NBER WORKING PAPER SERIES

THE GREAT ESCAPE:
INTERGENERATIONAL MOBILITY IN THE UNITED STATES SINCE 1940

Nathaniel G. Hilger

Working Paper 21217
http://www.nber.org/papers/w21217

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2015

Previously circulated as "The Great Escape: Intergenerational Mobility Since 1940." I thank Anna Aizer, Sam Asher, Nathaniel Baum-Snow, Leah Boustan, David Card, Kenneth Chay, Raj Chetty, Pedro Dal Bo, Eleanor Dillon, Andrew Foster, Claudia Goldin, Rachel Friedberg, Justine Hastings, Caroline Hoxby, Lawrence Katz, Brian Knight, Michael Kremer, David Lagakos, John Logan, Glenn Loury, Robert Margo, Paul Novosad, Claudia Olivetti, Emily Oster, Daniele Paserman, Emmanuel Saez, Jesse Shapiro, Matthew Turner, David Weil and Danny Yagan. Jenna Anders, Motaz Al-Chanati, James Bernard, Kian Ivey and Jorge Perez provided excellent research assistance. Weiwei Zhang provided helpful assistance with the 100% census data. Financial support from the Russell Sage Foundation is gratefully acknowledged (Grant No. 83-16-07). All errors are my own. The views expressed herein are those of the author and do not necessarily reflect the views of the National Bureau of Economic Research.

The Great Escape: Intergenerational Mobility in the United States Since 1940
Nathaniel G. Hilger
NBER Working Paper No. 21217
May 2015, Revised April 2016
JEL No. J1,J24,J62,N01,N3,O15

ABSTRACT

I develop a method to estimate intergenerational mobility (IM) in education on large cross-sectional surveys and apply the method to U.S. census data from 1940 to 2000. The method estimates IM directly for children age 26-29 who still live with parents and adjusts for independent children using a procedure that I validate extensively. Estimates imply large post-1940 gains in IM that were (1) driven primarily by large IM gains in the South for both whites and blacks, (2) larger for blacks due to their greater concentration in the South, and (3) driven by high school rather than college enrollment.

Nathaniel G. Hilger
Department of Economics
Brown University
64 Waterman Street
Providence, RI 02912
and NBER
nathaniel_hilger@brown.edu

1. Introduction

Intergenerational mobility (IM) is an important social objective for many individuals and policymakers, and may affect public attitudes toward other social objectives such as equality and growth (Piketty, 1995; Benabou and Ok, 2001; Corak, 2013). However, surprisingly little is known about IM variation over time, space and groups. The main empirical problem is that measuring IM requires data on labor market outcomes for both parents and children. Several panel datasets contain this information, but they begin in the late 1960s and are too small to examine mobility over time and subgroups with precision (e.g., Lee and Solon, 2009), while tax records only link children beginning in the 1990s (Chetty et al., 2014a). The scarcity of evidence on longer-term trends overall and for various subgroups is unfortunate because many interventions often thought to increase equality of opportunity such as the early GI Bills, Great Society programs, and the Civil Rights movement predate these panel datasets.

In this paper I develop a new method to estimate educational IM on U.S. census data–or any other large cross-sectional family survey–back to 1940. IM statistics typically characterize the joint distribution of parent and child outcomes. Unfortunately, census data only link parent and child outcomes while children still live with parents, and many children leave the parental household at ages before final educational attainment or earnings can be observed meaningfully (Cameron and Heckman, 1993). I develop a simple, semi-parametric adjustment for these "missing" independent children that allows me to estimate the conditional expectation function (CEF) of children's educational attainment in their late 20s (e.g., ages 26-29) with respect to parental income or education. After this adjustment, it is straightforward to recover point estimates of standard educational IM statistics based on the slopes and intercepts of these CEFs.

The adjustment for independent children rests on two simple and verifiable assumptions. To illustrate, consider an example with two parental groups in a fixed year. Let children have either "low-income" or "high-income" parents denoted 0 and 1, respectively. Among 27-year-olds, I observe 100 children living with high-income parents, 100 with low-income parents and 300 living independently, with average highest grade attained of 14, 12 and 12, respectively. I therefore observe a schooling CEF intercept of 12 years of schooling and a slope of 2 years of schooling across parental groups, but only for *dependent* children. I need to know two things to account for the remaining 60% of children who are independent: their parental group composition, and their average schooling by group. I first make a "parallel trends" assumption that the schooling CEF among independent children has the same slope as the CEF among dependent children: here 2 years. Now observe that the great majority of children up through age 17 still live with their parents. Suppose I observe 200 high-income and 300 low-income 17-year-olds. Under a second "smooth cohorts" assumption that parental group shares do not change across cohorts, I infer that 100 of the independent 27-year-old children have high-income parents and 200 have low-income parents. Let h equal average schooling of low-income 27-year-old independent children. We can now solve for h: $12 = \frac{100}{300}(h+2) + \frac{200}{300}h \implies h = 11.33$. Total schooling of low-income children is therefore $\frac{100}{300} \cdot 12 + \frac{200}{300} \cdot 11.33 = 11.55$ and total schooling of high-income children is $\frac{100}{200} \cdot 14 + \frac{100}{200} \cdot 13.33 = 13.665$. The *total* schooling CEF therefore has intercept of 11.55 and slope of 2.11. Below I formalize these two assumptions of parallel trends and smooth cohorts, generalize the method to more than two groups, and use multiple datasets spanning the entire 1940-2000 period to validate both the two assumptions and the resulting CEFs in the U.S. historical context.

I find that these assumptions appear surprisingly plausible in a wide range of panel datasets, though I also document cases in which they break down. Despite its limitations the approach here is useful because it opens up many new possibilities for research on IM in time periods, places and subgroups that can only be studied in large cross-sectional datasets such as population censuses. As a first application of the new method, I examine long-term IM trends in the U.S. nationally and for many subgroups.[1] Estimates imply that educational IM increased significantly after 1940 and stabilized sometime around 1960-80. Back of the envelope calculations suggest these IM gains generated significant annual aggregate earnings growth for several decades. I confirm that these large post-1940 IM gains are in fact consistent with analogous time trends in other data sources, and post-1940 time trends in *income-based* IM in census data based on the very different estimation strategy of Aaronson and Mazumder (2008). However, I find that multiple other data sources and prior research indicate a decline in educational IM after 1980 that my approach fails to detect. Despite this bias, I show that my educational IM estimates

[1] In Hilger (2016) I exploit the methodology to estimate educational IM for cohorts of Asian-Americans and blacks born in California as far back as 1911, an exercise which is only feasible in census data due to the very small sizes of these subgroups historically.

by state-of-birth since 1980 are highly correlated with income-based IM estimates across states in tax data (Chetty et al., 2014a), confirming the method's ability to detect spatial variation in IM.

The new approach allows me to extend this spatial variation back further in time than any prior study and reveals that post-1940 IM gains were driven primarily by large IM gains in the South from astonishingly low initial levels. Southern IM gains, in turn, were primarily driven by high school rather than college enrollment, and were similar for men and women. Strikingly, southern IM gains were also similar for whites and blacks, and thereby account for larger post-1940 IM gains of blacks nationally due to the greater geographic concentration of blacks in the South at this time.

The findings suggest a potentially important role for the "transformation" of the Southern economy after 1930 (Wright, 1986) in explaining post-1940 IM gains. To further explore potential mechanisms I construct a panel dataset of educational IM by state of birth for years 1940 to 2000 and show that higher IM correlates with larger black population shares, higher per-capita state income, lower income inequality, better K-12 schools, and several other factors, in many cases even conditional on state and year fixed effects. These correlations largely coincide with those reported in (Chetty et al., 2014a) with the exception that I find a larger historical correlation with income levels. This discrepancy likely relates to the broader trends of industrialization, urbanization, and political liberalization that took place over the long historical period studied here.

2. Related Literature

The primary advantage of the method developed here is that it potentially allows estimation of IM in time periods, places and subgroups that are not well-covered by existing panel datasets. For example, to my knowledge there is currently no method permitting estimation of any type of IM statistic separately by state of birth before the advent of tax data in the 2000s (Chetty et al., 2014b). The particular application I develop here also contributes to a broad literature on empirical estimation of IM. While historical trends in IM across states and demographic groups in the U.S. have not previously been available, a growing literature compares IM across places in recent decades.[2]

[2] A few of the important papers in this growing literature include Erikson and Goldthorpe (1992); Checchi et al. (1999); Jantti et al. (2006); Corak (2006); Hertz et al. (2008); Mayer and Lopoo (2008); Long and Ferrie (2013); Clark (2014); Chetty et al. (2014a); Chetty and Hendren (2015).

Many prior studies have estimated various measures of IM for the U.S. nationally as reviewed in Solon (1999) and Black and Devereux (2011). In this paper I focus on educational mobility out of necessity, but education also has the advantage of being less sensitive to well-known problems with measurement error and lifecycle bias afflicting measures of income (Solon, 1992; Mazumder, 2005, 2015; Mazumder and Acosta, 2015). Hertz et al. (2008) estimate educational IM in the U.S. for cohorts born as early as 1932, as well as for many other countries, using the World Bank Living Standards Measurement Surveys. They find a gradual increase in educational IM according to intergenerational elasticities, but no change in intergenerational correlations. I find that IM increased significantly for cohorts born between 1911 and 1932 in both elasticities and correlations, and I obtain results for cohorts born after 1932 that differ from their estimates. One possible explanation for this discrepancy is that Hertz et al. (2008) estimate time trends using variation in ages 20-69 from surveys conducted during years 1994-2000. As they discuss, their trends may therefore incorporate bias from high and selective mortality attrition over this period (Costa, 2015), as well as bias from changes in recall errors in own or parental education by age (Goldin, 1998). Hertz et al. (2008) are also unable to study sub-national IM variation across places or demographic groups due to sample size limitations.

Hertz (2007), Lee and Solon (2009) and Harding et al. (2009) document that intergenerational *income* elasticities exhibit no significant time trend for children born between 1950 and 1970. Chetty et al. (2014b) document stable rank-rank income IM for cohorts born between 1970 and 1990, and argue that national IM statistics based on ranks and logs are likely comparable in practice, implying stable income mobility for cohorts born 1950-1990. However, Mitnik et al. (2015) and Mazumder (2015) present evidence that IGE estimates and, to a lesser extent, rank-rank income-based IM statistics in Chetty et al. (2014b) are biased down by various measurement problems, which could also affect trend estimates if these biases vary significantly over time.

Aaronson and Mazumder (2008) estimate intergenerational income elasticities in census data back to 1940 by instrumenting for parental income with children's state of birth. They find that this measure of income-based IM increased after 1940 and, in contrast with the above studies, decreased after 1980. As Aaronson and Mazumder (2008) discuss in detail, their method may yield biased estimates of time trends if places have *time-varying* causal effects on children's income that are not captured by parental income. Putting concerns about this assumption aside, however, the results in Aaronson and Mazumder (2008) suggest similar post-1940 time trends for income IM as those I estimate for educational IM. However, my approach offers a significant advantage over that of Aaronson and Mazumder (2008) by facilitating estimation of IM *separately* across states of birth and demographic groups in every census beginning in 1940, rather than relying on state of birth as an allegedly exogenous instrument for parental income to estimate national IM trends. While my approach also relies on strong assumptions, it offers a new, inexpensive, and widely-available window onto a broad range of IM heterogeneity that could not otherwise be studied with existing data or methods.

Other researchers have estimated *occupational* IM in census data using information about names to construct matched panel datasets (Long and Ferrie, 2013) or pseudo-panels (Clark, 2014; Olivetti and Paserman, 2015). Findings in Long and Ferrie (2013) and Olivetti and Paserman (2015) suggest that occupational mobility declined significantly from the 19th to the mid-20th century, and may have begun to increase before 1940 depending on the method used to impute occupational income. An increase in IM after 1940, as I find here for educational IM, may therefore represent a return to higher 19th century levels. However, long-term comparisons of educational and occupational IM are difficult due to complex potential changes in returns to different occupations and to educational attainment.

Like most prior literature, I focus on two-generation mobility statistics relating child outcomes to parental outcomes. Recent work suggests that two-generation IM statistics likely overstate multi-generational IM (e.g., Clark, 2014; Olivetti et al., 2014; Stuhler, 2014; Braun and Stuhler, 2015; Solon, 2015).[3] Nonetheless, two-generation IM statistics are likely to remain critical benchmark measures of equality of opportunity due to data limitations in much the same way that GDP, Gini coefficients and poverty rates remain valuable measures of social objectives despite their well-known limitations.

Nybom and Stuhler (2014) provide important insights into the interpretation of time trends in IM. They show that one-time improvements in equality of opportunity (e.g., positive education supply shocks) will tend to increase two-generation IM statistics greatly at first when only child outcomes are affected, but less so in subsequent generations when both parent and child outcomes fully reflect the new environment. Counterintuitively, this effect can mechanically *decrease* two-generation IM statistics gradually starting about 30 years after a reform when the children first affected by the new environment begin to have their own children. Second, large social changes not directly related to equality of

[3] Note that much of this work rejects the extreme claim of Clark (2014) that multi-generational IM is similarly low across all times and places (e.g., Braun and Stuhler, 2015; Solon, 2015).

opportunity (e.g., industrial reallocations, wars, technology shocks, etc.) tend to increase two-generation IM statistics temporarily as advantages shift among dynasties with newly-favored skills and assets, again followed by declines in IM 30-60 years later. Below I argue that educational IM increased after 1940 most likely due to broad changes in the South, and that IM subsequently decreased after 1980. The arguments in Nybom and Stuhler (2014) imply that care must be taken in searching for causes of changes in IM, especially *declines* in IM. The post-1980 IM decline may represent new changes in the U.S. economy such as greater reliance on post-secondary educational institutions, but could also represent mechanical aftershocks of the much earlier forces that drove post-1940 IM gains. Distinguishing between these two explanations for the post-1980 decline in educational IM therefore represents an important empirical problem that I am not able to resolve in this paper.

A few papers also compare IM across races. Torche (2015) reviews some evidence of higher relative occupational IM for blacks than whites in the U.S. in recent decades, which would line up with my findings for educational IM. However, she concludes that little reliable evidence exists on time trends in black IM. Bhattacharya and Mazumder (2011) point out that comparisons of black and white intergenerational elasticities are problematic because they reflect mean-reversion to *group-specific* rather than population parental mean income, which can generate spurious racial differences in IM statistics if underlying CEFs are nonlinear.[4] I address this issue by estimating slopes of CEFs for whites and blacks on data collapsed to the same population distributions of parental SES, which amounts to reweighting all groups to the same parental SES distribution. I also find that these CEFs are approximately linear.[5] Therefore my estimated differences in IM statistics capture genuine differences in underlying group CEFs, not distributional shifts across regions of the same nonlinear CEF. Like Bhattacharya and Mazumder (2011), I find evidence of slightly lower upward mobility for blacks than whites in more recent decades. However, I uncover much lower black IM in earlier decades and show these gaps stem from greater geographic concentration of blacks in the South.[6]

[4]The same concerns about nonlinearities in CEFs have played an important role in comparisons of IM across countries. See e.g., Bratsberg et al. (2007); Bratberg et al. (2015).

[5]Given that these CEFs are approximately linear, in many results below I estimate intercepts and slopes of these CEFs weighting by cell sizes in order to increase precision. Apart from improvements in precision, results are not sensitive to these weights.

[6]Bhattacharya and Mazumder (2011) compare IM across races by focusing on upward mobility of children measured in terms of population quantile movements. My most relevant results for comparison are my estimates of rank-rank absolute upward educational mobility (CEF intercepts) in 1990 in Table A.5,

Finally, there are currently large-scale proposals emerging to link individuals across censuses and other datasets spanning the entire 20th century (e.g. Grusky et al., 2015; Johnson et al., 2015, and others). One might worry these efforts will render the methods here obsolete. However these proposals are still in early stages, and it is possible that linkage methodologies will introduce their own biases and exclude certain subgroups such as married women, minorities, people with low literacy rates, people with limited labor force participation, immigrants, and many other groups of interest. Even if these linkages prove highly successful in the U.S. (hopefully!), many other countries will continue to lack long-term, large-scale microdata linking adult children with their adult parents. Therefore the method here is likely to remain useful in a wide range of contexts well into the future, and can complement alternative data sources on IM as they develop.

3. Methodology

Consider an adult individual i in generation t with educational attainment $h_{i,t}$ in a fixed calendar year. Education can be measured in levels, ranks, z-scores, or any other transformation. Likewise let $y_{i,t-1}$ denote some measure of parental SES such as income or education, or any transformation of these variables. Most IM statistics in the literature characterize the CEF $E\left[h_{i,t}|y_{i,t-1}\right]$ in some parametric fashion. Much of the prior literature estimates best linear predictors of this CEF using regressions of the form:

$$h_{i,t} = \alpha + \beta y_{i,t-1} + e_{i,t} \tag{1}$$

interpreting estimates of β and possibly α as descriptive measures of IM.

Unfortunately, this regression cannot be estimated directly in census data because the census only contains parental characteristics for the subset of children who still live with parents. The key problem is that many children move out before standard ages of school completion, and the few children who do live with parents at older ages may not be representative of all children (e.g., Cameron and Heckman, 1993). Figure I illustrates this problem using age variation across cohorts in 1980 census data. Once children reach age 18 they rapidly begin to leave the parental home, yet educational attainment continues to increase well into children's mid-20s. I here develop a simple correction for the missing

Columns (3)-(4). These results are still only roughly comparable due to differences in child outcomes (income vs. education) as well as other differences in methodology and sample.

independent children at later ages to estimate $E\left[h_{i,t}|y_{i,t-1}\right]$, thereby recovering α, β, or any other statistic based on $E\left[h_{i,t}|y_{i,t-1}\right]$. I validate this correction empirically in detail below.

Let $h_{a,y}$ represent average years of completed schooling for children of fixed age a with parental income or education group y. Let $h_{a,y}^D$ and $h_{a,y}^I$ indicate average years of schooling for dependent children and independent children at age a, respectively. To fix ideas, think about a as an age in the range of 26-29 when (I confirm below) average educational attainment has stabilized, and think about y as a parental income decile. Let $N_{a,y}^D$ and $N_{a,y}^I$ indicate the number of dependent and independent children at age a. By definition,

$$h_{a,y} = d_{a,y}h_{a,y}^D + \left(1 - d_{a,y}\right)h_{a,y}^I. \tag{2}$$

where $d_{a,y} = \frac{N_{a,y}^D}{N_{a,y}^D + N_{a,y}^I}$, the dependency rate for children at age a in parental group y. For dependent children, I observe both average schooling, h_y^D, and number of children for each parental group, $N_{a,y}^D$. For independent children, I only observe the total number of children N_a^I and overall average schooling h_a^I, pooling all parental groups. I do not observe schooling or frequencies for independent children allocated to parental groups, $h_{a,y}^I$ and $N_{a,y}^I$. I therefore need to estimate these unobserved terms in order to impute overall schooling in parental groups, $h_{a,y}$.

I make two assumptions: (1) a ***parallel trends*** assumption for dependent and independent children by parental group status, and (2) ***smooth group cohort share trends (“smooth cohorts”)*** for parental groups. The parallel trends assumption states that:

$$f\left(h_y^D, h_y^I\right) = \rho \;\forall y \tag{3}$$

where $f\left(.\right)$ is any function known to the researcher based on evidence outside the census data. I refer to this as “parallel trends” because I implement this assuming $f\left(h_y^D, h_y^I\right) = h_y^D - h_y^I$. This function places no restriction on the shape of the CEFs of children’s education in parental income or education; it simply requires these CEFs to be equal up to a constant across dependent and independent children, where this constant is free to vary over time, place, demographic groups, etc. The economic underpinnings of this assumption depend on complex, unobserved relationships between schooling, dependent status, and parental groups. However, the assumption partly reflects a simple intuition: rich children exhibit better educational outcomes than poor children, wherever they choose to live in their late 20s. The simplest example would be $\rho = 0$, in which case children’s dependency choices in their late 20s do not depend on prior educational attainment after conditioning on parental group.

