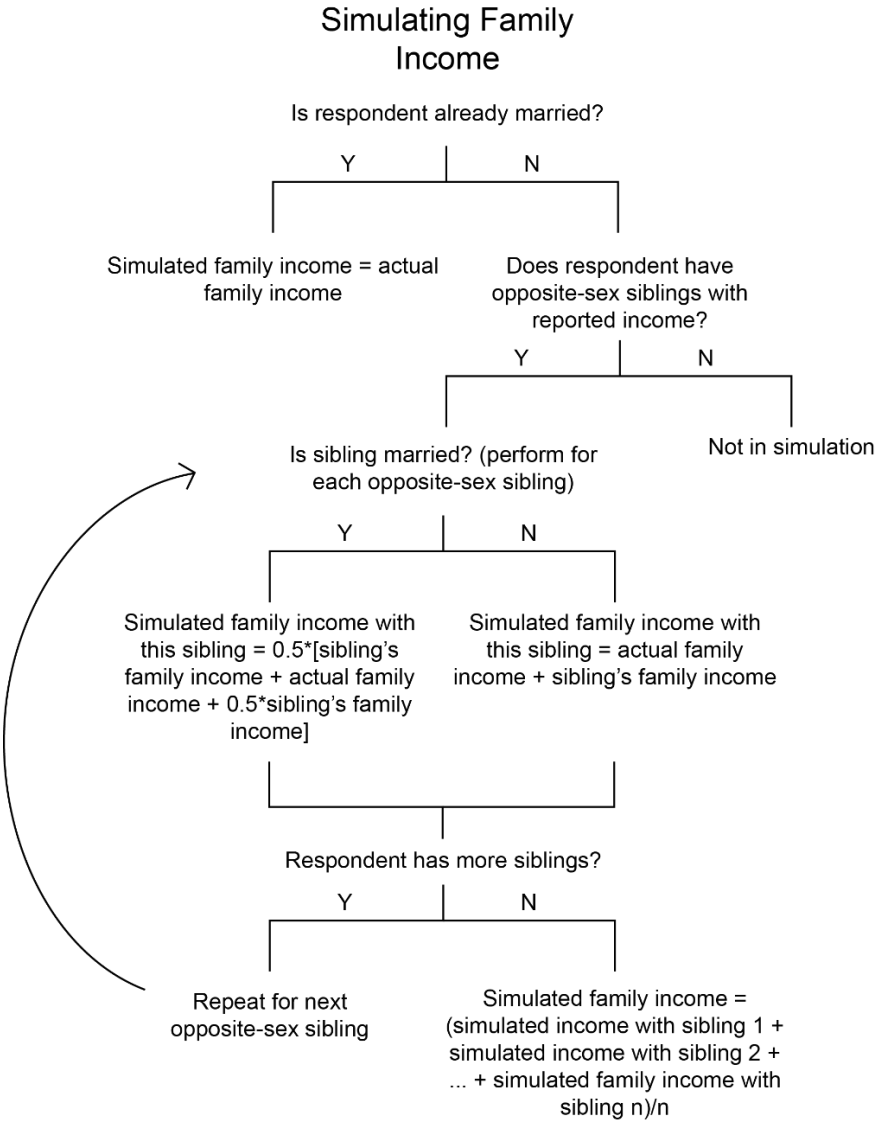
The NLSY79 topcodes family income, wage and salary income, and business and farm income at $75,001 through 1984, then $100,001 through 1988, then based on the average incomes of top earners in the remaining survey years, where top earners are defined as either outliers or as the top two percent of earners. The NLSY97 topcodes income for the top 2 percentiles of the distribution at the average of the topcoded incomes. To ensure that income topcoding is consistent across all years of the NLSY79 and NLSY97, we take a more restrictive approach and topcode all income variables at the 97th percentile of the unweighted distribution. We then adjust for inflation using the personal consumption expenditures (PCE) deflator. Because income is always reported for the year preceding the interview, income reported in 1979 is inflation-adjusted using the 1978 PCE deflator, income reported in 1980 is inflation-adjusted using the 1979 PCE deflator, and so on. The table below displays inflation-adjusted values of median parental income, earnings, family income, and simulated income in the NLSY79 and NLSY97 for various subgroups.

	Median Parental Income		Median Earnings		Median Family Income		Median Simulated Income	
	NLSY79	NLSY97	NLSY79	NLSY97	NLSY79	NLSY97	NLSY79	NLSY97
All youths	$42,283	$48,285	$27,942	$28,563	$49,066	$62,643	$67,504	$75,243
Non-black non-Hispanic	$55,197	$68,365	$33,579	$34,815	$61,134	$76,824	$74,404	$86,734
Men	$56,187	$67,038	$43,213	$42,581	$59,435	$76,017	$75,520	$85,860
Women	$54,956	$69,767	$22,892	$26,540	$62,871	$78,004	$73,965	$87,140
Black non-Hispanic	$27,807	$26,069	$21,178	$18,522	$34,107	$38,575	$55,198	$52,280
Men	$28,036	$26,773	$24,902	$20,228	$35,565	$40,541	$54,721	$50,902
Women	$26,929	$25,851	$16,897	$17,744	$32,679	$37,106	$55,396	$57,133
Men (all)	$41,990	$48,818	$36,195	$34,833	$48,732	$62,855	$66,305	$74,838
Women (all)	$42,570	$47,963	$20,736	$22,227	$49,393	$62,532	$68,685	$75,366

Simulation

The next step is to create a simulated family income variable that illustrates what would happen if single individuals were married. The simulation provides unmarried individuals with secondary earners who add to their family incomes, potentially improving their income mobility. We know that individuals are likely to marry partners with similar backgrounds, so we choose to supplement unmarried respondents' family incomes with their siblings' earnings, essentially "marrying off" respondents to their siblings. Because men and women's earnings distributions still differ and heterosexual marriages are more common than same-sex marriages, especially in the timeframes under study, we only simulate heterosexual marriages: that is, women "marry" their brothers and men "marry" their sisters.