The second assumption is smooth cohorts. This assumption allows me to impute the number of independent children in each parental group at older ages, using the total number of dependent children at earlier ages when most children are still dependent. Denote the total number of children in each parental group in cross-sectional data as $N_{a,y}$, where $N_{a,y} = N^D_{a,y} + N^I_{a,y}$. The assumption is that parental group shares in older cohorts can be approximated by functions of parental group sizes (and shares) in younger cohorts:

$$\frac{N_{a,y}}{N_a} \approx g_y \left(N_{a-k-1,y}, N_{a-k-2,y}, ..., N_{1,y}\right) \;\forall y \tag{4}$$

for any functions $g_y(.)$ that need not be known to the researcher but are smooth enough to be estimated parametrically in census data, and where k captures the distance between the target age and the ages used in estimation. Below I show that most children under 18 still live at home with parents in all census years, implying $N^D_{a,y} \approx N_{a,y}$ for $a < 18$. Under smooth cohorts, we can therefore estimate group cohort shares in cohorts that are old enough to have largely completed schooling by estimating the functions $g_y(.)$ on group cohorts under age 18. Once I have estimated $\frac{N_{a,y}}{N_a}$, estimates of the missing $N^I_{a,y}$ terms follow from the fact that I already know $\frac{N^D_{a,y}}{N^D_a}$, N^D_a, and N^I_a.

In practice I implement smooth cohorts assuming constant group cohort shares between cohorts age 17 and cohorts age 26-29 within each calendar year of census data, i.e. setting $g_y(.) = \frac{N_{17,y}}{N_{17}}$ where $N_{17} = \sum_y N_{17,y}$. This implementation requires that group cohort shares remain approximately stable across cohorts born 10 years apart. Sharp "baby booms" or "baby busts" that alter cohort sizes arbitrarily over the course of a decade are permitted as long as they are approximately balanced across parental groups. Below I show this simple implementation performs as well or better than more complex estimators of cohort share trends.

Under the assumption of parallel trends with $h^D_{a,y} - h^I_{a,y} = \rho$ and smooth group cohort shares, I can solve for the implied estimate of the unobserved parameter ρ as

$$\hat{\rho} = \left(\sum_{j=1}^{J} \frac{\hat{N}^I_{a,j}}{\hat{N}^I_a} h^D_{a,j}\right) - h^I \tag{5}$$

I can therefore estimate average schooling for independent children in parental group y as

$\hat{h}^I_{a,y} = h^D_{a,y} - \hat{\rho}$, and estimates of overall educational attainment by parental groups, $\hat{h}_{a,y}$, follow from equation (2). This yields a non-parametric estimate of the CEF $E\left[h_{i,t}|y_{i,t-1}\right]$ grouped into J bins of the parental characteristic (education or income). For parental education, I use all parental educational attainment levels available in that year's census data with some modest trimming described below, typically yielding J in the range of $10-12$. For parental income, I use $J = 10$ and divide the sample into parental income deciles based on the population adult income distribution. As long as CEFs are linear, regressions of $\hat{h}_{a,y}$ on y will yield unbiased estimates of α and β from equation (1) for each calendar year. This approach can also yield estimates of α and β in any subgroup with enough data to implement the correction.

I estimate educational IM in both parental education and parental income groups. In Appendix C, I discuss interpretation of these two different educational IM measures in relation to each other, and in relation to IM measures based on children's income, using a standard Beckerian model of parental investments in children's schooling (Solon, 2004). Under the strong assumptions of this stylized model, the coefficient from a regression of children's education on parental education equals the coefficient from a regression of children's log income on parental log income, i.e. the "intergenerational income elasticity." This result should not be interpreted literally, but only to suggest these two IM statistics may be similar in magnitude and highly correlated empirically. I find some support for this prediction across U.S. states in the main text below, as well as in survey data in Appendix C. Theory also suggests that my two measures of educational IM–in parental education and parental income–depend on different underlying parameters and therefore can in theory exhibit arbitrarily different variation over time and groups.

4. Data

I estimate IM on U.S. decennial census data from 1940 to 2000 (Ruggles et al., 2015). These censuses contains data on education, earnings, and relationship to head of all household members. I also make use of 100% complete-count census data in years 1930 and 1940 to construct a panel dataset of adult children in 1940 linked to parental characteristics in 1930 (Minnesota Population Center and Ancestry.com, 2013). I am not able to estimate educational IM in the 1950 census because it only contains income and education for one person per family. I also rely heavily on a number of additional panel datasets described below.

Educational attainment. To measure education in census data I rely on the detailed IPUMS variable EDUCD, which represents highest grade completed in all years.[7] For the "child" generations I focus on ages 26-29. At these ages most people have completed education, and experimentation in panel datasets revealed that educational mobility statistics stabilize around these ages. Table I presents highest grade attained by age and year for whites and blacks separately in census data. Education tends to increase as children age, though in several cases substantial cohort effects are apparent.

For "parent" generations I use average education of dependent children's mother and father, or education of the available parent in one-parent families.[8] I drop the small share of families with zero parental education; inspection suggested that many of these families likely represent measurement error. I also find that the bottom 2% of the parental education distribution, after excluding zeros, often yields zero or wrong-signed associations with child outcomes, and I therefore drop these parents as well.[9]

Margo (1986) documents that before 1920, many blacks and some whites, especially in the South and in rural areas, attended *ungraded schools*. For these relatively low-education individuals, educational attainment may represent years enrolled and therefore overstate highest grade attained. Note this problem does not affect CEFs with respect to parental income. Goldin (1998), consistent with earlier evidence in Denison (1985) and Folger and Nam (1967), documents an education *recall bias* whereby older cohorts report inflated high school graduation rates in the 1940 census. Under the plausible assumption that children's education correlates more strongly with actual parental education than with factors that motivate parents to exaggerate their education, this pattern would tend to flatten my estimated relationships between child and parent education in 1940, attenuating IM gains 1940-70. This problem also does not affect CEFs measured in parental income. The ability to compare educational IM estimates in both parental education and parental income illustrates another significant advantage of my approach. While other survey datasets

[7] Categories change slightly over the 1940-2000 period. I count GEDs and regular high school degrees as 12 years of schooling, associate's degrees as 14 years of schooling, college degrees as 16 years, and graduate or professional school as 17 years. Results are not sensitive to counting GEDs as 11 years, or counting "some college" anywhere in the 13-15 range.

[8] All main results are similar when I use mother's education, father's education, or head's education. I focus on average of mother's and father's education because it incorporates maximum information about parental SES while also permitting inclusion of single-parent families.

[9] Card and Krueger (1992a) also find that the bottom 2% of the education distribution behaves anomolously in a different application, and drop these observations from their sample.

discussed below do contain information on parental education for the earliest cohorts in my sample, there are no other datasets containing information on parental income in these early cohorts.

Parental income. I define parental income in census data as the sum of mother's and father's labor earnings using the IPUMS variable INCWAGE.[10] Parental income is missing or zero for a significant share of families in many years. Many zeros reflect self-employment and many also likely represent measurement error. Rather than impute income for these families, I exclude families with no reported labor income from estimates of educational mobility with respect to parental income. In the robustness section I show the main results are identical if I impute income for these families. Moreover, I show that mobility with respect to parental *education* is identical in families with missing and non-missing labor income, strongly suggesting that families with non-missing labor income exhibit mobility behavior representative of the full population. Throughout the text I focus on parental income in deciles both to facilitate comparability across years, and because schooling CEFs turn out to be more linear in parental income deciles than levels or logs.

Dependent status. I define children as "dependents" if they live in households headed by their mother or father. I therefore count children living with grandparents and other relatives as "independent" in the sense that I am not able to observe their parental income or education with confidence.[11] There is some ambiguity in dependent status of young adults in "group living" situations such as dormitories, prisons, and barracks in census data. In the robustness section I show the main results are robust to other reasonable choices in coding of dependent status.

Table II presents dependency rates by age, race and year in census data from 1940 to 2000. The table documents that over 87% of whites and over 73% of blacks still live with parents at age 17 in every year 1940 to 2000. Most "independent" children before age 18 live with grandparents or other relatives, rather than heading their their own households. As might be expected, panel data indicates that lower-SES children are more likely to be

[10] I exclude capital income because it is not available in 1940, apart from an indicator for the presence of non-wage income over $50.

[11] Children coded as "grandchildren" of heads may also live with their parents in the same three-generation household, but those parents will only be coded in relation to the head as "child" or "child-in-law." This makes parents of children indistinguishable from aunts and uncles of children in grandparent-headed households. Among children living with grandparents, in all years about 30% have no "candidate" parents in the household, 30-40% have one candidate parent, 20% have two candidate parents, and 10-20% have more than two candidate parents.

living apart from parents at these early ages. Though my methodology assumes I can link all children to parents at some age before 18, I show results are unlikely to be affected by this issue in the robustness section.

Additional Panel datasets. I incorporate several additional panel datasets both to assess the key "parallel trends" assumption underlying the empirical strategy, and to compare IM estimates on true panel data with those obtained from census data using my correction. These additinonal datasets include the Panel Study of Income Dynamics (PSID), the National Longitudinal Survey of Youth 1979 and 1997 (NLSY79 and NLSY97), the Occupational Change in a Generation 1973 survey (OCG73), and the General Social Survey (GSS). The PSID, NLSY79 and NLSY97 are panel data sets that track children after they split into new households and go back to 1968, 1979 and 1997, respectively. The OCG73 is a cross-sectional data set that collected information on adults and their retrospective parental characteristics during adolescence. The GSS is an annual cross-sectional survey that collects retrospective information on parental income and education during adolescence, and begins in 1972 for the US. I provide details of sample selection and variable choices in these datasets as well as the main census dataset in Appendix D. [12]

5. Validation of Two Assumptions

5.1. Validation of Parallel Trends

Figure II presents non-parametric visual evidence on the validity of the parallel trends assumption in the PSID, NLSY79 and NLSY97 for CEFs of children's education with respect to parental income deciles and parental education levels, pooling child ages 26-29 and restricting to whites. The assumption appears surprisingly reasonable, though clearly not perfect. In addition to being approximately parallel, the curves are not far apart from each other in levels, implying results will be relatively insensitive to the smooth cohorts assumption.

Figure II suggests that these CEFs are approximately linear. I therefore test the parallel

[12] I omit several other datasets for various reasons. The Wisconsin Longitudinal Survey does not contain information on children's dependent status (correspondence with personnel at WLS). The NLSY Original Cohorts have highly incomplete data on parental income and education. I only make limited use of OCG62 data because this survey only contains father's education, and only in 2-year bins. The Children of the NLSY79 survey still only contains children whose mothers gave birth at relatively young ages, and therefore do not yield a representative sample for IM estimation.

trends assumption more formally and quantify potential violations using regressions of the following form:

$$h_{i,y}^{j} = \alpha + \beta y + \rho 1\{j = D\} + \gamma \times y \times 1\{j = D\} + e_{i,y}^{j} \tag{6}$$

for individuals i with parental group y and dependent status j, where β captures the effect of parental group status on children's education, ρ captures a level shift in children's education across dependent and independent children, and γ captures differences in the trend in parental group status across dependent and independent children. Taken literally, the parallel trends assumption expresses the null hypothesis that $\gamma = 0$. However, the key question is not whether we reject this null hypothesis, which is inevitable given sufficient data. The key question is whether γ appears sufficiently small and stable relative to β to allow useful inferences about overall IM variation based on the subsample of children living with parents at ages 26-29.

Table III presents estimates from this regression in parental education for every available dataset with reliable information on parental education during adolescence and children's dependency status in young adulthood. Estimates of the shared slope term β are large and highly statistically significant in every sample. In contrast, estimates of the interaction term γ are typically small and often statistically insignificant. The few major exceptions– the 1990s in the PSID, blacks in the NLSY97–appear somewhat random relative to similar estimates in other datasets.

Table IV presents analogous tests for parallel trends in parental income deciles. Once again, estimates of the shared slope term β are large and highly significant, while estimates of the interaction term γ are typically small and often statistically insignificant. Moreover, the large estimate of γ for blacks in the parental education regressions is not replicated in the parental income regressions. In the results below using dependent children in census data, I estimate IM variation (across time, regions, races, and other dimensions) that is often too large to be easily explained by estimates of γ in this range .[13]

In order to test parallel trends before the 1980s, I create a panel dataset by linking children ages 10 to 17 (when most children live at home) in the 1930 census with children

[13] As a point of comparison, I have re-run these parallel trend tests in children's *earnings* at ages 26-29, rather than children's education. For this child outcome I find overwhelmingly evidence to reject parallel trends. This suggests that the evidence failing to strongly reject parallel trends in children's education is a substantive finding capturing something about reality, not an artifact of how I construct the data.

ages 20 to 27 in the 1940 census. This allows me to plot children's schooling outcomes by parental home value and rent groups (income and education are not available in the 1930 census). I also restrict to boys due to changes in surnames of girls after marriage.[14]

Figure III plots children's final schooling at ages 24-27 by parental home value and rent deciles, and for both whites and blacks. For whites, dependent and independent children at ages 24-27 have virtually identical schooling CEFs. For blacks, the parallel trends assumption also holds, though the data are noisy in higher deciles. For blacks, though not for whites, allowing for a level shift fits the data significantly better.[15] Similar patterns arise when cutting each race on region of birth. These results line up well with the results for later decades. Therefore, the parallel trends assumption is surprisingly reasonable over the entire sample period and for nearly all subgroups and datasets with sufficient power to implement a meaningful test.

5.2. Intuition for Parallel Trends

When testing the parallel trends assumption in Section 5.1, I am often unable to reject not only parallel but *overlapping* trends in dependent and independent children's education at ages 26-29 with respect to parental SES ($\rho = 0$), or in other words exogeneity of dependent status conditional on parental SES. This finding conflicts with the intuition that children often live with parents in order to finance additional education. However, this issue typically arises in studies that focus on earlier ages of young adulthood (e.g, ages 16-24 as in Card and Lemieux, 2000). By ages 26-29, children may choose to live at home for reasons that are much less directly related to education.

To explore this possibility I examine determinants of dependent status in young adulthood in the PSID and NLSY79 at different ages. In Table V I regress a dummy for depen-

[14] I link children across censuses based on five variables: year of birth, state of birth, sex, race, first name and last name. I require exact, unique matches, except for allowing year of birth to be off by one year in either direction. Out of 8.8 million children in the 1940 census, I match 2.5 million or 28%, and about 60% of these matches are unique for a final match rate of about 17%, which is similar to prior research (Long and Ferrie, 2013; Abramitzky et al., 2014). The resulting panel contains 1.5 million children aged 20-27 with outcomes observed in 1940 matched to their parental characteristics at ages 10-17 in 1930. For simplicity I forego more sophisticated matching algorithms that incorporate machine learning (e.g., Feigenbaum, 2015).

[15] Why does schooling decline so dramatically for blacks with the highest parental rent expenditures? There are very few blacks in these cells, and many of them may have reported rent incorrectly, for example reporting annual rent in place of monthly rent. This type of measurement error would generate the observed pattern, and is also consistent with the lack of a similar decline for blacks with the highest home values.

dent status on education, marital status, sex, and parental income decile as well as a full set of year dummies. Column (1) presents the results for children age 18-25 in the PSID, clustering standard errors at the individual level. All four regressors are highly significant, though marital status is overwhelmingly the most important variable both economically and statistically.[16] In Column (2), I repeat this regression for children ages 26-29. Interestingly, at these later ages the economic significance of the education coefficient is much reduced, and the R^2 of the regression falls dramatically.

As marital status remains the overwhelmingly most important predictor of dependent status at these later ages, I examine determinants of marital status in Columns (3)-(4). Once again, observable variables have statistically and economically significant impacts on marital status at earlier ages, but much weaker impacts on marital status at later ages. In Columns (5)-(8) I repeat the analysis in NLSY79 data and document similar findings. These results suggest that, among the 10-20% of children who choose to coreside with parents in their late 20s, factors only indirectly related or unrelated to education potentially play a large role. Such factors may include quality of parent-child relationships, early-career labor market shocks (Kaplan, 2012), marriage market luck[17], availability of spare bedrooms in the parental home, job matches within commuting distance of parents, and so forth.

Parallel trends with $\rho \neq 0$, as opposed to conditional exogeneity, are harder to understand from an economic perspective. What is the intuition for this specific form of endogeneity? Some insight can be gained from a simple two-type example. Let g represent a continuous measure of parental SES. Suppose there are two types of children: high-ability types H disposed toward higher levels of schooling $h_H(g)$, and low-ability types L disposed toward lower levels of schooling $h_L(g) < h_H(g)\,\forall g$. Assume both types exhibit higher schooling in higher-status parental households such that $h'_H, h'_L > 0$. Let $p_D(g) \in [0,1]$ indicate the probability a dependent child is the high type, and likewise let $p_I(g)$ indicate the probability an independent child is a high type. We can now write average schooling among dependent and independent children as

$$
\begin{aligned}
h_D &= p_D(g)\, h_H(g) + (1 - p_D(g))\, h_L(g) \\
h_I &= p_I(g)\, h_H(g) + (1 - p_I(g))\, h_L(g)\,.
\end{aligned}
$$

[16]Without marital status, the R^2 of the regression falls from 0.406 to 0.134.

[17]Observables become statistically and economically insignificant predictors of age-at-first-marriage when restricting to individuals marrying after age 26.

The parallel trends assumption requires:

$$\frac{d\left(h_D - h_I\right)}{dg} = 0, \tag{7}$$

which can be shown to imply that

$$h_H^{'} - h_L^{'} = -\frac{\left(h_H - h_L\right)^2}{\rho}\left(p_D^{'} - p_I^{'}\right) \tag{8}$$

where $\rho = h_D\left(g\right) - h_I\left(g\right) \neq 0$ equals the constant gap between parallel schooling CEFs as above. One way to satisfy this condition is with $h_H^{'} - h_L^{'} = p_D^{'} - p_I^{'} = 0$, or parallel trends with respect to parental SES for both schooling of high versus low types, and prevalence of high types among dependents versus independents. The second way to satisfy equation (8) requires violations in parallel trends in these two terms to cancel each other out.

I examine the key terms in equation (8) empirically in the NLSY79. I operationalize "high types" H empirically as children with above-median AFQT scores. Therefore let p denote the probability of exhibiting above-median AFQT, and let h denote educational attainment. The top panel of figure IV plots p_D and p_I by parental income deciles on the left, and plots h_H and h_L by parental income deciles in the right, all for children at age 27, based on parental income at age 17. The bottom panel repeats this exercise for parental education rather than parental income. These trends are roughly consistent with the notion that $h_H^{'} - h_L^{'} \approx p_D^{'} - p_I^{'} \approx 0$ for both measures of parental SES, roughly supporting the plausibility of equation (8). Of course, this result just pushes the puzzle back one step. Therefore an interesting question for future research is what types of behavioral economic models might generate parallel but non-overlapping trends as well as these additional parallel trend relationships.

5.3. Validation of Smooth Cohorts

The smooth cohort shares assumption allows me to impute parental SES shares to the pool of unallocated independent children at each age 26-29. I employ a simple method to select and validate an estimator of group cohort shares using only information about cohort shares of children under age 18, who all have high dependency rates. In this section I assume 100% dependency rates under age 18; in the robustness section I address the fact that a significant minority of children under 18 also do not live with parents. The approach

I take is to evaluate potential estimators of total group cohort shares at ages 16 and 17 using group cohort shares up through age 7. If estimators perform well at these ages when true group cohort shares are observed with accuracy, then it is more plausible that the same estimators will perform well when using group cohort shares up through age 17 to predict group cohort shares at ages 26-29.[18]

The approach is easy to understand visually. Figure V plots the share of children living with parents in different income deciles by age in the 1940 100% sample for whites and blacks. While group shares up through age 7 do not perfectly predict group shares at age 17, the deviations are economically small. This suggests that in 1950, we should be able to predict group cohort shares at ages 26-29 using group cohort shares at ages up through 17. To formalize this intuition I regress group cohort shares in cohorts age 16 to 17 on group cohort shares in cohorts younger than age 8. Tables VII and VI present the results of these regressions. Columns (1)-(3) experiment with different estimators, pooling all years 1940-2000. The simplest estimator based on group cohort share at age 7 performs better than more complex estimators. I therefore rely on this simple estimator for all main results, both due to its performance in this exercise, and because it is more stable for smaller subgroups.

Columns (6)-(12) examine this estimator by year. The estimators are highly statistically significant in every year. If group cohort shares are stable over ten-year periods, these results indicate that the composition of dependent children in cohorts age 17 contains a great deal of information about the composition of children in cohorts ages 26-27 (and probably 26-29). Second, the coefficients on the estimators are typically close to one, with some variation over time that does not line up sharply with the main patterns documented below in any particularly disturbing way. The predictions for parental education groups

[18]It might seem that I could observe parental group cohort sizes almost perfectly for children under age 28 in the prior census, since ten years previously most of these children were still dependents under age 18. This is not true for several reasons. First, both income and education are not observed in 1930, preventing the use of this method to estimate parental group cohort sizes in 1940. Since CEFs cannot be estimated in the 1950 census, it is critical that I develop a method that can be applied to the 1940 census to study long-term mobility trends in the U.S. context. Second, parental group status may change in systematic ways over ten-year intervals. For example, parents of 16-17 year-olds in the bottom income decile in 1960 may not systematically be in the bottom income decile as parents of 26-27 year-olds in 1970. This consideration is less important for parental education, but still may exist due to survey methodology variation or reporting biases discussed above. A less serious problem is that ten years of death and international migration take place between censuses. This problem would be small in my application because few 16-17 year-old children die before turning 26-27 during this period, and because I restrict samples to native-born children.

are somewhat better than for parental income groups in the sense of having coefficients close to one and high R^2, though both are quite good. Tables A.1 and A.2 display similar patterns for black children. These results suggest the smooth cohorts assumption is quite reasonable. Moreover, in the robustness section below I find that the main results are not highly sensitive to plausible degrees of mismeasurement in group cohort shares.

6. Results

6.1. Historical Trends

Figure VI presents the two estimated CEFs in 1940 before and after the correction for independent children. The correction turns out to affect levels much more than slopes because dependency rates turn out to be roughly stable across parental groups at ages 26-29 in census data. The correction also affects blacks more than whites due to the larger share of young adult blacks living independently. The figure illustrates the fact that relative mobility is strongly correlated with absolute upward mobility because the CEFs "pivot" at high levels of parental income and education, and that education is approximately linear in parental education and parental income rank.[19] These patterns are similar to findings in Chetty et al. (2014a), but 60 years earlier in time.

In order to summarize many CEFs parsimoniously I estimate intercepts and slopes using the simple regression equation:

$$h_{y,t} = \sum_{t=1940}^{2000} \alpha_t \cdot 1\{\text{year} = t\} + \sum_{t=1940}^{2000} \beta_t \cdot 1\{\text{year} = t\} \cdot y + \varepsilon_{y,t}, \tag{9}$$

where $h_{y,t}$ represents a child outcome measure in census year t for children in parental group y (either education or income). I focus primarily on the slope coefficients β_t as measures of relative IM, because intercepts largely reflect secular trends in schooling. When I re-estimate equation (9) in ranks below, intercepts provide more useful measures of absolute upward IM separate from secular trends in education levels. I estimate equation 9 on data collapsed to the level of year $\times$ parental SES group $\times$ race.[20] Given that I am collapsing

[19] Note that dropping the bottom 2% of the parental education distribution as discussed in Section 4 is important for achieving approximate linearity of CEFs in parental education.

[20] National results pooling races are very similar to the results for whites only, as expected given that approximately 90% of the U.S. population is white over the sample period.

millions of observations into dozens of bins, standard errors are likely conservative relative to those that would be obtained from similar regressions in microdata even with various forms of clustering.

Tables VIII-IX display estimated intercepts and slopes of CEFs in parental education levels, for whites and blacks separately. Tables X-XI present analogous estimates for mobility in parental income decile. Column (1) from these tables contains estimated intercepts and slopes for whites and blacks. For whites, the slope in parental education falls from 0.50 in 1940 to 0.39 in 1960, or about 20%, and remains relatively stable up through 2000.[21] The slope in parental income deciles similarly falls from 0.37 in 1940 to 0.25 in 1960, or by about 25%. The post-1940 mobility gains of black children are especially remarkable, with slopes in parental income and education both falling by over 50%. These results show that 20th century black-white economic convergence (see, for example, Smith, 1984; Margo, 1986) can be understood not only as blacks converging toward whites, but as poor blacks converging toward rich blacks, and rising mobility more generally.[22] These post-1940 mobility gains are economically large, potentially contributing about 0.25 percentage points of aggregate earnings growth over the 1940-70 period.[23]

How do these estimates compare to estimates in panel/retrospective datasets that do not require any adjustment for independent children? Figure VII plots the IM estimates in parental education from Column (1) in Tables VIII-IX alongside estimates from many other sources. For comparability I average children's education at ages 26-29 in these other datasets across cells defined by decade $\times$ parental SES $\times$ race, and then estimate intercepts and slopes on these aggregated data. I define decade "1980", for example, to include cohorts turning ages 26-29 in 1980-89. Note that survey data estimates for cohorts in decades before 1970 rely entirely on retrospective reports from increasingly elderly respondents, and may

[21] Formal tests for equality of parameter estimates across years with very different point estimates generally yield p-values well below 5%. In particular, slopes in 1940 are always statistically different from slopes in 1960-80.

[22] Recall that I am collapsing the data into population distributions of parental income and education, so differences in black and white IM statistics reflect genuine differences in the shapes of estimated CEFs, not reversion to different group means along identical and nonlinear CEFs.

[23] To see this, consider the impact of the increase in relative mobility with respect to parental income. Suppose relative educational mobility in 1970 remained at the 1940 level, so that schooling at the top decile in 1970 were held constant at its observed value but schooling of all lower deciles were decreased to reflect the steeper slope from 1940. This would reduce average schooling in 1970 by about 0.75 years. If annual earnings increase by 10% for each additional year of education, this change would account for 0.25 percentage points of aggregate earnings growth over the 1940-1970 period. I estimate that total household earnings over this period grew at an annual rate of 3.4%, suggesting increasing relative mobility increased annual earnings growth by 8% over this high-growth period.

therefore suffer from mortality attrition (Costa, 2015) and recall error (National Center for Education Statistics, 1984).

Reassuringly, Figure VII indicates that estimates from these various datasets are similar in magnitude and exhibit a decline in slopes after 1940 that is larger for blacks. However, there are also some important discrepancies. First, the decline in IM in census data is somewhat larger over the 1940-60 period (ignoring 1930) than that in the OCG73 and GSS55, for both whites and blacks. This pattern could easily be generated by selective mortality of lower-SES respondents at older ages. Second, the overall magnitude of the slopes in most survey datasets appear somewhat higher in most years for whites, relative to census estimates, suggesting different sample selection or additional measurement error. Most importantly, survey datasets strongly suggest a substantial increase in slopes for whites after 1980 that is not captured by census estimates. A decrease in educational mobility since 1980 would be consistent with prior work documenting post-1980 increases in gaps between high-income and low-income children in educational attainment (Acemoglu and Pischke, 2001; Belley and Lochner, 2007; Bailey and Dynarski, 2011) as well as test scores (Reardon, 2011). [24]

What explains this post-1980 discrepancy in time trends? Inspection of underlying CEFs reveals that children of parents with college degrees or higher in census data have conspicuously "too little" education in many decades compared both to similar children in panel datasets and to children of less educated parents in census data. If I exclude children of parents with college and graduate degrees, the estimated trend in slopes remains similar up through 1990 but exhibits a significant increase 1990-2000 that is consistent with the time trend in panel data (increasing from 0.39 in 1990 to 0.47 in 2000, nearly as high as the slope of 0.51 in 1940). I conclude that the imputation method is accurately detecting a post-1940 mobility increase, but failing to detect a decrease in mobility in recent decades due to a particular violation of parallel trends by children in the highest-education families.

Figure VIII plots IM estimates in parental *income deciles* from Column (1) in Tables X-XI. There is no other nationally representative dataset containing credible information about parental income for cohorts who reach their late 20s before 1968. It is therefore both novel and reassuring to document that IM with respect to parental income also increases after 1940 for both races, and more so for blacks. Once again the census estimates appear

[24] As discussed in the literature review, most research on *income-based* IM has found no trend since 1980 (Hertz, 2007; Lee and Solon, 2009; Harding et al., 2009; Chetty et al., 2014b), with the exception of Aaronson and Mazumder (2008).

to miss a possible decline in mobility after 1980, although estimates from other datasets are less clear on this point.

So far I have focused on estimates of IM based on education levels. Increases in IM based on education levels may reflect changes in educational inequality across generations rather than changes in positional transition matrices. In Appendix B, I adjust for changes in educational inequality by estimating two alternative educational IM statistics that are not subject to this concern: (1) rank-rank elasticities in parental education and income, and (2) correlation coefficients in parental education levels. While I find that the parallel trends assumption looks more tenuous in education ranks, results for relative educational IM based on ranks are very similar to results already discussed. Moreover, intercepts of rank-rank CEFs can be interpreted as measures of absolute upward mobility distinct from secular gains in education (Chetty et al., 2014a), and I find that by these measures absolute upward mobility increased modestly for whites and dramatically for blacks after 1940 much like relative mobility. The broad finding of large post-1940 IM gains are therefore quite robust to the choice of IM statistic and are not driven by changes in educational inequality across generations.

6.2. Heterogeneity

Even if the method provides estimates of post-1980 national time trends in educational IM that are apparently biased, it may still shed light on cross-sectional variation in IM across places and subgroups in all years if biases are similar across subgroups within years. To examine this possibility as it applies to spatial variation in IM, I compare my educational IM estimates with newly-available estimates of income-based IM across states in the 2000s. Chetty et al. (2014a) have recently made available rank-rank income CEF slopes and intercepts by "commuter zone" (CZ) in the U.S. using the population of U.S. tax records spanning 1996-2012. I average their income-based rank-rank slopes up to the state level.[25] I then construct education rank-rank slopes on census data by state, adjusted to account for independent children, as described in Appendix B. Note that Chetty et al. (2014a) measure children's residential location around age 15. I can either measure children's location at

[25]The resulting state-level intercepts and slopes should come close to what would be obtained from a rank-rank regression on state-wide micro-data. Such a regression would average the CZ-level slopes with weights proportional to the variance of parental income rank in each CZ (Angrist and Pischke, 2009). I use unweighted averages, although I experimented with weighting by CZ-level Gini coefficients and interquartile ranges as two proxies for parental income rank variance, and found these alternative weights had virtually no impact on the results.

ages 26-29, or at time of birth. I choose time of birth because many individuals will have left their home states as of ages 26-29.[26] This comparison is far from ideal because educational IM and income-based IM measure different concepts. Unfortunately, the tax data do not contain educational attainment and I am unaware of any other dataset containing state-level IM. As discussed above, however, economic theory suggests educational and income-based IM may be closely linked.

Figure VII plots educational rank-rank mobility estimates from census data against income rank-rank mobility estimates from tax data. The correlation between the intercepts in Panel (a) is 0.61, while the correlation between the slopes in Panel (b) is 0.55. These results strongly suggest that the method here is indeed capturing genuine variation in IM across states in the recent period, and is therefore likely useful for studying spatial variation in IM in earlier years. This finding is valuable because there is currently no alternative dataset available for estimating IM across states in earlier decades.

I now present estimates of educational IM across various subgroups by year from 1940 to 2000 in Tables VIII-IX. Given the results above, time trends should be interpreted with some caution, especially after 1980. I focus on relative educational IM estimates with respect to parental education, noting any important differences for results based on parental income deciles. Columns (2)-(3) show IM gains were similar for boys and girls. Columns (4)-(5) break out results into South and North, where "North" includes all non-southern regions. These results replicate lower IM in the South in the 2000s (Chetty et al., 2014a), but show this gap is quite small in longer-term historical context. Surprisingly, IM gains in the South were similarly large for whites and blacks with respect to parental education, and only slightly larger for blacks with respect to parental income. Given that over 75% of blacks in 1940 lived in the South compared to 25% of whites, these regional convergence trends account for most of the larger mobility gains of blacks nationally. Columns (6)-(7) of Tables VIII-IX compare mobility in areas defined by the census as "Urban" and "Rural" in years where this variable is available. Both absolute and relative mobility are higher in urban areas for both whites and blacks. The urban/rural mobility gap has also tended to decline over time, much like the North-South mobility gap.[27]

[26]Note that state-level schooling mobility estimates in census data are quite noisy because IPUMS only provides small random samples after 1940, the education distribution is much lumpier than the income distribution, and the state size distribution is skewed. To increase precision, I therefore average the separately-adjusted CEFs from the 1980, 1990 and 2000 censuses into one pooled CEF for each year, race and state of birth before estimating slopes and intercepts.

[27]The one exception to this convergence pattern is that the urban/rural mobility gap increased for blacks

Columns (8)-(9) of Tables VIII-IX report estimated intercepts and slopes for CEFs of *annual enrollment* at "high school ages" 16-18 and "college ages" 19-21, rather than final educational attainment at ages 26-29 as in columns (1)-(7). Note that the vast majority of children live at home at ages 16-18, implying much weaker concerns about representativeness of dependent children. For both whites and blacks, in both parental income and education, high school enrollment accounts for *all* of the post-1940 increase in relative educational mobility. In contrast, relative mobility in terms of college enrollment exhibits no decline over this period, and once again fails to detect a post-1980 decline in college enrollment mobility that others have previously estimated in survey data (e.g., Lochner and Monge-Naranjo, 2011; Bailey and Dynarski, 2011; Belley and Lochner, 2007). This long-term pattern suggests that low-SES children have plausibly never successfully gained ground on high-SES children in *college* attainment since the start of the 20th century, in sharp contrast to high school attainment. These findings on high school enrollment mobility provide strong additional evidence that post-1940 mobility gains are not driven by changes in selective dependency decisions.

7. Determinants of Trends in Educational IM

What accounts for the large gains in educational IM after 1940? Findings above suggest the importance of increasing supply of and/or demand for *high school* education for cohorts reaching high school in 1930-1950 (born ~1915-35), especially factors affecting all genders and races in the South.

These findings immediately cast doubt on several potential explanations. Post-1940 G.I. Bills almost exclusively benefited men, and yielded few benefits for southern blacks (Turner and Bound, 2003). School integration reforms beginning in the late 1950s and the Civil Rights gains of the 1960s arrive too late to explain the time trend in IM, and also cannot explain major IM gains for southern whites. The Great Migration and other regional population shifts are also hard to square with black and white mobility both rising to similar extents *within* the South and *within* rural areas.[28] For similar reasons, the belated black high school movement (Anderson, 1988) is not a likely explanation, and

between 1940-60 before converging 1960-1990. Note that intercept and slope estimates for rural blacks in 1960 only cover income deciles 1-7 due to a lack of sufficient high-income rural blacks.

[28] Holding constant state of birth (or state of residence) population shares at their 1940 level has almost no impact on national mobility trends for whites or blacks. Of course, migration may have played an important role in sustaining wage gains in the South by tightening labor markets (Wright, 1986).

in Appendix E I use newly-digitized data on the spread of black high schools to document additional evidence casting doubt on this explanation. The Great Depression could have temporarily constrained educational demand among lower-SES families, leaving room for "catch-up" after 1940; however the Great Depression was not a disproportionately southern phenomenon (Rosenbloom and Sundstrom, 1999).

Growth, Inequality, and Other Factors: State-Level Panel Data Analysis. Prior literature suggests many other factors that may relate to the observed mobility trends, including per-capita income levels and inequality (Loury, 1981; Becker and Tomes, 1986; Murphy et al., 1991; Galor and Tsiddon, 1997; Owen and Weil, 1998), urbanization (Goldin, 1998), black population shares (Margo, 1990; Card and Krueger, 1992b), teen birth rates (Edin and Kefalas, 2011; Kearney and Levine, 2012), compulsory schooling laws (Acemoglu and Angrist, 2001; Lleras-Muney, 2002), educational inputs (Card and Krueger, 1992b,a; Chetty et al., 2011; Jackson et al., 2015), the demand for teen labor (Cogan, 1982; Margo and Finegan, 1993), and migration (Long and Ferrie, 2013). Many of these factors changed simultaneously during the "transformation" of the southern economy after 1930 (e.g. Wright, 1986, 1987).

In order to explore basic descriptive correlations of IM with this wide range of factors, I construct a long-term panel dataset of IM statistics by state of birth and year. In order to probe robustness of these correlations I estimate three bivariate regression models for each explanatory variable: OLS, fixed effects (FE) and first-differences (FD). Formally, for mobility statistic $M_{s,t}$ and covariate $X_{s,t}$ calculated on individuals born in state s and ages 20-29 in year $t \in [1940, 2000]$, I estimate the following regressions:[29]

$$\begin{aligned} M_{s,t} &= \beta_{OLS} X_{s,t} + e_{s,t} \text{ (OLS)} \\ M_{s,t} &= \beta_{FE} X_{s,t} + \gamma_s + \lambda_t + \epsilon_{s,t} \text{ (FE)} \\ M_{s,t} - M_{s,t-1} &= \beta_{FD} \left(X_{s,t} - X_{s,t-1} \right) + \eta_t + \varepsilon_{s,t} \text{ (FD)} \end{aligned}$$

Even with controls for state and year effects, the estimates here cannot be interpreted as causal determinants of IM. Nonetheless, the panel dimension of these data are new and therefore enrich descriptive understanding of IM variation in the U.S. context.

Table A.6 presents summary statistics for the dependent and independent variables in

[29] I combine ages 20-29 in this section, as opposed to 26-29 in the main section, in order to maximize statistical power.

the analysis for whites and blacks separately. The immense variation in mobility and all independent variables highlights the novel 60-year timeframe of the analysis and the wide spatial heterogeneity of the United States. I focus on mobility $M_{s,t,}$ defined as slopes with respect to parental education.[30] In order to characterize these variables in years that are most plausibly related to IM, I match them to IM statistics by the year in which children underlying the IM statistics would have been teenagers when possible, and otherwise by the year in which these children were in their 20s.

Table XII presents results separately for whites and blacks. For both races, higher IM is robustly associated with higher state earnings, lower inequality, lower black population share, higher minimum school dropout age, higher relative teacher wages, and to some extent higher out-migration rates, while correlations appear less robust for share urban, teen employment rates, and class size . [31] In many cases, coefficients are similar even when including state and year fixed effects. Most of these findings are consistent with correlates of income IM in Chetty et al. (2014a) for the 2000s, including large correlations with black population share (for both whites and blacks), inequality, and K-12 school quality. The negative association between inequality and IM has been termed the "Great Gatsby Curve" (Krueger, 2012; Corak, 2013), and to my knowledge has only been documented previously in cross-sectional data in the modern period. Results here extend this result to mid-20th century variation across U.S. states, and show that the association remains significant in panel data regressions controlling for state-of-birth and year fixed effects. The estimates suggest that a 10 percentage point increase in the interquartile household earnings gap within a state-of-birth group has historically decreased relative educational IM $(1-\beta)$ by .03, suggesting that the post-1940 decline in inequality (Goldin and Margo, 1992) by itself would predict a large share of observed post-1940 educational IM gains.

The most striking discrepancy with earlier work is that I find a robust positive correlation of educational IM with state household earnings, whereas Chetty et al. (2014a) find a near-zero correlation of IM with mean household income levels in the 2000s. In Table XIII, I probe this result further by regressing IM *jointly* on state household earnings and household earnings inequality, using IM defined in terms of parental education and parental

[30] Results are similar for absolute mobility (intercepts of CEFs), and for CEFs in parental income deciles.

[31] The results presented above almost all become underpowered if I allow for linear time trends that vary by state. Some of the results survive time trends that vary by region. The results are largely robust to dropping 1940, and dropping the South, but become underpowered when dropping both 1940 and the South together. The results become more significant if I weight by the precision of the regressions used to estimate the mobility statistics.

income separately. While precision declines, the point estimates are robust and in some specifications both variables retain statistical significance. Therefore *broad-based economic growth*, rather than just growth *per se,* is historically associated with educational IM.

One plausible reason for this discrepancy with earlier work is that, over the longer-run sample period examined here, state household earnings levels correlate with broad processes of industrialization, urbanization, and political liberalization, particularly in the South (Wright, 1986). There are many reasons why the "Old South" may have yielded lower educational IM for both whites and blacks. Political institutions in the South actively excluded poor whites as well as blacks using poll taxes, literacy tests, and other impediments to voting (e.g., Husted and Kenny, 1997), and many states loosened these restrictions by the 1950s well before the Voting Rights Acts of 1964 and 1965.[32] Wright (1986) presents anecdotal evidence that many elite southerners opposed large investments in public education. Perhaps as a consequence of these political factors, the South spent less on public schooling, had lower-quality public schools by many other measures, and had much lower high school graduation rates than other regions of the U.S. throughout the first half of the 20th century, and many of these differences decline dramatically across cohorts attending high school between 1930 and 1950-60 when southern IM increases in my data (Ayres, 1920; Wright, 1986; Card and Krueger, 1992a; Goldin, 1998). An interesting question for future research, then, is whether broad economic and political transformations in other times and places (e.g., Taiwan, South Korea, Botswana, etc.) have also been associated with improvements in IM.