If an individual is already married with a spouse present in the household, simulated family income is the same as actual family income. Since



unmarried partners may be counted as family members in the NLSY97, we treat cohabiting individuals as married in the NLSY97 (but not in the NLSY79). The remaining unmarried individuals are assigned family incomes that combine their own family incomes with their siblings' family incomes. Individuals without opposite-sex siblings or with opposite-sex siblings who do not report income in any of the three survey years under study are not assigned simulated family incomes. In other words, unmarried women with no reported brother income and unmarried men with no reported sister income are omitted when we tabulate simulated family income. In case the omission of those without opposite-sex siblings reporting earnings impacts mobility estimates, we also tabulate actual (not simulated) family income among the subsample of individuals for which simulated family income is defined.

For unmarried individuals with siblings, the calculation to create simulated family income depends upon the number of siblings and the siblings' marital statuses. The simplest case is for an individual with one unmarried opposite-sex sibling. In this situation, simulated family income is the sum of the two siblings' separate family incomes. If a respondent's sibling is married (or cohabiting, in the NLSY97), the sibling's family income may include a spouse's income as well. To avoid marrying off an already-married sibling, we take the average of two cases: first, that a respondent's married family income would simply equal his or her married sibling's family income (in other words, the respondent is essentially assigned the sibling's spouse's income); and second, that a respondent's married family income would equal his or her current family income plus half of the married sibling's family income (supposing that siblings and their spouses contribute

	NLSY79		NLSY97	
	Black	White	Black	White
In simulation	**64%**	**80%**	**60%**	**82%**
Married in all years	25%	50%	26%	49%
Not married, matched with sibling	33%	23%	10%	3%
Married in some years, not married and matched with sibling in other years*	5%	6%	24%	30%
Not in simulation	**36%**	**20%**	**40%**	**18%**
Married in at least one year**	2%	1%	2%	1%
Not married, no opposite-sex siblings	26%	15%	30%	12%
Not married, cannot be matched with sibling	3%	1%	1%	0%
Unreported marital status	5%	3%	7%	4%

*Incomes are averaged over 2011, 2013, and 2015. Individuals who are married in one or two of those three years are given simulated incomes with siblings in the years in which they are not married. If it is not possible to match them with siblings, their incomes are coded as missing in the years in which they are not married.

**A married individual can be present in the family income distribution but absent from the simulated family income distribution if he or she is married in at least one but not all of the three years, has missing income during the year(s) in which he or she is married but reported income in the year(s) in which he or she is not married, and cannot be matched with a sibling in the years in which he or she is not married. This number includes only people who reported their marital status in all three years.

about equally to family income). If a respondent has more than one sibling, final simulated family income is the average of all sibling-specific simulated family incomes.

Note that we have not included any age restrictions on siblings when calculating simulated income. When we construct earnings distributions, we will limit the sample to respondents who are between the ages of 14 and 16 in the first survey year. However, respondents may be assigned the incomes of siblings who are older or younger than 14 to 16 in the survey year, as long as these siblings are still within the age range present in the NLSY. If two siblings are both between the ages of 14 and 16 and have non-missing simulated income, then both will be present in the subsample for which we calculate income ranks.

In the table on the previous page, we show the numbers of black and white individuals who are "married off" as a result of our simulation.

Presentation of Results

We create income quintiles using Stata's "_pctile" command and cutting quintiles at the 20th, 40th, 60th, and 80th percentiles of the resulting distributions. We use person weights in the latest survey year for which we measure adult income (1998 for the NLSY79 and 2015 for the NLSY97). Respondents are included in the distributions if they are between the ages of 14 and 16 in the first survey year, have non-missing incomes for the relevant income measure, and have non-missing person weights for the relevant year. We calculate separate distributions for each income variable, as well as separate earnings (but not family income) distributions for men and women. We separate men and women in the earnings distributions because women's earnings are lower on average than men's, and this paper is primarily interested in comparing mobility rates across race within gender.

When creating distributions for parental income, there is a danger of including some households more than once since some households have more than one child between the ages of 14 and 16. To avoid this, we create parental income distributions only for the first child listed in each household, then assign the resulting parental income quintile to each child in the household. The following table shows the lower bounds of the second through fifth quintiles of each income distribution.

For each income variable, we cross-tabulate parental income quintile and child income quintile by race and gender. We run tabulations for 28- to 35-year-olds in the NLSY97 and NLSY79, as well as for 41- to 47-year-olds in the NLSY79 (denoted "older" in the accompanying tables), using the following income variables:

1. Parent family income rank versus child earnings rank
2. Parent family income rank versus child earnings rank excluding respondents with zero earnings
3. Parent family income rank versus child family income rank
4. Family-size-adjusted parent family income rank versus family-size-adjusted child family income rank
5. Parent family income rank versus simulated child family income rank
6. Parent family income rank versus child family income rank among respondents who have non-missing simulated incomes

A summary of our findings is here. Our full results are available in data tables in the accompanying document, Results: Data, in which tabs are labeled according to the survey year (79 or 97) and income variables (1 through 6 according to the list above). Our datasets and code are available available on request from KGuyot@brookings.edu.

The Center on Children and Families studies policies that affect the well-being of America's children and their parents, especially children in less advantaged families. The Center addresses the issues of poverty, inequality, and lack of opportunity in the United States and seeks to find more effective means of addressing these problems.

Questions about the research? Email communications@brookings.edu.
Be sure to include the title of this paper in your inquiry.