High School, College and Educational Institutions. The fact that high school attainment has equalized across SES groups over time, while college attainment has if anything become more unequal, points to a potential role for institutional differences in these two educational margins. High schools, in contrast with colleges, are characterized by public finance, compulsary initiation, and automatic enrollment. Prior work on college access documents the importance of financial obstacles (Dynarski and Scott-clayton, 2013; Fack and Grenet, 2015) and frictions arising from voluntary and active enrollment (Dynarski and Scott-Clayton, 2008; Bettinger et al., 2012; Carrell and Sacerdote, 2013), while other work shows that defaults can be sticky even in high-stakes choice environments (e.g., Carroll et al., 2009; Chetty et al., 2013). It is therefore possible that institutions governing college

[32]Florida, Georgia, Louisiana, and North Carolina dropped their poll tax in 1945, followed by South Carolina and Tennessee in 1951.

education in the U.S. are less amenable to educational IM than institutions governing high school education.

A different explanation points to quality problems in the K-12 system that may affect low-SES children disproportionately. Table XII documents a strong association between relative teacher pay and IM. Relative teacher pay and teacher quality have declined dramatically since 1940 (Hoxby and Leigh, 2004; Bacolod, 2007), and high school graduation rates have been stagnant since the 1960s (Heckman and Lafontaine, 2010). It is therefore also possible that the public primary school system successfully prepared many low-SES children to attend high school in the 1940-70 period, but the public K-12 system in recent decades has failed to prepare low-SES children to attend college.

8. Robustness

Missing/zero parental income and education. There is substantial variation in the fraction of children with zero or missing values for parental characteristics. Table A.8 displays this variation for whites and blacks, restricting to dependent children age 26-29. The fraction of parental education values that are missing or zero is small in all years for both whites and blacks. The fraction of missing parental income observations follows a U-shaped time trend for both whites and blacks, and therefore does not seem likely to explain the observed decline in mobility. The fraction of children with parents reporting zero income raises more serious concerns, as these shares are very high in 1940 and fall significantly and steadily over time.

While it is reassuring that educational IM tells the same basic story in both parental education and parental income, I examine robustness of results using parental income in three ways. First, I impute income for families with missing or zero reported income, recalculate population income deciles, and include these families in the analysis.[33] Table XIV replicates my basic results in columns (1)-(2) and presents the results based on imputed income in columns (3)-(4). The main patterns of rising IM after 1940 and larger IM gains

[33] I implement this by calculating average household earnings among all individuals with non-zero earnings in cells defined by year, race, age, sex, education, marital status, and state or country of birth. I assign cell means to individuals with zero and missing earnings based on this set of characteristics, roughly following Autor et al. (1998). In households with two earners I take the maximum of these two predictions. For households with zero earnings I follow Neal (2006) and adjust for selection by multiplying imputed household earnings by 0.6. While this method is obviously *ad hoc*, it provides a rough check on whether households with zero and missing earnings are likely to be driving the main results.

for blacks are unchanged. Second, I examine mobility with respect to parental education separately for families reporting zero and positive income. If IM with respect to parental education is similar in these two groups, it suggests that the excluded families are unlikely to exhibit such different IM from the included families as to generate substantial bias in the main results. Figure A.6 plots education elasticities for children with positive and zero/missing parental income, by year, for whites and blacks separately, and indicates that mobility patterns are indistinguishable in these two samples.

Third, basic calculations suggest the increase in mobility for whites is too large to be accounted for by the decline in the share of families reporting zero income. Note that the share of families reporting zero income falls by 14 percentage points between 1940 and 1960. The worst-case scenario is that these 14 percentage points of families are "perfectly mobile" with relative IM of zero. In that case the true 1940 slope would actually be $(0.14\,(0) + 0.86\,(0.5) =)$ 0.43, which is still higher than the estimated slope of 0.39 in 1960. It should be clear from the result in Figure A.6 discussed above that perfect mobility for this group is so conservative as to be implausible. Finally, note that the trend in share of missing and zero income is similar for whites and blacks, despite the large differences in white and black estimated mobility trends.

Definition of dependent status. There is some ambiguity in dependent status of young adults in "group living" situations such as college dormitories, prisons, and military barracks in census data (e.g., Bureau of the Census, 1988; National Research Council, 2006, p. 47). For my primary results I count all children living in dormitories, prisons and military barracks at ages 26-29 as independents. Figures A.4 and A.5 compare the estimated slopes and intercepts of education CEFs in parental income and education for the primary sample and an alternative sample that excludes children in "group living" situations. The results are nearly identical with the one exception of an anomalously flat slope of the education CEF in 1970, which reflects an oddly low level of estimated final schooling among children of high-education parents in that year.

Missing dependents at age 17. I rely on cohorts of age 17 in each year to assign the share of independent children to parental SES groups in cohorts of ages 26-29. However, as discussed above, a significant minority of children are not living with identifiable parents at age 17, especially in earlier years and especially among black children. Examination of this issue in panel datasets revealed, unsurprisingly, that low-SES teenagers are less likely to be living with parents. In order to explore the sensitivity of the results to this type of problem, I scale up the number of 17-year-olds in lower-SES groups for my imputation, using the simple formula $N_y\left(1+0.4\times\left(1-\frac{y}{y_{\max}}\right)\right)$ where N_y denotes the number of dependent children in a given year and race in parental group y and $\frac{y}{y_{\max}}$ rises linearly in parental SES group from 0 to 1. I therefore increase the relative number of 17-year-olds in lower-SES groups by a large amount. Columns (5)-(6) in Table XIV present the results based on parental income decile groups. Estimates are barely affected.[34]

9. Conclusion

In this paper I develop a new method to estimate educational IM on cross-sectional U.S. census data. The method overcomes the problem that most children cannot be linked to parents by ages of school completion, and thereby allows for estimation of final educational outcomes by parental income and education. The new methodology in conjunction with multiple additional datasets yields several important new historical facts. Educational IM increased significantly after 1940 (1911-14 birth cohorts) before stabilizing and then declining after 1980 (1951-54 birth cohorts). Post-1940 educational IM gains plausibly increased aggregate annual earnings growth by 0.25 percentage points over the 1940-70 period. IM gains were particularly large in the South for both whites and blacks, implying larger IM gains for blacks nationally due to their greater geographic concentration in the South.

The increase in relative educational IM after 1940 stemmed from greater high school enrollment, rather than college enrollment. The GI Bills, the Civil Rights Movement, school desegregation, the black high school movement, and the Great Migration do not account for post-1940 IM gains. I construct a novel long-term panel dataset by year and state of birth to explore descriptive determinants of IM. As Chetty et al. (2014a) find for income-based IM across places in the 2000s, I find that educational IM (for both whites and blacks) correlates with black population shares, income inequality, and educational quality, even conditional on state and year fixed effects. Unlike Chetty et al. (2014a), I find a robust positive association of state income levels with educational IM. This discrepancy likely stems from the much larger forces of modernization accompanying economic growth after 1940, particularly in the South where lower-SES voters gained political power and K-12 public school input gaps narrowed dramatically. I also point out that institutional differences in

[34]This adjustment also barely affects IM estimates based on parental education groups, which I omit for brevity.

the K-12 and college systems including public finance, defaults and compulsion, as well as quality problems in the K-12 system, may play a role in the contrast between post-1940 gains and post-1980 declines in educational IM. However, it is also possible that post-1980 declines in IM represent a mechanical "aftershock" of the same forces underlying post-1940 gains in IM (Nybom and Stuhler, 2014).

Identifying causal determinants of long-term trends in educational IM therefore remains an important task for future research. The methods developed here can further this research by expanding IM measurement to times, places and groups that are only accessible in large, cross-sectional household datasets such as population censuses.

NOT FOR PUBLICATION

A. Mobility Statistics in a Model of Parental Borrowing Constraints

How do educational mobility statistics relate to earnings mobility statistics? And how does mobility with respect to parental education relate to mobility with respect to parental income? In this section I derive steady-state expressions for these different mobility statistics in a stylized model of family borrowing constraints and educational investments developed in Solon (2004) based on Becker and Tomes (1979, 1986), generalizing the setup in Solon (2004) slightly to allow for heritable determinants of child income other than human capital (e.g., family connections).

Let a parent with one child maximize a Cobb-Douglas utility function

$$U_i = (1-\alpha)\ln C_{i,t-1} + \alpha \ln y_{i,t} \tag{10}$$

where i indexes individuals, t indexes a generation, $C_{i,t-1}$ denotes parent's own consumption, $y_{i,t}$ denotes the child's future pre-tax income, and α governs the trade-off between own consumption and children's income. The parent maximizes utility subject to a budget constraint

$$(1-\tau)\cdot y_{i,t-1} = C_{i,t-1} + I_{i,t-1} \tag{11}$$

where τ is the average and marginal tax rate on parental income, $y_{i,t-1}$ denotes parental pre-tax income, and $I_{i,t-1}$ denotes financial investments in children's human capital. These financial investments yield decreasing marginal returns subject to the human capital production function

$$h_{i,t} = \delta + \theta \ln\left(I_{i,t-1} + G_{i,t-1}\right) + e_{i,t} \tag{12}$$

where δ represents the minimum schooling level in society, θ represents the productivity of financial investments in human capital, $G_{i,t-1}$ represents government spending on human capital of child i, and $e_{i,t}$ captures human capital transmitted to children from parents through channels other than financial investment. Assume that government education spending is allocated progressively such that

$$\frac{G_{i,t-1}}{y_{i,t-1}} \approx \varphi - \gamma \ln\left(y_{i,t-1}\right), \tag{13}$$

where φ indicates the universal subsidy as a share of income, and γ captures progressivity of the subsidy schedule.

Assume a log-linear earnings equation in schooling in the tradition of Mincer:

$$\ln y_{i,t} = \mu + p h_{i,t} + \varepsilon_{i,t} \tag{14}$$

where p indicates the return to schooling, μ is the minimal income level in society, and $\varepsilon_{i,t}$ captures income transmitted to children from parents through channels other than observed human capital.

Let heritability of both $e_{i,t}$ and $\varepsilon_{i,t}$ be governed by the same AR(1) process such that

$$e_{i,t} = \lambda e_{i,t-1} + \nu_{i,t} \tag{15}$$

$$\varepsilon_{i,t} = \lambda \varepsilon_{i,t-1} + u_{i,t} \tag{16}$$

where λ indicates the degree of human capital and income inherited from parents outside of monetary investment channels. The assumption that one parameter governs both these inheritance processes is made for analytical convenience.

As pointed out in Becker and Tomes (1986), if parental income y_{t-1} exceeds a certain cut-off level, then parents in this model will leave financial bequests to children, and marginal parental income has no causal effect on children's human capital or income.[35] I assume that parental income is below this threshold, such that parents wish to borrow from their children's future income but are prevented from doing so by a complete failure of the human capital loans market.

Assume all dynasties are in steady state. Using known results on autoregressive models (Greene, 2002, pg. 266), and letting $\beta_{x,x'}$ denote the OLS coefficient from a regression of x on x', it can be shown that

$$\begin{aligned} \beta_{h_t,h_{t-1}} &= \frac{p\theta\,(1-\gamma)+\lambda}{1+p\theta\,(1-\gamma)\,\lambda} \\ &= \beta_{\ln y_t,\ln y_{t-1}}. \end{aligned}$$

This result suggests that intergenerational education and income elasticities reflect similar underlying features of social systems and should be similar in magnitude. While this result cannot be taken literally for many reasons[36], I find it is roughly consistent with the data. My estimated education elasticities since 1980 (around 0.4) are quite similar to

[35] As the model is written, parents cannot increase child income directly with bequests because human capital is the only instrument available for transfers. To add savings, augment the parent's budget constraint to $y_{i,t} = C_{i,t} + I_{i,t} + S_{i,t}$ and augment the child's earnings function to $\ln(y_{i,t} - S_{i,t}) = \mu + ph$. In this extended model, for parental income above a critical value saving is positive and parental income has no causal impact on children's schooling but still has a positive regression coefficient due to the non-financial transmission parameter λ, as expected.

[36] For example, changes in the distribution of earnings Autor et al. (2008) and the curvature of earnings functions (Lemieux, 2006; Heckman et al., 2006; Goldin and Katz, 2010) directly violate the steady state and functional form assumptions of the model.

intergenerational income elasticities in prior literature (Solon, 1999). Figure IX.b is also consistent with this result, establishing a strong correlation between rank-rank education and income elasticities across U.S. states in recent decades. In Appendix C, I provide additional support for a historical link between educational and income mobility using the OCG62 and OCG73. Of course, educational and income mobility are clearly distinct objects and both are interesting in their own right.

Again assuming steady state, it can also be shown that slopes of education CEFs in parental education and parental income are related by:

$$\beta_{h_t,\ln y_{t-1}} = \frac{1}{p}\left(\beta_{h_t,h_{t-1}} - \frac{\lambda}{1-\lambda\theta\left(1-\gamma\right)}\sigma_\varepsilon^2\right) \tag{17}$$

where σ_ε^2 denoting the variance of income conditional on human capital. Due to this additional parameter, comparative statics of $\beta_{h_t,\ln y_{t-1}}$ are more ambiguous than for $\beta_{h_t,h_{t-1}}$. In this simple model, the finding that these two mobility statistics exhibit similar historical trends does contain some empirical content, suggesting for example that trends are not driven by changes in σ_ε^2.[37]

B. Mobility or Inequality? Rank-Rank Elasticities and Correlation Coefficients

Above I focus on educational elasticities for convenience and ease of interpretation. However, elasticity trends depend both on trends in positional mobility and trends in cross-sectional inequality. For some purposes, we may wish to measure these two social objectives–mobility and inequality–with statistics that are *mechanically* independent. I estimate two such additional mobility statistics: rank-rank elasticities, and correlation coefficients.

I use the same method developed above to estimate intergenerational elasticities in education *ranks*. Educational attainment can be mapped into ranks by choosing a method to resolve ties; I choose the midpoint of the probability mass interval occupied by an educational category. I rank parental education (average of mother plus father, calculated as described in text) separately by year, and I rank children's education during ages 26-29 separately by age and year. I continue ranking parental income as in main text. I re-test the parallel trends assumption for education ranks in Tables A.3-A.4. The assumption appears approximately true for education ranks in parental income ranks, though education ranks in parental education ranks look more problematic, with dependent children often displaying statistically and economically flatter slopes (i.e., higher mobility). These violations suggest that the census method will likely overstate mobility overall, and may overstate mobility for blacks, in recent decades. These problems must be kept in mind when interpreting trends in this section.

[37] I have confirmed empirically that the conditional variance of income does not evolve in a way that would explain trends in $\beta_{h_t,\ln y_{t-1}}$.

Despite these problems, education ranks afford two advantages: they facilitate comparison of IM over time as the underlying distributions of educational attainment evolve, and they allow interpretation of intercepts as measures of absolute upward mobility distinct from secular gains in education. The disadvantages are that rank-based IM cannot be interpreted in terms of human capital units, and ranks can be unstable for discrete random variables with lumpy distributions such as educational attainment.

Table A.5 displays rank-based absolute and relative mobility estimates in census data that are analogous to those in Tables VIII-IX. For whites (Columns 1 and 3), rank CEFs in parental income suggest limited gains in absolute upward mobility, but significant gains in relative mobility. Rank CEFs in parental education indicate no gains in absolute upward mobility, but do suggest gains in relative mobility that are roughly consistent with results for education levels. For blacks (Columns 2 and 4), rank CEFs exhibit large improvements in both absolute upward and relative mobility over time, especially during the 1940-60 period.

While ranks do a better job than levels at distinguishing mobility trends from inequality trends, due to the lumpiness of the education distribution ranks are still not conclusive in this respect. The intergenerational *correlation* also abstracts from changes in educational inequality, and equals the intergenerational elasticity multiplied by the ratio of the standard deviation in parental education over the standard deviation in children's education. I construct these standard deviation ratios for whites and blacks in every year, and adjust the estimated elasticities accordingly. Figure A.1 displays the time trends in child-parent educational correlations for both whites and blacks.[38] The trends are similar to those displayed in Figure VII for education elasticities.

C. Did Higher Educational Mobility Lead to Higher Income Mobility?

How have changes in educational mobility translated into income mobility? Validation of state-level estimates against income mobility in tax data indicates the answer is yes for recent cohorts (see Section 6). Here I also address this question for earlier cohorts in the OCG1962 and OCG1973 surveys, which contain parental education levels and larger samples than the PSID or NLSY79. I first ask if education affects income of children from different parental groups in similar ways. If that were the case, it would suggest we can link these two concepts together with this shared return to schooling as assumed in the model of Appendix (A).

To proceed, I decompose children's earnings into three factors: returns to education, returns to parental group status unrelated to education, and differential returns to education

[38]For correlation coefficients, the conservative standard errors based on collapsed data used above result in confidence intervals that are too wide to be useful. In this case I therefore make a less conservative assumption of a Moulton structure in children's education with respect to parental education, and adjust standard errors of correlation coefficients with Moulton factors by race and year.

by parental group status, by estimating regressions of the form

$$\text{logEarnings}_{g,\text{educ}} = \alpha + \beta \cdot \text{educ} + \gamma_g \cdot 1\{\text{fatherEduc} = g\} + \delta_g \cdot 1\{\text{fatherEduc} = g\} \cdot \text{educ} \tag{18}$$

for individuals in 10-year birth cohort groups separately on OCG73 and OCG62 data. Here β captures a shared return to schooling, γ_g captures effect of parental background on earnings through non-education channels such as family connections, and δ_g captures differential returns to schooling by parental status due to factors such as educational quality.

Table A.7 Columns (1)-(5) present the results. I do not reject the hypothesis that returns to schooling (β) are the only determinant of children's earnings for any cohort in either OCG data set. I am unable to reject the hypothesis that other factors (γ_g and δ_g) changed in ways that could have offset the educational gains of children from low-SES parents. However, the point estimates decline across cohorts, which would amplify effects of increasing educational mobility on income mobility.

Therefore changes in children's education likely imply changes in children's earnings across all parental status groups. I now ask if higher income mobility can be observed directly in the OCG data. For this exercise, I estimate children's education and income slopes separately with respect to father's education, allowing the intercept and slope of this relationship to change across cohorts. Specifically, I estimate equations of the form

$$\text{childOutcome}_{g,\text{c}} = \pi + \phi \cdot \text{fatherEduc} + \eta_c \cdot 1\{\text{cohort} = c\} + \lambda_c \cdot 1\{\text{cohort} = c\} \cdot \text{fatherEduc} \tag{19}$$

for individuals in the same 10-year cohorts as before, where $\text{childOutcome}_{g,c}$ is either log earnings or education, father's education varies from 7 to 17 years of completed schooling, π and ϕ capture the intercept and slope of the outcome CEF, respectively, and the η_c and λ_c terms capture changes in the intercept and slope, respectively, across cohorts. I select cohorts that correspond roughly to cohorts of 26-29 year-olds in the 1940, 1950 and 1960 censueses. These cohorts have earnings that can be observed after age 27 in OCG data sets (except for the 1940 birth cohort in OCG62) and span the key educational mobility gains documented above.

Table A.7 Columns (6)-(9) present the results. The education IM statistics are similar to those estimated above on census data, and display similar increases in intercepts and decreases in slopes as in census data, although much less precisely. First note that a return to one year of schooling around 10% per year suggests that education by itself can explain about 75% ($= 0.1 \times 0.429/0.055$) of the gains from having higher-education parents. I have also replicated this pattern in the 1930-40 matched census panel for parental groups defined by home value and rent; there education by itself can explain about 50% of the gains from having higher-status parents.

Second, note that the gains in educational mobility with respect to parental education that I document above suggest the CEF rotated up by about one year for children of the lowest-education parents. Returns to schooling of 10% per year therefore imply that the income CEF in parental education should increase by 0.1 log points in the intercept and, given the domain of father's education from 7-17 years, should decrease the slope by about 0.01 log points. This is close to the results in Column (7) for the OCG1973, though again results are imprecise. In OCG1962, I cannot observe income for the cohort corresponding to the 1960 census with precision, and results are too imprecise to be useful for the cohort corresponding to 1950. Overall, these results do suggest that gains in educational mobility imply gains in income mobility, but are too noisy to demonstrate this conclusively. This is not surprising given that the motivation for this paper stems from a lack of any precise, long-term historical time series data on intergenerational income mobility.

D. Data Appendix

Census

I use publicly-available U.S. Decennial Census data from IPUMS (Ruggles et al., 2015). I also use 100% samples of the 1930 and 1940 censuses from the IPUMS collaboration with Ancestry.com.

I set age, relation to head, sex, race, earnings, and education to missing if allocated using quality flags.

In some cases I work with random samples of households that have no minorities or foreign-born members to facilitate analysis, scaling person and household weights accordingly.

I recode "relation to head" as "head" for all individuals who are not reporting relation to head as one of head, spouse, or child. I then define "dependent" as children living with parent(s), defining all remaining children as independent. As discussed in the text, this method treats all individuals who cannot be directly linked to their parents as "independent."

I set earnings to missing for individuals outside the age range of 25-65. All earnings are deflated with CPI-Urban to 2000 dollars.

Panel Study of Income Dynamics

Include all observations in the PSID in order to maximize power (including the Survey Research Center, the Survey of Economic Opportunity), which is especially important for results on blacks.

Use individual weights for current years to adjust for differential selection probabilities and attrition.

Impute schooling in future years for individuals who previously reported nonmissing schooling values.

Construct household earnings as in the census: set labor earnings to missing for individuals outside ages 25-65, use labor earnings only, sum earnings of both parents when present otherwise use earnings of single parent. Exclude households with zero total household income from percentile calculations. Use family income when children are 17, and either 16 or 17 in years after 1997 when PSID becomes biannual. Calculate deciles in this parental income variable separately in each calendar year, using current-year individual sample weights of children.

Define "dependent" as any individual living with parent(s), otherwise independent.

For education of parents, I check for any nonmissing mother and father education responses during years when children are ages 12-17 and dependent, and take median of reported education values if they differ across these years within individual parents, and then average these values for head and spouse to obtain parental education. I construct education ranks for parents in each calendar year separately using current-year individual sample weights of children.

Define child race based on parent race. I first identify the mode of races reported by individual parents when children are ages 12-17. If both parents are white, I define the child as white. If either parent is black, I define the child as black. I drop children for whom neither parent is either white or black.

Define decade labelled "1980" as outcomes observed in calendar years 1980-89, decade 1990 as outcomes observed in calendar years 1990-99, etc.

National Longitudinal Survey of Youth 1979

Restrict to nationally representative sample using "sample type" between values 1 and 8.

Recode all education variables to conform to census: years over 20 coded as 17.

Use mother and father's education as reported by adult respondents from memory, and average them as with census.

Define "dependent" as any individual for whom head reports relation as 4-5 (mother or father), 24-25 (mother or father-in-law), 47 ("parent"), 50-51 (foster parents before 1993) or 51-52 (foster parents after 1993).

Use sampling weights provided by NLSY in all calculations.

National Longitudinal Survey of Youth 1997

All choices similar to NLSY79.

Define "dependent" as any individual for whom head reports relations 3-12 (various types of parents), and as "independent" otherwise.

General Social Survey

Recode all education variables to span 1-17 grades as in census.

Take average of adult respondents reported mother's and father's education.

Use number of adults in household as weights, as recommended in GSS documentation.

Define "dependent" as any individual for reporting relation to head as "child," and as "independent" otherwise.

Occupational Change in a Generation 1973

Similar to GSS. Use weights provided in survey.

Occupational Change in a Generation 1962

Similar to GSS. Use weights provided in survey. Note only father's education available, and only in two-year intervals.

E. The Black High School Movement

One potential explanation for the sharp increase in black mobility after 1940 is the extreme scarcity of black high schools in Southern states with segregated school systems (Anderson, 1988).[39] In 1940, many southern blacks would have had to pay out of pocket for private and often faraway high schools. Much prior research has documented the importance of supply-side educational quality improvements in accounting for black-white education and earnings gaps (e.g., Smith and Welch, 1989; Margo, 1990; Card and Krueger, 1992b; Donohue III et al., 2002; Aaronson and Mazumder, 2011), but I am not aware of prior research quantifying impacts of the black high school movement in southern states. Once again, black high schools are unlikely to provide a full explanation given the similar IM gains of whites in the South.

To explore the role of black high schools in more detail I have compiled archival data on the evolution of black high schools by state from the series "Accredited Secondary Schools in the United States" from various years 1928-1944 (Phillips and United States Office of Education, 1929; Carr and United States Office of Education, 1930; United States Office of Education, 1933; Carr and United States Office of Education, 1934; United States Office of Education, 1937; Carr and of Education, 1939; United States Office of Education, 1943; Carr and of Education, 1944) and the series "Directory of Secondary Day Schools" from years 1949 and 1952 (Rice and United States Office of Education, 1949, 1952). The volumes from 1928-44 contain data on the number of *accredited* white and black high schools in every state, while the last two volumes from 1949-52 contain data on all white and black high schools, allowing me to assess whether accredited public high schools proxy well for all public high schools. The share of all black high schools that are accredited in 1949 and 1952 varies around 30-50% across states with segregated schools, though it is lower in a few states. To obtain a measure of high school density, I divide the number of white and black high schools by the number of age 14-17 white and black children, respectively. I then regress total public black high schools per-capita on accredited black high schools per

[39] I thank Robert Margo for suggesting this explanation.

capita. I obtain coefficients of 1.03 (SE=.18) in 1949 and .82 (SE=.15) in 1952. These findings suggest that accredited black high schools per capita are plausibly a good proxy for total black high schools per capita in earlier years. I match high school density in 1928 and 1952 to 26-29 year-olds by state of birth in the 1940 and 1970 censuses, respectively.[40]

Figure A.2 plots black and white public high schools per capita, by state, over the years 1928-1952. The figure illustrates the extreme relative scarcity of black high schools in the U.S. South in 1928, along with striking heterogeneity in convergence over the next 24 years. I exploit this variation using a difference-in-differences approach by plotting changes in black mobility across southern states with above- and below-median changes in black high school density. The "treatment" implied by this comparison is large: an additional increase of .002 high schools per black high-school age child, which is about two-thirds of average white high school density in 1952 in the South. Figure A.3 plots non-parametric education CEFs with respect to parental education in these two groups of states in 1940 and 1970. The figure shows roughly equal gains in mobility for blacks in states with small vs. large gains in black high school density. This result casts further doubt on the idea that the black high school movement accounts for black educational mobility gains 1940-70.[41]

[40] I use the 1970 census rather than the 1960 census for reasons of statistical power; results are similar in either case.

[41] I find similar results using a variety of additional difference-in-difference estimators that compare states with high vs. low growth in black high school density, high vs. low initial black high school density in 1940, high vs. low growth in relative density of black vs white high schools.

References

Aaronson, D. and B. Mazumder (2008): "Intergenerational Economic Mobility in the United States, 1940 to 2000," *Journal of Human Resources*, 43, 139–172.

——— (2011): "The Impact of Rosenwald Schools on Black Achievement," *Journal of Political Economy*, 119, 821–888.

Abramitzky, R., L. P. Boustan, and K. Eriksson (2014): "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration," *Journal of Political Economy*, 122, 467–506.

Acemoglu, D. and J. Angrist (2001): "How large are the social returns to education? Evidence from compulsory schooling laws," in *NBER Macroeconomics Annual*, ed. by B. S. Bernanke and K. S. Rogoff, National bureau of economic research, vol. 15.

Acemoglu, D. and J.-S. Pischke (2001): "Changes in the wage structure, family income, and children's education," *European Economic Review*, 45, 890–904.

Anderson, J. D. (1988): *The Education of Blacks in the South, 1860-1935*, Chapel Hill: The University of North Carolina Press, 1 edition ed.

Angrist, J. D. and S. Pischke (2009): *Mostly Harmless Econometrics : An Empiricist's Companion*, Princeton University Press.

Autor, D., L. Katz, and M. Kearney (2008): "Trends in U.S. Wage Inequality: Revising the Revisionists," *Review of Economics and Statistics*, 90, 300–323.

Autor, D. H., L. F. Katz, and A. B. Krueger (1998): "Computing Inequality: Have Computers Changed the Labor Market?" *Quarterly Journal of Economics*, 113, 1169–1214.

Ayres, L. P. (1920): *An Index Number for State School Systems*, Department of education, Russell Sage foundation.

Bacolod, M. P. (2007): "Do Alternative Opportunities Matter? The Role of Female Labor Markets in the Decline of Teacher Quality," *The Review of Economics and Statistics*, 89, 737–751.

Bailey, M. J. and S. M. Dynarski (2011): "Inequality in Postsecondary Education," in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, ed. by G. J. Duncan and R. J. Murnane, Russell Sage Foundation, 117–132.

Becker, G. S. and N. Tomes (1979): "An equilibrium theory of the distribution of income and intergenerational mobility," *Journal of Political Economy.*

——— (1986): "Human capital and the rise and fall of families." *Journal of labor economics*, 4, 1–47.

Belley, P. and L. Lochner (2007): "The Changing Role of Family Income and Ability in Determining Educational Achievement," *Journal of Human Capital*, 1, 37–89.

Benabou, R. and E. A. Ok (2001): "Social Mobility and the Demand for Redistribution: the POUM Hypothesis," *The Quarterly Journal of Economics*.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012): "The role of application assistance and information in college decisions: results from the H&R Block FAFSA Experiment," *The Quarterly Journal of Economics*, 1205–1242.

Bhattacharya, D. and B. Mazumder (2011): "A nonparametric analysis of black-white differences in intergenerational income mobility in the United States," *Quantitative Economics*, 2, 335–379.

Black, S. E. and P. J. Devereux (2011): "Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics, Volume 4B*, Elsevier B.V., vol. 4, 1487–1541.

Bratberg, E., J. Davis, B. Mazumder, K. Vaage, M. Nybom, and D. Schnitzlein (2015): "A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden and the U. S," *URL: http://folk. uib. no/secaa/Public/Trygd2014/InternationalMobilityCurves20Feb2015. pdf*.

Bratsberg, B., K. RÃžed, O. Raaum, R. Naylor, M. JaÂšntti, T. Eriksson, and E. OÂšsterbacka (2007): "Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons*," *The Economic Journal*, 117, C72–C92.

Braun, S. and J. Stuhler (2015): "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany," Tech. rep., Mimeo, November 2014, Universidad Carlos III de Madrid.

Bureau of the Census (1988): *1980 Census of Population and Housing: The coverage of population in the 1980 census*, U.S. Department of Commerce, Bureau of the Census.

Cameron, S. V. and J. J. Heckman (1993): "Comment on Trends in College Entry among Whites, Blacks, and Hispanics by Robert M. Hauser," in *Studies of Supply and Demand in Higher Education*, ed. by C. T. Clotfelter and M. Rothschild, University of Chicago Press, January, 61–120.

Card, D. and A. B. Krueger (1992a): "Does School Quality Matter ? Returns to Education and the Characteristics of Public Schools in the United States," *The Journal of Political Economy*, 100, 1–40.

——— (1992b): "School Quality and Black-White Relative Earnings: A Direct Assessment," *The Quarterly Journal of Economics*, 107, 151–200.

Card, D. and T. Lemieux (2000): "Adapting to Circumstances: The Evolution of Work, School,and Living Arrangements among North American Youth," in *Youth Employment and Joblessness in Advanced Countries*, ed. by D. Blanchflower and R. B. Freeman, 171–214.

Carr, M. J. S. and U. S. O. of Education (1939): *Accredited Secondary Schools in the United States. Bulletin, 1939, No. 2*, Office of Education, United States Department of the Interior.

——— (1944): *Accredited Secondary Schools in the United States, 1944. Bulletin, 1944, No. 4*, US Office of Education, Federal Security Agency.

Carr, M. J. S. and United States Office of Education (1930): *Accredited Secondary Schools in the United States. Bulletin, 1930, No. 24*, Office of Education, United States Department of the Interior.

——— (1934): *Accredited Secondary Schools in the United States. Bulletin, 1934, No. 17*, Office of Education, United States Department of the Interior.

Carrell, S. E. and B. Sacerdote (2013): "Why Do College Going Interventions Work?" Working Paper 19031, National Bureau of Economic Research.

Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick (2009): "Optimal Defaults and Active Decisions." *The Quarterly Journal of Economics*, 124, 1639–1674.

Checchi, D., A. Ichino, and A. Rustichini (1999): "More equal but less mobile?: Education financing and intergenerational mobility in Italy and in the US," *Journal of Public Economics*, 74, 351–393.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011): "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star," *The Quarterly Journal of Economics*, 126, 1593–1660.

Chetty, R., J. N. Friedman, S. Leth-Peterson, T. H. Nielsen, and T. Olsen (2013): "Active vs. Passive Decisions and Crowd-Out in Retirement Savings Accounts: Evidence from Denmark," *The Quarterly Journal of Economics*, 0–83.

Chetty, R. and N. Hendren (2015): "The Impacts of Neighborhoods on Intergenerational Mobility: Childhood Exposure Effects and County-Level Estimates," *Unpublished Manuscript*.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014a): "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States," *The Quarterly Journal of Economics*, 129, 1553–1623.

Chetty, R., N. Hendren, P. Kline, E. Saez, and N. Turner (2014b): "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," *American Economic Review*, 104, 141–147.

Clark, G. (2014): *The Son Also Rises: Surnames and the History of Social Mobility*, Princeton University Press.

Cogan, J. (1982): "The Decline in Black Teenage Employment: 1950-70," *The American Economic Review*, 72, 621–638.

Corak, M. (2006): "Do Poor Children Become Poor Adults ? Lessons from a Cross Country Comparison of Generational Earnings Mobility," *IZA Discussion Paper 1993*.

——— (2013): "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *Journal of Economic Perspectives*, 27, 79–102.

Costa, D. L. (2015): "Health and the Economy in the United States from 1750 to the Present," *Journal of Economic Literature*, 53, 503–570.

Denison, E. F. (1985): *Trends in American Economic Growth, 1929-1982*, Brookings Institution Press.

Donohue III, J. J., J. J. Heckman, and P. E. Todd (2002): "The Schooling of Southern Blacks: The Roles of Legal Activism and Private Philanthropy, 1910-1960," *The Quarterly Journal of Economics*.

Dynarski, S. and J. Scott-clayton (2013): "Financial Aid Policy: Lessons from Research," *The Future of Children*, 23.

Dynarski, S. M. and J. E. Scott-Clayton (2008): "Complexity and Targeting in Federal Student Aid: A Quantitative Analysis," *Tax Policy and the Economy*, 22, 109–150.

Edin, K. and M. J. Kefalas (2011): *Promises I Can Keep: Why Poor Women Put Motherhood before Marriage*, Berkeley: University of California Press, revised ed.

Erikson, R. and J. H. Goldthorpe (1992): *The Constant Flux: A Study of Class Mobility in Industrial Societies*, Oxford England : New York: Oxford University Press.

Fack, G. and J. Grenet (2015): "Improving College Access and Success for Low-Income Students: Evidence from a Large Need-Based Grant Program," *American Economic Journal: Applied Economics*, 7, 1–34.

Feigenbaum, J. J. (2015): "Automated Census Record Linking," *Unpublished Manuscript.*

Folger, J. K. and C. B. Nam (1967): *Education of the American Population*, U.S. Department of Commerce, Bureau of the Census.

Galor, O. and D. Tsiddon (1997): "Technological Progress, Mobility, and Economic Growth," *The American Economic Review*, 87, 363–382.

Goldin, C. (1998): "America's Graduation from High School: The Evolution and Spread of Secondary Schooling in the Twentieth Century," *The Journal of Economic History*, 58, 345–374.

Goldin, C. and L. F. Katz (2010): *The Race between Education and Technology*, Cambridge, Mass.: Belknap Press.

Goldin, C. and R. A. Margo (1992): "The Great Compression: The Wage Structure in the United States at Mid-Century," *The Quarterly Journal of Economics*, 107, 1–34.

Greene, W. H. (2002): *Econometric Analysis*, Prentice Hall.

Grusky, D. B., T. M. Smeeding, and C. M. Snipp (2015): "A New Infrastructure for Monitoring Social Mobility in the United States," *The ANNALS of the American Academy of Political and Social Science*, 657, 63–82.

Harding, D. J., C. Jencks, L. M. Lopoo, and S. E. Mayer (2009): "The changing effect of family background on the incomes of American adults," in *Unequal Chances: Family Background and Economic Success*, ed. by S. Bowles, H. Gintis, and M. O. Groves, Princeton University Press.

Heckman, J. J. and P. a. Lafontaine (2010): "the American High School Graduation Rate: Trends and Levels." *The review of economics and statistics*, 92, 244–262.

Heckman, J. J., L. J. Lochner, and P. E. Todd (2006): "Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond," *Handbook of the Economics of Education*, 1.

Hertz, T. (2007): "Trends in the Intergenerational Elasticity of Family Income in the United States," *Industrial Relations*, 46, 22–50.

Hertz, T., T. Jayasundera, P. Piraino, S. Selcuk, N. Smith, and A. Verashchagina (2008): "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends," *B.E. Journal of Economic Analysis & Policy*, 7.

Hilger, N. G. (2016): "Upward Mobility and Discrimination: the Case of Asian-Americans," *Unpublished Manuscript.*

Hoxby, C. M. and A. Leigh (2004): "Pulled Away or Pushed Out? Explaining the Decline of Teacher Aptitude in the United States," *The American Economic Review*, 94, 236–240.

Husted, T. A. and L. W. Kenny (1997): "The Effect of the Expansion of the Voting Franchise on the Size of Government," *Journal of Political Economy*, 105, 54–82.

Jackson, C. K., R. C. Johnson, and C. Persico (2015): "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms*," *The Quarterly Journal of Economics*, qjv036.

Jantti, M., B. Bratsberg, K. Roed, O. Raaum, R. Naylor, E. Osterbacka, A. Bjorklund, and T. Eriksson (2006): "American Exceptionalism in a New Light: A Comparison of Intergenerational Earnings Mobility in the Nordic Countries, the United Kingdom and the United States," *IZA DP 1938*.

Johnson, D. S., C. Massey, and A. O'Hara (2015): "The Opportunities and Challenges of Using Administrative Data Linkages to Evaluate Mobility," *The ANNALS of the American Academy of Political and Social Science*, 657, 247–264.

Kaplan, G. (2012): "Moving Back Home: Insurance against Labor Market Risk," *Journal of Political Economy*, 120, 446–512.

Kearney, M. S. and P. B. Levine (2012): "Why is the Teen Birth Rate in the United States So High and Why Does It Matter?" *The Journal of Economic Perspectives*, 26, 141–66.

Krueger, A. B. (2012): "The Rise and Consequences of Inequality in the United States," Tech. rep., Washington D.C.

Lee, C.-I. and G. Solon (2009): "Trends in intergenerational income mobility," *The Review of Economics and Statistics*, 91, 766–772.

Lemieux, T. (2006): "The "Mincer Equation" Thirty Years after Schooling, Experience and Earnings"," in *Jacob Mincer: A Pioneer of Modern Labor Economics*.

Lleras-Muney, A. (2002): "Were Compulsory Attendance and Child Labor Laws Effective? An Analysis from 1915 to 1939," *Journal of Law and Economics*, 45, 401–435.

Lochner, L. J. and A. Monge-Naranjo (2011): "The Nature of Credit Constraints and Human Capital," *American Economic Review*, 101, 2487–2529.

Long, J. and J. Ferrie (2013): "Intergenerational Occupational Mobility in Great Britain and the United States Since 1850," *American Economic Review*, 103, 2041–2049.

LOURY, G. C. (1981): "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49, 843–867.

MARGO, R. A. (1986): "Race and Human Capital : Comment," *The American Economic Review*, 76, 1221–1224.

——— (1990): *Race and Schooling in the South, 1880-1950: An Economic History*, Chicago: University Of Chicago Press, 1 edition ed.

MARGO, R. A. AND T. A. FINEGAN (1993): "The Decline in Black Teenage Labor-Force Participation in the South, 1900- 1970: The Role of Schooling," *The American Economic Review*, 83, 234–247.

MAYER, S. E. AND L. M. LOPOO (2008): "Government spending and intergenerational mobility," *Journal of Public Economics*, 92, 139–158.

MAZUMDER, B. (2005): "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data," *Review of Economics and Statistics*, 87, 235–255.

——— (2015): "Estimating the Intergenerational Elasticity and Rank association in the US: Overcoming the current limitations of tax data," *Federal Reserve Bank of Chicago Working Paper 2015-04*.

MAZUMDER, B. AND M. ACOSTA (2015): "Using Occupation to Measure Intergenerational Mobility," *The ANNALS of the American Academy of Political and Social Science*, 657, 174–193.

MINNESOTA POPULATION CENTER AND ANCESTRY.COM (2013): "IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]," Tech. rep., University of Minnesota, Minneapolis.

MITNIK, P. A., V. BRYANT, M. WEBER, AND D. B. GRUSKY (2015): "New Estimates of Intergenerational Mobility Using Administrative Data," *Working Paper*.

MURPHY, K. M., A. SHLEIFER, AND R. W. VISHNY (1991): "The allocation of talent: implications for growth," *The Quarterly Journal of Economics*, 106, 503–530.

NATIONAL CENTER FOR EDUCATION STATISTICS (1984): "High School and Beyond, a National Longitudinal Study for the 1980s: Quality of Responses of High School Students to Questionnaire Items," Tech. rep.

NATIONAL RESEARCH COUNCIL (2006): *Once, Only Once, and in the Right Place: Residence Rules in the Decennial Census*, Washington, D.C: National Academies Press.

Neal, D. (2006): "Chapter 9: Why Has Black-White Skill Convergence Stopped?" in *Handbook of the Economics of Education*, Elsevier, vol. 1, 511–576.

Nybom, M. and J. Stuhler (2014): "Interpreting Trends in Intergenerational Mobility," Tech. rep.

Olivetti, C. and M. D. Paserman (2015): "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850-1940," *American Economic Review*, 105, 2695–2724.

Olivetti, C., M. D. Paserman, and L. Salisbury (2014): "Intergenerational Mobility Across Three Generations in the 19th Century : Evidence from the US Census," *Unpublished Manuscript*.

Owen, A. L. and D. N. Weil (1998): "Intergenerational earnings mobility, inequality and growth," *Journal of Monetary Economics*, 41, 71–104.

Phillips, F. M. and United States Office of Education (1929): *Accredited Secondary Schools in the United States. Bulletin, 1928, No. 26*, Bureau of Education, Department of the Interior.

Piketty, T. (1995): "Social Mobility and Redistributive Politics," *The Quarterly Journal of Economics*, 110.

Reardon, S. F. (2011): "The Widening Academic Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations Sean F. Reardon," in *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, ed. by G. J. Duncan and R. J. Murnane, Russell Sage Foundation.

Rice, M. C. and United States Office of Education (1952): *Directory of secondary day schools, 1951-52: showing accredited status, enrollment, staff, and other data*, Washington, D.C.: Federal Security Agency, Office of Education.

Rice, M. F. C. and United States Office of Education (1949): *Directory of secondary schools in the United States*, Circular (United States. Office of Education)250, Washington: U. S. Govt. Print. Off.

Rosenbloom, J. L. and W. A. Sundstrom (1999): "The Sources of Regional Variation in the Severity of the Great Depression: Evidence from U.S. Manufacturing, 1919-1937," *The Journal of Economic History*, 59, 714–747.

Ruggles, S. J., K. Genadek, R. Goeken, J. Grover, and M. Sobek (2015): *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database].*, Minneapolis: University of Minnesota.

Smith, J. P. and F. R. Welch (1989): "Black Economic Progress After Myrdal," *Journal of Economic Literature*, 27, 519–564.

Smith, P. J. (1984): "Race and Human Capital," *The American Economic Review*, 74, 685–698.

Solon, G. (1992): "Intergenerational Income Mobility in the United States," *The American Economic Review*, 82, 393–408.

——— (1999): "Intergenerational Mobility in the Labor Market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, Elsevier Science B.V., vol. 3.

——— (2004): "A Model of Intergenerational Mobility Variation over Time and Place," in *Generational Income Mobility in North America and Europe*, ed. by M. Corak, Cambridge, UK: Cambridge University Press, 1979, 38–47.

——— (2015): "What Do We Know So Far about Multigenerational Mobility?" Working Paper 21053, National Bureau of Economic Research.

Stephens, Jr., M. and D.-y. Yang (2014): "Compulsory Education and the Benefits of Schooling," *The American Economic Review*, 104, 1777–1792.

Stuhler, J. (2014): "Mobility Across Multiple Generations," *Unpublished Manuscript*.

Torche, F. (2015): "Analyses of Intergenerational Mobility An Interdisciplinary Review," *The ANNALS of the American Academy of Political and Social Science*, 657, 37–62.

Turner, S. and J. Bound (2003): "closing the gap or widening the divide: the effects of the G.I. bill and world war II on the educational outcomes of black americans," *The Journal of Economic History*, 63.

United States Office of Education (1933): *Accredited Secondary Schools in the United States. Supplement to Bulletin, 1930, No. 24*, Office of Education, United States Department of the Interior.

——— (1937): *Accredited Secondary Schools in the United States. Supplement to Bulletin, 1934, No. 17*, Office of Education, United States Department of the Interior.

——— (1943): *Accredited Secondary Schools in the United States. Bulletin, 1942, No. 3. Supplement to Bulletin, 1939, No. 2*, US Office of Education, Federal Security Agency.

Wright, G. (1986): *Old South, New South: Revolutions in the Southern Economy since the Civil War*, New York: Basic Books.

——— (1987): "The economic revolution in the American South," *The Journal of Economic Perspectives*, 161–178.

	White						Black					
Age	1940	1960	1970	1980	1990	2000	1940	1960	1970	1980	1990	2000
10	3.73	3.56	3.49	3.47	3.19	2.91	2.85	3.34	3.43	3.47	3.57	3.18
11	4.61	4.54	4.48	4.47	4.96	4.29	3.52	4.25	4.37	4.45	5.24	4.59
12	5.47	5.53	5.47	5.47	6.23	5.69	4.12	5.16	5.32	5.41	6.18	5.84
13	6.39	6.47	6.46	6.46	6.43	6.66	4.81	6.03	6.23	6.39	6.45	6.85
14	7.32	7.52	7.42	7.45	6.8	7.55	5.54	7.07	7.16	7.36	7	7.64
15	8.16	8.41	8.43	8.43	8.11	8.35	6.14	7.81	8.12	8.27	8.27	8.56
16	8.9	9.33	9.4	9.38	9.55	9.49	6.62	8.64	9	9.2	9.51	9.58
17	9.55	10.2	10.32	10.33	10.54	10.48	7	9.29	9.82	10.07	10.51	10.51
18	9.97	10.93	11.16	11.22	11.58	11.39	7.09	9.74	10.53	10.92	11.38	11.26
19	10.26	11.31	11.75	11.83	12.17	11.96	7.25	9.92	10.93	11.44	11.82	11.69
20	10.31	11.52	12.14	12.2	12.37	12.27	7.06	10.14	11.14	11.74	12.02	11.87
21	10.39	11.65	12.45	12.51	12.51	12.47	7.06	10.08	11.33	11.94	12.11	12.03
22	10.42	11.75	12.58	12.68	12.81	12.82	6.91	10.16	11.37	12.07	12.19	12.23
23	10.36	11.79	12.63	12.85	13.04	13.11	6.92	10.1	11.41	12.13	12.28	12.39
24	10.32	11.74	12.54	12.91	13.12	13.3	6.86	9.93	11.18	12.11	12.36	12.54
25	10.25	11.69	12.51	12.99	13.18	13.4	6.72	9.79	11.19	12.15	12.37	12.64
26	10.19	11.68	12.5	13.11	13.18	13.47	6.81	9.8	11.23	12.17	12.43	12.66
27	10.14	11.66	12.48	13.15	13.18	13.53	6.78	9.77	11.19	12.21	12.46	12.72
28	10.03	11.65	12.42	13.24	13.17	13.57	6.56	9.65	11.17	12.23	12.49	12.77
29	10.01	11.64	12.34	13.29	13.21	13.61	6.6	9.49	10.96	12.24	12.52	12.83

Table I: Educational Attainment by Age, Year, and Race

Notes: Table presents highest grade attained by age and year for whites and blacks separately in census data.

	White						Black					
Age	1940	1960	1970	1980	1990	2000	1940	1960	1970	1980	1990	2000
10	0.93	0.97	0.97	0.96	0.95	0.95	0.81	0.84	0.9	0.88	0.83	0.83
11	0.93	0.96	0.98	0.97	0.95	0.95	0.81	0.83	0.89	0.88	0.83	0.84
12	0.93	0.96	0.97	0.97	0.95	0.95	0.81	0.83	0.9	0.89	0.84	0.84
13	0.93	0.96	0.97	0.97	0.95	0.95	0.81	0.83	0.89	0.89	0.84	0.83
14	0.92	0.95	0.96	0.96	0.95	0.95	0.81	0.82	0.88	0.89	0.84	0.83
15	0.92	0.94	0.96	0.96	0.94	0.94	0.79	0.8	0.87	0.88	0.83	0.82
16	0.9	0.92	0.95	0.95	0.93	0.93	0.77	0.8	0.86	0.87	0.81	0.8
17	0.87	0.86	0.92	0.91	0.91	0.91	0.73	0.74	0.83	0.85	0.8	0.78
18	0.8	0.62	0.72	0.71	0.73	0.74	0.66	0.64	0.72	0.73	0.69	0.66
19	0.73	0.47	0.54	0.54	0.56	0.51	0.58	0.53	0.59	0.61	0.59	0.53
20	0.65	0.37	0.42	0.44	0.49	0.43	0.48	0.44	0.47	0.54	0.52	0.47
21	0.57	0.32	0.32	0.36	0.42	0.38	0.41	0.36	0.38	0.46	0.48	0.4
22	0.49	0.26	0.26	0.31	0.37	0.32	0.34	0.3	0.31	0.4	0.42	0.36
23	0.41	0.2	0.2	0.24	0.31	0.29	0.28	0.24	0.26	0.33	0.38	0.32
24	0.35	0.16	0.15	0.18	0.26	0.23	0.23	0.21	0.2	0.28	0.34	0.28
25	0.3	0.14	0.12	0.14	0.21	0.19	0.2	0.19	0.16	0.23	0.3	0.24
26	0.26	0.12	0.1	0.12	0.17	0.15	0.18	0.16	0.13	0.2	0.26	0.2
27	0.23	0.1	0.08	0.09	0.13	0.13	0.15	0.13	0.12	0.17	0.24	0.19
28	0.2	0.09	0.07	0.08	0.11	0.11	0.13	0.12	0.1	0.14	0.22	0.17
29	0.17	0.08	0.06	0.07	0.1	0.09	0.11	0.11	0.1	0.12	0.19	0.15

Table II: Percent of Children Living at Home with Parents by Age, Year, and Race

Notes: Table presents share of children living at home with identifiable parents by age and year for whites and blacks separately in census data.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
				PSID							NLSY79					NLSY97			GSS	OCG73
Child Educ	All	1980-89	1990-99	2000-09	Male	Female	White	Black	All	Male	Female	White	Black	All	Male	Female	White	Black	All	All
Par Educ	0.436**	0.393**	0.490**	0.521**	0.431**	0.442**	0.441**	0.256**	0.506**	0.507**	0.506**	0.513**	0.323**	0.497**	0.496**	0.501**	0.510**	0.398**	0.336**	0.441**
	(0.00599)	(0.00759)	(0.0110)	(0.0146)	(0.00784)	(0.00702)	(0.00736)	(0.00888)	(0.00618)	(0.00920)	(0.00832)	(0.00683)	(0.0202)	(0.00919)	(0.0129)	(0.0130)	(0.0109)	(0.0268)	(0.0169)	(0.0172)
Dep	0.551*	0.287	1.273**	-0.509	0.224	1.137**	-0.249	0.298	0.118	-0.0471	0.555	0.00185	-0.838	0.0192	-0.201	0.444	0.298	-3.077**	1.294*	-0.0113
	(0.269)	(0.300)	(0.412)	(0.593)	(0.283)	(0.336)	(0.340)	(0.241)	(0.206)	(0.277)	(0.317)	(0.238)	(0.455)	(0.274)	(0.354)	(0.430)	(0.339)	(0.662)	(0.599)	(0.525)
Par Educ * Dep	-0.0711**	-0.0399	-0.131**	0.0139	-0.0550*	-0.0946**	-0.0104	-0.0417	-0.0231	-0.0196	-0.0427	-0.0129	0.0634	-0.0288	-0.0111	-0.0525	-0.0523*	0.232**	-0.114*	-0.0568
	(0.0222)	(0.0259)	(0.0336)	(0.0449)	(0.0230)	(0.0272)	(0.0268)	(0.0218)	(0.0168)	(0.0227)	(0.0256)	(0.0192)	(0.0401)	(0.0206)	(0.0265)	(0.0325)	(0.0253)	(0.0518)	(0.0459)	(0.0496)
Constant	7.980**	8.507**	7.304**	6.702**	8.002**	7.952**	7.974**	9.841**	6.978**	6.935**	7.006**	6.902**	8.787**	7.107**	6.875**	7.281**	6.966**	8.041**	9.295**	8.208**
	(0.0734)	(0.0906)	(0.138)	(0.197)	(0.0981)	(0.0871)	(0.0936)	(0.0978)	(0.0772)	(0.116)	(0.103)	(0.0859)	(0.230)	(0.125)	(0.176)	(0.177)	(0.150)	(0.348)	(0.223)	(0.188)
Observations	21,147	9,695	6,140	3,296	9,998	11,149	11,203	7,747	20,986	10,198	10,788	17,388	2,459	17,116	8,679	8,437	12,245	2,747	3,401	2,985
R-squared	0.262	0.230	0.259	0.309	0.255	0.271	0.259	0.110	0.269	0.266	0.274	0.270	0.134	0.180	0.189	0.173	0.183	0.136	0.111	0.201

Robust standard errors in parentheses

** p<0.01, * p<0.05

Table III: Parallel Trends in Parental Education

Notes: Each column displays estimates of equation (6) for children's schooling at ages 26-29 by parental income deciles. Parental characteristics measured when children are age 17, or earlier in adolescence if not observed at age 17. Children's schooling excludes zeros. Parental education defined as average of mother's and father's educational attainment, or educational attainment of the single parent when second parent not observed. Children with parental education in the bottom 2% of the parental education distribution excluded. Sample weights used in all calculations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
					PSID						NLSY79					NLSY97		
Child Educ	All	1980-89	1990-99	2000-09	Male	Female	White	Black	All	Male	Female	White	Black	All	Male	Female	White	Black
Par Income	0.237**	0.196**	0.256**	0.272**	0.257**	0.218**	0.229**	0.206**	0.266**	0.266**	0.265**	0.264**	0.161**	0.319**	0.308**	0.334**	0.318**	0.251**
	(0.00622)	(0.00939)	(0.0114)	(0.0149)	(0.00899)	(0.00858)	(0.00709)	(0.0179)	(0.00854)	(0.0122)	(0.0120)	(0.00953)	(0.0292)	(0.00925)	(0.0138)	(0.0124)	(0.0110)	(0.0284)
Dep	-0.102	-0.224	0.185	-0.853**	0.0890	-0.203	-0.108	0.102	-0.403**	-0.387*	-0.413	-0.427**	-0.206	-0.365**	-0.168	-0.499**	-0.468**	-0.752**
	(0.118)	(0.219)	(0.168)	(0.237)	(0.150)	(0.194)	(0.154)	(0.192)	(0.135)	(0.178)	(0.211)	(0.164)	(0.203)	(0.128)	(0.173)	(0.191)	(0.165)	(0.258)
Par Inc * Dep	-0.0411	-0.0131	-0.102**	0.0683	-0.0996**	0.0390	-0.0341	-0.0883	0.0237	-0.00249	0.0685	0.0288	0.0421	-0.0378	-0.0392	-0.0365	-0.0297	0.0672
	(0.0220)	(0.0419)	(0.0335)	(0.0398)	(0.0289)	(0.0328)	(0.0261)	(0.0495)	(0.0226)	(0.0297)	(0.0355)	(0.0264)	(0.0594)	(0.0219)	(0.0295)	(0.0327)	(0.0273)	(0.0552)
Constant	12.08**	12.08**	11.97**	12.35**	11.89**	12.26**	12.16**	11.94**	11.77**	11.71**	11.82**	11.80**	11.77**	11.90**	11.68**	12.06**	11.93**	12.02**
	(0.0385)	(0.0579)	(0.0696)	(0.0949)	(0.0553)	(0.0533)	(0.0462)	(0.0649)	(0.0519)	(0.0740)	(0.0725)	(0.0588)	(0.108)	(0.0581)	(0.0874)	(0.0769)	(0.0711)	(0.134)
Observations	17,793	8,093	5,104	2,797	8,554	9,239	10,911	6,773	9,463	4,813	4,650	7,752	1,106	9,934	4,819	5,115	7,168	1,478
R-squared	0.113	0.083	0.131	0.165	0.127	0.104	0.100	0.076	0.111	0.111	0.113	0.106	0.041	0.133	0.120	0.150	0.132	0.084

Robust standard errors in parentheses

** p<0.01, * p<0.05

Table IV: Parallel Trends in Parental Income Deciles

Notes: Each column displays estimates of equation (6) for children's schooling at ages 26-29 by parental income deciles. Parental characteristics measured when children are age 17, or earlier in adolescence if not observed at age 17. Children's schooling excludes zeros. Parental income defined as sum of mother's and father's earnings when possible, otherwise as total family income. Children with zero parental income at age 17 excluded from the parental income figures. Parental income deciles calculated separately by year. Sample weights used in all calculations.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	PSID				NLSY79			
VARIABLES	Dep 22-25	Dep 26-29	Married 18-25	Married 26-29	Dep 22-25	Dep 26-29	Married 18-25	Married 26-29
Education	-0.0473**	-0.00769**	0.0251**	-0.00796*	-0.00517	-0.00628**	-0.0340**	0.00129
	(0.00370)	(0.00187)	(0.00446)	(0.00301)	(0.00317)	(0.00243)	(0.00352)	(0.00435)
Married	-0.609**	-0.209**			-0.462**	-0.231**		
	(0.0221)	(0.0107)			(0.0104)	(0.0122)		
Female	-0.0474**	-0.0524**	0.0890**	0.0532**	-0.0314**	-0.0332**	0.135**	0.0966**
	(0.00932)	(0.00827)	(0.0131)	(0.0131)	(0.0112)	(0.0112)	(0.0123)	(0.0193)
Par Income Decile	0.0116**	0.00002	-0.0122**	0.00631*	0.00598**	-0.000101	-0.00414	0.00150
	(0.00298)	(0.00109)	(0.00187)	(0.00291)	(0.00200)	(0.00211)	(0.00220)	(0.00351)
Constant	1.494**	0.314**	-0.265**	0.738**	0.448**	0.387**	0.866**	0.462**
	(0.0431)	(0.0440)	(0.0444)	(0.0983)	(0.0467)	(0.0364)	(0.0516)	(0.0586)
Observations	37,983	17,504	37,983	17,504	20,088	9,687	20,088	9,687
R-squared	0.406	0.152	0.081	0.016	0.323	0.120	0.126	0.010
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

** p<0.01, * p<0.05

Table V: Determinants of Dependent and Marital Status

Notes: Table reports regression estimates in PSID and NLSY79 samples described in text. Columns (1) and (5) regress dummies for dependent status on various regressors in sample restricting to ages 22-25. Columns (2) and (6) repeat these regressions for samples age 26-29, and indicate much smaller coefficients and R^2 at these older ages. Columns (3) and (7) regress dummies for married on various regressions in sample restricting to ages 18-25. Columns (4) and (8) repeat these regressions for samples age 26-29, and again indicate much smaller coefficients and R^2 at these older ages. All regressions use sample weights and restrictions described in text.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep Var: Educ Group Cohort Share Ages 16-17	All Years	All Years	All Years	1940	1950	1960	1970	1980	1990	2000
Age 7	0.950**			1.163**	0.800**	0.660**	0.780**	0.941**	1.046**	1.039**
	(0.0152)			(0.0441)	(0.109)	(0.0729)	(0.0188)	(0.0160)	(0.0220)	(0.0137)
Ages 4-7, linear		0.882**								
		(0.0104)								
Ages 1-7, linear			0.241**							
			(0.0275)							
Constant	0.00667*	0.0131**	0.115**	-0.0139*	0.0337	0.0386**	0.0227**	0.00682	-0.00716	-0.00566
	(0.00295)	(0.00211)	(0.00824)	(0.00585)	(0.0207)	(0.00957)	(0.00338)	(0.00345)	(0.00478)	(0.00323)
Observations	190	190	190	32	34	28	26	24	24	22
R-squared	0.954	0.975	0.291	0.959	0.629	0.759	0.986	0.994	0.990	0.997

Standard errors in parentheses
** p<0.01, * p<0.05

Table VI: Validation of Group Cohort Size Predictors in Parental Education, Whites

Notes: Documents ability to predict group cohort sizes at later ages with group cohort sizes at earlier ages. Columns (1)-(3) regress actual parental education group cohort shares at ages 16-17 on predicted parental education group cohort shares at ages 16-17 and a constant, where each column uses a different predictor. Columns (4)-(10) run the regression in Column (1) separately by census year 1940-2000. Regressions for white-only sample. All regressions weighted by the square root of the cell size.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep Var: Income Group										
Cohort Share Ages 16-17	All Years	All Years	All Years	1940	1950	1960	1970	1980	1990	2000
Age 7	0.929**			0.753**	0.569**	0.692**	0.768**	0.826*	1.398**	1.327**
	(0.0800)			(0.141)	(0.0556)	(0.161)	(0.219)	(0.318)	(0.216)	(0.0716)
Ages 4-7, linear		0.880**								
		(0.0444)								
Ages 1-7, linear			0.883**							
			(0.0568)							
Constant	0.0117	0.0145**	0.0152*	0.0252	0.0439**	0.0335	0.0291	0.0262	-0.0373	-0.0324**
	(0.00879)	(0.00509)	(0.00641)	(0.0144)	(0.00637)	(0.0176)	(0.0245)	(0.0347)	(0.0236)	(0.00809)
Observations	140	140	140	20	20	20	20	20	20	20
R-squared	0.494	0.740	0.636	0.613	0.853	0.508	0.406	0.273	0.700	0.950

Standard errors in parentheses

** p<0.01, * p<0.05

Table VII: Validation of Group Cohort Size Predictors in Parental Income, Whites

Notes: Documents ability to predict group cohort sizes at later ages with group cohort sizes at earlier ages. Columns (1)-(3) regress actual parental income group cohort shares at ages 16-17 on predicted parental income group cohort shares at ages 16-17 and a constant, where each column uses a different predictor. Columns (4)-(10) run the regression in Column (1) separately by census year 1940-2000. Regressions for white-only sample. All regressions weighted by the square root of the cell size.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Var	Highest Grade Attained, Age 26-29							Enrollment	
Sample	All	Boys	Girls	South	North	Urban	Rural	Age 16-18	Age 19-21
Intercept 1940	7.267**	7.065**	7.558**	6.043**	7.881**	8.128**	6.454**	0.373**	-0.0791*
	(0.141)	(0.170)	(0.161)	(0.208)	(0.186)	(0.140)	(0.152)	(0.0129)	(0.0377)
Intercept 1960	9.311**	8.886**	9.171**	7.825**	9.457**	9.430**	8.119**	0.553**	-0.0426
	(0.146)	(0.211)	(0.169)	(0.232)	(0.216)	(0.161)	(0.195)	(0.0143)	(0.0391)
Intercept 1970	10.12**	9.546**	10.14**	9.735**	9.764**			0.712**	-0.00600
	(0.133)	(0.184)	(0.168)	(0.282)	(0.179)			(0.0128)	(0.0357)
Intercept 1980	10.84**	10.57**	10.46**	10.11**	10.72**			0.642**	-0.0444
	(0.135)	(0.181)	(0.171)	(0.224)	(0.196)			(0.0129)	(0.0360)
Intercept 1990	10.72**	10.70**	10.80**	10.57**	10.86**	10.99**	10.05**	0.728**	0.0706
	(0.172)	(0.209)	(0.196)	(0.248)	(0.231)	(0.163)	(0.216)	(0.0165)	(0.0460)
Intercept 2000	11.45**	11.08**	11.14**	10.54**	10.89**			0.802**	0.168**
	(0.151)	(0.207)	(0.187)	(0.281)	(0.253)			(0.0146)	(0.0403)
Slope 1940	0.499**	0.516**	0.469**	0.647**	0.432**	0.428**	0.562**	0.0442**	0.0411**
	(0.0216)	(0.0263)	(0.0244)	(0.0340)	(0.0278)	(0.0203)	(0.0247)	(0.00198)	(0.00578)
Slope 1960	0.388**	0.414**	0.334**	0.533**	0.329**	0.340**	0.453**	0.0344**	0.0487**
	(0.0218)	(0.0278)	(0.0222)	(0.0326)	(0.0278)	(0.0202)	(0.0289)	(0.00217)	(0.00582)
Slope 1970	0.371**	0.426**	0.295**	0.361**	0.372**			0.0255**	0.0577**
	(0.0196)	(0.0238)	(0.0216)	(0.0367)	(0.0230)			(0.00190)	(0.00523)
Slope 1980	0.387**	0.391**	0.371**	0.422**	0.358**			0.0322**	0.0675**
	(0.0206)	(0.0243)	(0.0228)	(0.0308)	(0.0259)			(0.00199)	(0.00550)
Slope 1990	0.365**	0.363**	0.363**	0.373**	0.353**	0.343**	0.399**	0.0238**	0.0665**
	(0.0244)	(0.0298)	(0.0277)	(0.0362)	(0.0323)	(0.0224)	(0.0330)	(0.00236)	(0.00652)
Slope 2000	0.404**	0.380**	0.411**	0.406**	0.378**			0.0220**	0.0707**
	(0.0271)	(0.0320)	(0.0287)	(0.0387)	(0.0336)			(0.00263)	(0.00723)
Observations	77	80	81	82	80	43	43	77	77
R-squared	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	0.988

Standard errors in parentheses
** p<0.01, * p<0.05

Table VIII: Mobility Estimates in Parental Education, Whites

Notes: Displays estimated intercepts and slopes of children's schooling CEFs with respect to parental education for whites. Estimates correspond to α_t and β_t from Equation (9). Regressions weighted by square root of estimated cell sizes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Var	Highest Grade Attained, Age 26-29							Enrollment	
Sample	All	Boys	Girls	South	North	Urban	Rural	Age 16-18	Age 19-21
Intercept 1940	4.193**	3.697**	4.607**	3.901**	6.291**	5.117**	3.700**	0.336**	0.0115
	(0.141)	(0.144)	(0.219)	(0.151)	(0.352)	(0.203)	(0.149)	(0.0121)	(0.0324)
Intercept 1960	7.694**	6.103**	8.248**	6.684**	8.589**	8.183**	6.050**	0.574**	0.130**
	(0.144)	(0.176)	(0.236)	(0.179)	(0.321)	(0.203)	(0.204)	(0.0128)	(0.0326)
Intercept 1970	9.242**	9.217**	9.336**	9.105**	9.312**			0.727**	0.0391
	(0.135)	(0.148)	(0.206)	(0.195)	(0.149)			(0.0119)	(0.0309)
Intercept 1980	10.45**	9.822**	9.996**	10.11**	10.25**			0.711**	0.103**
	(0.124)	(0.159)	(0.227)	(0.162)	(0.194)			(0.0113)	(0.0285)
Intercept 1990	11.04**	10.37**	10.85**	11.01**	11.10**	10.55**	11.00**	0.748**	0.131**
	(0.154)	(0.209)	(0.296)	(0.190)	(0.207)	(0.191)	(0.338)	(0.0136)	(0.0352)
Intercept 2000	10.98**	10.72**	10.92**	10.64**	10.81**			0.794**	0.123**
	(0.145)	(0.167)	(0.248)	(0.203)	(0.219)			(0.0130)	(0.0334)
Slope 1940	0.635**	0.639**	0.627**	0.632**	0.419**	0.530**	0.585**	0.0371**	0.0228**
	(0.0282)	(0.0291)	(0.0431)	(0.0326)	(0.0533)	(0.0333)	(0.0357)	(0.00242)	(0.00648)
Slope 1960	0.409**	0.583**	0.285**	0.464**	0.250**	0.289**	0.519**	0.0174**	0.0119*
	(0.0244)	(0.0275)	(0.0341)	(0.0299)	(0.0396)	(0.0271)	(0.0423)	(0.00210)	(0.00540)
Slope 1970	0.285**	0.291**	0.275**	0.314**	0.273**			0.0115**	0.0298**
	(0.0181)	(0.0198)	(0.0278)	(0.0283)	(0.0192)			(0.00159)	(0.00415)
Slope 1980	0.275**	0.277**	0.271**	0.283**	0.269**			0.0162**	0.0340**
	(0.0177)	(0.0180)	(0.0255)	(0.0217)	(0.0230)			(0.00161)	(0.00406)
Slope 1990	0.239**	0.242**	0.227**	0.236**	0.240**	0.244**	0.139**	0.0186**	0.0482**
	(0.0243)	(0.0254)	(0.0360)	(0.0308)	(0.0316)	(0.0229)	(0.0451)	(0.00216)	(0.00557)
Slope 2000	0.374**	0.339**	0.362**	0.371**	0.346**			0.0208**	0.0662**
	(0.0299)	(0.0288)	(0.0430)	(0.0358)	(0.0369)			(0.00270)	(0.00687)
Observations	85	88	89	88	83	46	45	85	85
R-squared	1.000	1.000	0.999	0.999	1.000	1.000	0.998	1.000	0.984

Standard errors in parentheses
** p<0.01, * p<0.05

Table IX: Mobility Estimates in Parental Education, Blacks

Notes: Displays estimated intercepts and slopes of children's schooling CEFs with respect to parental education for blacks. Estimates correspond to α_t and β_t from Equation (9). Regressions weighted by square root of estimated cell sizes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Var	Highest Grade Attained, Age 26-29							Enrollment	
Sample	All	Boys	Girls	South	North	Urban	Rural	Age 16-18	Age 19-21
Intercept 1940	8.123**	8.311**	8.704**	7.389**	8.972**	9.141**	8.144**	0.461**	0.0119
	(0.209)	(0.291)	(0.297)	(0.322)	(0.295)	(0.292)	(0.239)	(0.00865)	(0.0434)
Intercept 1960	10.25**	10.48**	10.39**	9.654**	10.84**	10.80**	9.986**	0.665**	0.120**
	(0.239)	(0.281)	(0.240)	(0.284)	(0.261)	(0.230)	(0.218)	(0.00818)	(0.0383)
Intercept 1970	11.13**	11.20**	11.55**	11.12**	11.37**			0.793**	0.194**
	(0.218)	(0.212)	(0.213)	(0.298)	(0.193)			(0.00651)	(0.0313)
Intercept 1980	12.01**	12.17**	12.08**	11.62**	12.34**			0.762**	0.220**
	(0.189)	(0.213)	(0.216)	(0.244)	(0.211)			(0.00660)	(0.0317)
Intercept 1990	12.20**	12.20**	12.48**	11.89**	12.52**	12.53**	11.64**	0.829**	0.331**
	(0.154)	(0.242)	(0.244)	(0.259)	(0.247)	(0.197)	(0.224)	(0.00746)	(0.0358)
Intercept 2000	12.49**	12.46**	12.69**	12.27**	12.79**			0.860**	0.355**
	(0.189)	(0.256)	(0.247)	(0.280)	(0.250)			(0.00767)	(0.0371)
Slope 1940	0.365**	0.397**	0.320**	0.536**	0.305**	0.338**	0.463**	0.0391**	0.0302**
	(0.0336)	(0.0539)	(0.0542)	(0.0662)	(0.0525)	(0.0610)	(0.0633)	(0.00160)	(0.00799)
Slope 1960	0.254**	0.286**	0.242**	0.420**	0.200**	0.217**	0.278**	0.0183**	0.0280**
	(0.0409)	(0.0481)	(0.0405)	(0.0548)	(0.0426)	(0.0366)	(0.0444)	(0.00140)	(0.00650)
Slope 1970	0.223**	0.289**	0.152**	0.266**	0.221**			0.0149**	0.0309**
	(0.0373)	(0.0352)	(0.0351)	(0.0512)	(0.0317)			(0.00108)	(0.00519)
Slope 1980	0.199**	0.227**	0.208**	0.292**	0.188**			0.0154**	0.0297**
	(0.0308)	(0.0355)	(0.0359)	(0.0421)	(0.0345)			(0.00110)	(0.00526)
Slope 1990	0.164**	0.184**	0.152**	0.232**	0.147**	0.162**	0.198**	0.0121**	0.0371**
	(0.0245)	(0.0408)	(0.0409)	(0.0451)	(0.0408)	(0.0319)	(0.0409)	(0.00126)	(0.00602)
Slope 2000	0.175**	0.187**	0.199**	0.223**	0.169**			0.0115**	0.0364**
	(0.0301)	(0.0422)	(0.0404)	(0.0475)	(0.0406)			(0.00127)	(0.00611)
Observations	60	60	60	60	60	28	28	60	60
R-squared	1.000	0.999	0.999	0.999	0.999	1.000	0.999	1.000	0.987

Standard errors in parentheses
** p<0.01, * p<0.05

Table X: Mobility Estimates in Parental Income Deciles, Whites

Notes: Displays estimated intercepts and slopes of children's schooling CEFs with respect to parental income deciles for whites. Presents estimates of α_t and β_t from Equation (9). Regressions weighted by square root of estimated cell sizes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent Var	Highest Grade Attained, Age 26-29							Enrollment	
Sample	All	Boys	Girls	South	North	Urban	Rural	Age 16-18	Age 19-21
Intercept 1940	5.170**	5.335**	6.214**	5.466**	7.850**	6.634**	5.059**	0.404**	0.0543**
	(0.314)	(0.327)	(0.344)	(0.251)	(0.398)	(0.339)	(0.208)	(0.00769)	(0.0171)
Intercept 1960	8.271**	8.265**	9.075**	8.088**	10.03**	9.350**	7.400**	0.623**	0.168**
	(0.264)	(0.287)	(0.254)	(0.213)	(0.262)	(0.239)	(0.193)	(0.00632)	(0.0136)
Intercept 1970	10.39**	10.59**	10.83**	10.57**	10.72**			0.764**	0.175**
	(0.253)	(0.195)	(0.204)	(0.199)	(0.119)			(0.00466)	(0.0101)
Intercept 1980	11.62**	11.82**	11.84**	11.66**	12.20**			0.789**	0.260**
	(0.203)	(0.208)	(0.211)	(0.186)	(0.142)			(0.00500)	(0.0106)
Intercept 1990	11.81**	11.82**	12.16**	11.87**	12.16**	12.03**	11.57**	0.838**	0.333**
	(0.160)	(0.233)	(0.244)	(0.206)	(0.167)	(0.179)	(0.247)	(0.00561)	(0.0121)
Intercept 2000	12.16**	12.15**	12.49**	12.33**	12.31**			0.865**	0.321**
	(0.203)	(0.243)	(0.242)	(0.212)	(0.170)			(0.00577)	(0.0124)
Slope 1940	0.648**	0.682**	0.624**	0.676**	0.291**	0.588**	0.744**	0.0475**	0.0270**
	(0.0867)	(0.116)	(0.120)	(0.104)	(0.0989)	(0.120)	(0.120)	(0.00271)	(0.00600)
Slope 1960	0.524**	0.688**	0.486**	0.833**	0.173**	0.336**	1.015**	0.0218**	0.0150**
	(0.0739)	(0.102)	(0.0818)	(0.100)	(0.0609)	(0.0671)	(0.142)	(0.00202)	(0.00434)
Slope 1970	0.255**	0.272**	0.200**	0.284**	0.243**			0.0166**	0.0279**
	(0.0639)	(0.0540)	(0.0575)	(0.0672)	(0.0299)			(0.00127)	(0.00277)
Slope 1980	0.207**	0.192**	0.217**	0.217**	0.149**			0.0135**	0.0260**
	(0.0451)	(0.0485)	(0.0499)	(0.0496)	(0.0297)			(0.00118)	(0.00249)
Slope 1990	0.183**	0.174**	0.194**	0.204**	0.157**	0.181**	0.189**	0.0114**	0.0321**
	(0.0325)	(0.0521)	(0.0544)	(0.0500)	(0.0339)	(0.0390)	(0.0642)	(0.00125)	(0.00269)
Slope 2000	0.166**	0.165**	0.163**	0.177**	0.159**			0.0104**	0.0358**
	(0.0381)	(0.0532)	(0.0541)	(0.0488)	(0.0356)			(0.00129)	(0.00273)
Observations	60	60	60	60	60	28	25	60	60
R-squared	0.999	0.999	0.999	0.999	0.999	0.999	0.998	1.000	0.996

Standard errors in parentheses
** p<0.01, * p<0.05

Table XI: Mobility Estimates in Parental Income Deciles, Blacks

Notes: Displays estimated intercepts and slopes of children's schooling CEFs with respect to parental income deciles for blacks. Estimates correspond to α_t and β_t from Equation (9). Regressions weighted by square root of estimated cell sizes.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
		Whites				Blacks			
Covariate	Model	Estimate	SE	N	R-squared	Estimate	SE	N	R-squared
Log Household Earnings	OLS	-0.120**	(0.00943)	248	0.389	-0.152**	(0.0132)	159	0.519
	FE	-0.289**	(0.0445)	248	0.691	-0.349**	(0.0811)	159	0.730
	FD	-0.276**	(0.0800)	191	0.167	-0.216**	(0.0357)	104	0.550
Earnings Inequality: p75-p25	OLS	0.305**	(0.0377)	248	0.290	0.476**	(0.0410)	159	0.481
	FE	0.308**	(0.0578)	248	0.691	0.372**	(0.0768)	159	0.731
	FD	0.201	(0.121)	191	0.143	0.304**	(0.0914)	104	0.547
Urban Share	OLS	-0.463**	(0.103)	130	0.337	-0.776**	(0.106)	77	0.483
	FE	-0.162	(0.153)	130	0.756	-0.921**	(0.307)	77	0.817
	FD	-0.129	(0.124)	42	0.021	-0.160	(0.312)	16	0.012
Share Black	Basic	0.345**	(0.0552)	249	0.128	0.264**	(0.0802)	157	0.074
	FE	0.650**	(0.191)	249	0.658	0.938*	(0.360)	157	0.743
	FD	1.081**	(0.254)	192	0.152	0.777*	(0.348)	104	0.519
Teen Birth Rate	OLS	1.359**	(0.292)	249	0.111	0.823*	(0.306)	157	0.028
	FE	1.282*	(0.548)	249	0.644	2.431**	(0.743)	157	0.691
	FD	0.303	(0.589)	192	0.117	-0.145	(0.973)	104	0.494
Teen Employment Rate	OLS	-0.339**	(0.101)	249	0.048	0.233	(0.169)	157	0.014
	FE	0.244	(0.138)	249	0.632	0.866**	(0.210)	157	0.723
	FD	0.0254	(0.175)	192	0.115	0.155	(0.294)	104	0.495
Dropout Age	OLS	-0.0305*	(0.0147)	178	0.035	-0.0711**	(0.0253)	106	0.125
	FE	-0.0319	(0.0180)	178	0.681	-0.0529	(0.0282)	106	0.753
	FD	-0.0373*	(0.0154)	122	0.164	-0.0322	(0.0321)	53	0.367
Class Size	OLS	0.0164**	(0.00209)	136	0.316	0.0263**	(0.00260)	77	0.520
	FE	0.00533	(0.00544)	136	0.690	0.0169	(0.0127)	77	0.844
	FD	0.00383	(0.00484)	84	0.092	0.0111	(0.0108)	32	0.193
Relative Teacher Wage	OLS	-0.224**	(0.0740)	136	0.114	-0.163	(0.0859)	77	0.056
	FE	-0.144*	(0.0576)	136	0.701	-0.257**	(0.0848)	77	0.875
	FD	-0.180**	(0.0558)	84	0.141	-0.241**	(0.0670)	32	0.267
Share Move State	OLS	0.168**	(0.0497)	248	0.083	0.131*	(0.0614)	159	0.033
	FE	0.242*	(0.115)	248	0.651	0.597**	(0.139)	159	0.739
	FD	0.139	(0.144)	191	0.117	0.253	(0.197)	104	0.501

Table XII: Mobility Regressions on State-Year Panel 1940-2000

Notes: Table displays estimates from bivariate regressions of educational mobility CEF slopes (child education on parent education) on various covariates. OLS, fixed effect and first-difference models described in text. Standard errors clustered at state-of-birth level. All regressions unweighted. Mobility varies by year, state-of-birth and race. All other variables vary by year and state-of-birth, or year and state. White and black samples vary due to requirement that CEFs underlying mobility statistics be non-missing for at least 60% of parental group levels. Mobility statistics for state-of-birth panel analysis constructed on children age 20-29. Dropout age, class size and relative teacher wage matched to to census from year at which age 20-29 year-old children would have been approximately age 14, and are only matched to censuses 1940-70 using data from Stephens and Yang (2014). Log household earnings and earnings inequality constructed from household earnings distribution (head + spouse) for heads of household age 30-65 in year children turn 20-29 and living in child's state of birth. Migration and urbanization constructed from age 20-29 year-olds by state of birth and year.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	White						Black					
Household Earnings	Slope w.r.t. Parental Education			Slope w.r.t. Parental Income Decile			Slope w.r.t. Parental Education			Slope w.r.t. Parental Income Decile		
Log Household Earnings	-0.0907**	-0.173*	-0.229*	-0.138**	-0.0459	-0.107	-0.0974**	-0.200	-0.148**	-0.154**	-0.368**	0.0207
	(0.00882)	(0.0706)	(0.111)	(0.0118)	(0.109)	(0.124)	(0.0143)	(0.109)	(0.0389)	(0.0243)	(0.132)	(0.0725)
Earnings Inequality: p75-p25	0.133**	0.185	0.107	0.109*	0.228	0.204	0.257**	0.221*	0.200**	0.305**	0.406**	0.440*
	(0.0400)	(0.0931)	(0.158)	(0.0422)	(0.122)	(0.137)	(0.0391)	(0.100)	(0.0696)	(0.0658)	(0.136)	(0.180)
Observations	248	248	191	294	294	234	159	159	104	204	204	151
R-squared	0.421	0.702	0.174	0.460	0.675	0.210	0.592	0.742	0.568	0.457	0.735	0.210
Model	OLS	FE	FD	OLS	FE	FD	OLS	FE	FD	OLS	FE	FD

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table XIII: Mobility Regressions on Income Level and Inequality in State-Year Panel, 1940-2000

Notes: Table presents OLS, FE and FD regressions of mobility by state-of-birth, year and race on average log of household earnings and interquartile gap in log of household earnings. Constants omitted from table for convenience.

	(1)	(2)	(3)	(4)	(5)	(6)
Specification:	Basic	Basic	Impute Inc	Impute Inc	Scaling Deps	Scaling Deps
Race:	White	Black	White	Black	White	Black
Intercept 1940	8.123**	5.170**	7.948**	5.186**	8.579**	5.854**
	(0.209)	(0.314)	(0.250)	(0.252)	(0.240)	(0.283)
Intercept 1960	10.25**	8.271**	10.20**	8.461**	10.53**	8.821**
	(0.239)	(0.264)	(0.265)	(0.268)	(0.211)	(0.225)
Intercept 1970	11.13**	10.39**	11.17**	10.40**	11.42**	10.75**
	(0.218)	(0.253)	(0.224)	(0.192)	(0.173)	(0.168)
Intercept 1980	12.01**	11.62**	11.95**	11.42**	12.24**	11.91**
	(0.189)	(0.203)	(0.219)	(0.167)	(0.175)	(0.177)
Intercept 1990	12.20**	11.81**	11.98**	11.68**	12.40**	12.07**
	(0.154)	(0.160)	(0.252)	(0.187)	(0.198)	(0.201)
Intercept 2000	12.49**	12.16**	12.25**	11.89**	12.70**	12.40**
	(0.189)	(0.203)	(0.256)	(0.195)	(0.205)	(0.206)
Slope 1940	0.365**	0.648**	0.456**	0.693**	0.362**	0.658**
	(0.0336)	(0.0867)	(0.0457)	(0.0829)	(0.0460)	(0.105)
Slope 1960	0.254**	0.524**	0.287**	0.474**	0.265**	0.531**
	(0.0409)	(0.0739)	(0.0444)	(0.0759)	(0.0373)	(0.0762)
Slope 1970	0.223**	0.255**	0.244**	0.287**	0.227**	0.244**
	(0.0373)	(0.0639)	(0.0364)	(0.0499)	(0.0298)	(0.0488)
Slope 1980	0.199**	0.207**	0.235**	0.298**	0.210**	0.203**
	(0.0308)	(0.0451)	(0.0350)	(0.0398)	(0.0302)	(0.0441)
Slope 1990	0.164**	0.183**	0.218**	0.246**	0.168**	0.181**
	(0.0245)	(0.0325)	(0.0406)	(0.0428)	(0.0345)	(0.0474)
Slope 2000	0.175**	0.166**	0.228**	0.240**	0.183**	0.165**
	(0.0301)	(0.0381)	(0.0400)	(0.0427)	(0.0350)	(0.0481)
Observations	60	60	60	60	60	60
R-squared	1.000	0.999	0.999	0.999	1.000	0.999

Standard errors in parentheses
** p<0.01, * p<0.05

Table XIV: Educational IM Under Alternative Assumptions, 1940-2000

Notes: Displays estimated intercepts and slopes of children's schooling CEFs with respect to parental income deciles for blacks. Estimates correspond to α_t and β_t from Equation (9). Regressions weighted by square root of estimated cell sizes. Columns (1)-(2) replicate basic results with respect to parental income deciles which exclude families with zero parental income. Columns (3)-(4) include families with zero parental income, imputing their income based on demographic characteristics as described in text. Columns (5)-(6) scale up the number of dependents at age 17 disproportionately for lower-SES families to account for children who have already become independent at these ages, then re-calculates the adjustment for missing independent children.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep Var: Educ Group Cohort Share Ages 16-17	All Years	All Years	All Years	1940	1950	1960	1970	1980	1990	2000
Age 7	1.023**			0.974**	1.162**	1.094**	1.146**	1.078**	0.786**	0.647**
	(0.0292)			(0.0196)	(0.132)	(0.0877)	(0.0660)	(0.0373)	(0.0378)	(0.0412)
Ages 4-7, linear		0.833**								
		(0.0429)								
Ages 1-7, linear			0.919**							
			(0.0305)							
Constant	-0.00230	0.0167**	0.00811*	0.00262	-0.0162	-0.00944	-0.0146	-0.00784	0.0214**	0.0353**
	(0.00343)	(0.00520)	(0.00368)	(0.00269)	(0.0159)	(0.0110)	(0.00753)	(0.00397)	(0.00404)	(0.00449)
Observations	140	140	140	20	20	20	20	20	20	20
R-squared	0.899	0.732	0.868	0.993	0.811	0.896	0.944	0.979	0.960	0.932

Standard errors in parentheses
** p<0.01, * p<0.05

Table A.1: Validation of Group Cohort Size Predictors in Parental Education, Blacks

Notes: Columns (1)-(3) regress actual parental education group cohort shares at ages 16-17 on predicted parental education group cohort shares at ages 16-17 and a constant, where each column uses a different predictor. Columns (4)-(10) run the regression in Column (1) separately by census year 1940-2000. Regressions for black-only sample. All regressions weighted by the square root of the cell size.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Dep Var: Income Group Cohort Share Ages 16-17	All Years	All Years	All Years	1940	1950	1960	1970	1980	1990	2000
Age 7	1.052**			1.022**	1.134**	1.205**	1.165**	1.051**	0.839**	0.761**
	(0.0162)			(0.0141)	(0.123)	(0.0552)	(0.0472)	(0.0192)	(0.0174)	(0.0234)
Ages 4-7, linear		0.903**								
		(0.0225)								
Ages 1-7, linear			0.983**							
			(0.0159)							
Constant	-0.00531*	0.0134**	0.00252	-0.00561	-0.0102	-0.0253*	-0.0174*	-0.00511*	0.0164**	0.0246**
	(0.00232)	(0.00329)	(0.00231)	(0.00290)	(0.0188)	(0.00970)	(0.00675)	(0.00231)	(0.00208)	(0.00288)
Observations	280	280	280	40	40	40	40	40	40	40
R-squared	0.938	0.853	0.933	0.993	0.692	0.926	0.941	0.987	0.984	0.965

Standard errors in parentheses
** p<0.01, * p<0.05

Table A.2: Validation of Group Cohort Size Predictors in Parental Income, Blacks

Notes: Columns (1)-(3) regress actual parental income group cohort shares at ages 16-17 on predicted parental income group cohort shares at ages 16-17 and a constant, where each column uses a different predictor. Columns (4)-(10) run the regression in Column (1) separately by census year 1940-2000. Regressions for black-only sample. All regressions weighted by the square root of the cell size.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
				PSID							NLSY79					NLSY97		
Child Educ	All	1980-89	1990-99	2000-09	Male	Female	White	Black	All	Male	Female	White	Black	All	Male	Female	White	Black
Par Income	0.349**	0.274**	0.383**	0.424**	0.367**	0.333**	0.338**	0.299**	0.327**	0.316**	0.338**	0.324**	0.237**	0.357**	0.351**	0.369**	0.355**	0.283**
	(0.00886)	(0.0133)	(0.0166)	(0.0214)	(0.0129)	(0.0121)	(0.0116)	(0.0156)	(0.0115)	(0.0161)	(0.0163)	(0.0129)	(0.0387)	(0.00996)	(0.0148)	(0.0133)	(0.0120)	(0.0294)
Dep	-0.289	-0.187	0.00920	-1.055**	-0.194	-0.169	-0.287	0.103	-0.649**	-0.754**	-0.517	-0.751**	-0.194	-0.453**	-0.270	-0.515*	-0.588**	-0.764**
	(0.178)	(0.286)	(0.330)	(0.374)	(0.226)	(0.293)	(0.265)	(0.172)	(0.181)	(0.234)	(0.289)	(0.223)	(0.269)	(0.136)	(0.184)	(0.204)	(0.180)	(0.267)
Par Inc * Dep	-0.0479	-0.00973	-0.123*	0.0766	-0.0867*	-0.000913	-0.0366	-0.144**	0.0592	0.0367	0.104*	0.0755*	0.0797	-0.0462*	-0.0413	-0.0572	-0.0355	0.0399
	(0.0290)	(0.0478)	(0.0523)	(0.0618)	(0.0371)	(0.0469)	(0.0407)	(0.0375)	(0.0304)	(0.0392)	(0.0487)	(0.0359)	(0.0792)	(0.0235)	(0.0316)	(0.0351)	(0.0298)	(0.0573)
Constant	2.697**	3.028**	2.511**	2.431**	2.457**	2.917**	2.804**	2.528**	2.906**	2.950**	2.857**	2.944**	2.769**	2.891**	2.596**	3.121**	2.952**	3.105**
	(0.0556)	(0.0830)	(0.104)	(0.133)	(0.0809)	(0.0762)	(0.0754)	(0.0673)	(0.0693)	(0.0974)	(0.0984)	(0.0795)	(0.142)	(0.0622)	(0.0934)	(0.0825)	(0.0774)	(0.139)
Observations	15,001	6,555	4,386	2,797	7,230	7,771	9,267	5,633	9,687	4,926	4,761	7,815	1,134	10,201	4,959	5,242	7,198	1,485
R-squared	0.103	0.066	0.118	0.148	0.113	0.094	0.091	0.067	0.095	0.094	0.099	0.090	0.049	0.141	0.132	0.154	0.139	0.093

Robust standard errors in parentheses
** p<0.01, * p<0.05

Table A.3: Parallel Trends for Child Education Rank in Parental Income Decile

Notes: Each column displays estimates of equation (6) for children's education rank at ages 26-29 by parental income deciles. Parental characteristics measured when children are age 17, or earlier in adolescence if not observed at age 17. Children's schooling excludes zeros. Parental income defined as sum of mother's and father's earnings when possible, otherwise as total family income. Children with zero parental income at age 17 excluded from the parental income figures. Parental income deciles calculated separately by year. Sample weights used in all calculations. Regressions on collapsed data weighted by square of cell size.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
				PSID							NLSY79					NLSY97			GSS	OCG73
Child Educ	All	1980-89	1990-99	2000-09	Male	Female	White	Black	All	Male	Female	White	Black	All	Male	Female	White	Black	All	All
Par Educ	0.536**	0.513**	0.574**	0.520**	0.551**	0.525**	0.533**	0.422**	0.564**	0.566**	0.562**	0.568**	0.352**	0.392**	0.401**	0.387**	0.396**	0.308**	0.353**	0.496**
	(0.00935)	(0.0201)	(0.0151)	(0.0182)	(0.0137)	(0.0128)	(0.0124)	(0.0185)	(0.00970)	(0.0140)	(0.0135)	(0.0108)	(0.0340)	(0.00710)	(0.00988)	(0.0100)	(0.00847)	(0.0203)	(0.0188)	(0.0175)
Dep	0.0655	0.579	0.554*	-0.872**	0.0349	0.310	0.0499	-0.0169	0.0508	-0.152	0.424	0.0888	-0.730*	-0.433**	-0.459**	-0.186	-0.403**	-1.013**	0.299	0.0839
	(0.165)	(0.424)	(0.266)	(0.296)	(0.212)	(0.268)	(0.253)	(0.172)	(0.146)	(0.192)	(0.230)	(0.172)	(0.294)	(0.0880)	(0.113)	(0.138)	(0.114)	(0.187)	(0.279)	(0.289)
Par Educ * Dep	-0.101**	-0.123	-0.192**	0.0761	-0.0908*	-0.133**	-0.0891*	-0.122**	-0.0672*	-0.0441	-0.108*	-0.0727*	0.232**	-0.00119	0.0121	-0.0272	-0.0177	0.198**	-0.0718	-0.104*
	(0.0295)	(0.0817)	(0.0453)	(0.0542)	(0.0377)	(0.0483)	(0.0417)	(0.0467)	(0.0266)	(0.0346)	(0.0421)	(0.0303)	(0.0730)	(0.0162)	(0.0206)	(0.0259)	(0.0202)	(0.0402)	(0.0514)	(0.0516)
Constant	1.721**	1.725**	1.425**	1.990**	1.514**	1.890**	1.798**	1.974**	1.876**	1.812**	1.931**	1.863**	2.256**	3.009**	2.657**	3.331**	3.056**	3.011**	3.160**	2.091**
	(0.0558)	(0.119)	(0.0910)	(0.107)	(0.0828)	(0.0753)	(0.0775)	(0.0758)	(0.0554)	(0.0811)	(0.0757)	(0.0625)	(0.142)	(0.0419)	(0.0588)	(0.0589)	(0.0517)	(0.0998)	(0.105)	(0.104)
Observations	10,715	2,219	4,364	2,770	4,999	5,716	6,222	4,203	10,786	5,394	5,392	8,769	1,265	16,993	8,627	8,366	11,852	2,671	3,456	3,015
R-squared	0.251	0.234	0.261	0.262	0.268	0.238	0.242	0.125	0.261	0.266	0.258	0.261	0.131	0.190	0.210	0.174	0.189	0.146	0.101	0.228

Robust standard errors in parentheses
** $p<0.01$, * $p<0.05$

Table A.4: Parallel Trends for Child Education Rank in Parental Education Rank

Notes: Each column displays estimates of equation (6) for children's education rank at ages 26-29 by parental education rank. Parental characteristics measured when children are age 17, or earlier in adolescence if not observed at age 17. Children's schooling excludes zeros. Parental education rank defined as the percentile of the average of mother's and father's educational attainment, where ties are assigned the minimum rank over the tied interval. Sample weights used in all calculations. Regressions on collapsed data weighted by square of cell size.

	(1)	(2)	(3)	(4)
Dependent Var:	Child's Education Rank, Ages 26-29			
Parental Groups:	Parental Income Rank		Parental Education Rank	
Sample:	White	Black	White	Black
Intercept 1940	31.74**	14.24**	34.01**	13.83**
	(2.981)	(2.181)	(1.386)	(1.514)
Intercept 1960	38.85**	24.79**	36.79**	26.75**
	(2.628)	(1.735)	(1.469)	(1.458)
Intercept 1970	37.55**	30.92**	33.54**	29.82**
	(2.149)	(1.288)	(1.390)	(1.327)
Intercept 1980	37.25**	33.01**	33.14**	30.69**
	(2.175)	(1.356)	(1.372)	(1.146)
Intercept 1990	36.26**	31.32**	32.70**	29.93**
	(2.459)	(1.541)	(1.518)	(1.328)
Intercept 2000	35.77**	31.39**	32.87**	28.55**
	(2.549)	(1.578)	(1.525)	(1.350)
Slope 1940	0.374**	0.520**	0.439**	0.554**
	(0.0548)	(0.0766)	(0.0310)	(0.0546)
Slope 1960	0.260**	0.431**	0.321**	0.350**
	(0.0446)	(0.0555)	(0.0266)	(0.0448)
Slope 1970	0.268**	0.262**	0.352**	0.307**
	(0.0356)	(0.0354)	(0.0231)	(0.0327)
Slope 1980	0.249**	0.234**	0.365**	0.310**
	(0.0361)	(0.0318)	(0.0234)	(0.0267)
Slope 1990	0.246**	0.271**	0.369**	0.328**
	(0.0413)	(0.0344)	(0.0262)	(0.0291)
Slope 2000	0.259**	0.239**	0.363**	0.339**
	(0.0420)	(0.0349)	(0.0259)	(0.0287)
Observations	60	60	59	59
R-squared	0.996	0.995	0.998	0.997

Standard errors in parentheses
** p<0.01, * p<0.05

Table A.5: Rank-Rank Educational Mobility Estimates in Parental Income and Education

Notes: Displays estimated intercepts and slopes of age 26-29 children's schooling rank CEFs with respect to parental income and education ranks. All ranks on scale from 0 to 100. Ranks computed on full population for each age and year for children, and for each year for parents pooling ages 26-65. Parental income rank ignores zeros. Ties in all rankings are assigned midpoints of rank intervals. Bottom 2% of reported educational attainment in each year dropped from sample. Estimates correspond to α_t and β_t from Equation (9). Regressions weighted by square root of estimated cell sizes.

	White					Black				
Variable	Mean	Std Dev	Min	Max	Count	Mean	Std Dev	Min	Max	Count
Mobility: Parental Educ	0.37	0.11	0.00	0.72	256	0.31	0.13	0.05	0.69	159
Mobility: Parental Income	0.21	0.14	-0.28	0.70	305	0.22	0.20	-0.40	0.90	204
Log Household Earnings	10.21	0.56	8.31	11.01	297	10.20	0.60	8.31	11.01	246
Earnings Inequality: p75-p25	0.99	0.19	0.61	1.66	297	1.01	0.19	0.65	1.66	246
Urban Share	0.64	0.17	0.00	0.94	153	0.66	0.14	0.36	0.94	123
Share Black	0.10	0.13	0.00	0.71	297	0.12	0.13	0.00	0.71	241
Teen Birth Rate	0.07	0.03	0.01	0.17	297	0.07	0.03	0.01	0.16	241
Teen Employment Rate	0.36	0.08	0.16	0.56	297	0.36	0.08	0.17	0.53	241
Dropout Age	16.22	0.75	14.00	18.00	192	16.17	0.75	14.00	18.00	150
Class Size	24.85	4.23	17.30	37.59	147	25.79	4.31	17.30	37.59	109
Relative Teacher Wage	1.00	0.18	0.64	2.04	147	1.00	0.20	0.64	2.04	109
Share Move State	0.54	0.20	0.02	0.86	297	0.55	0.20	0.02	0.86	246
Black HS per Capita	0.0025	0.0019	0.0000	0.0091	143	0.0021	0.0017	0.0000	0.0091	105
White HS per Capita	0.0029	0.0016	0.0000	0.0076	143	0.0026	0.0014	0.0004	0.0068	105
Black HS PC - South	0.0015	0.0018	0.0000	0.0091	47	0.0016	0.0018	0.0000	0.0091	46
White HS PC - South	0.0027	0.0012	0.0006	0.0061	47	0.0027	0.0012	0.0006	0.0061	46

Table A.6: Summary Statistics for Variables in Panel Analysis

Notes: Table displays summary statistics for whites and blacks separately. Mobility varies by year, state-of-birth and race. All other variables vary by year and state-of-birth, or year and state. White and black samples vary due to requirement that CEFs underlying mobility statistics be non-missing for at least 60% of parental group levels. Mobility statistics for state-of-birth panel analysis constructed on children age 20-29. "HS per capita" and "HS PC" refer to number of public high schools per age 14-17 year-old children in state-of-birth and year, and are only matched to census years 1940, 1960 and 1970 from high school data sets in years 1928, 1949 and 1952, respectively to match age at high school attendance as closely as possible with available data. Dropout age, class size and relative teacher wage matched to to census from year at which age 20-29 year-old children would have been approximately age 14, and are only matched to censuses 1940-70 using data from Stephens and Yang (2014). Log household earnings and earnings inequality constructed from household earnings distribution (head + spouse) for heads of household age 30-65 in year children turn 20-29 and living in child's state of birth. Migration and urbanization constructed from age 20-29 year-olds by state of birth and year.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		OCG1973		OCG1962		OCG1973		OCG1962	
	Cohort 1920	Cohort 1930	Cohort 1940	Cohort 1920	Cohort 1930				
VARIABLES			Log Earnings			Education	Log Earnings	Education	Log Earnings
Education	0.094**	0.085**	0.075**	0.075*	0.080**				
	(0.012)	(0.015)	(0.015)	(0.032)	(0.017)				
Father HS Grad	-0.188	-0.039	-0.016	-0.200	0.207				
	(0.223)	(0.264)	(0.279)	(0.599)	(0.322)				
Father Some College	0.197	-0.195	0.099	-0.406	0.144				
	(0.223)	(0.264)	(0.279)	(0.599)	(0.322)				
Education * Father HS Grad	0.017	0.002	0.003	0.021	-0.004				
	(0.017)	(0.021)	(0.022)	(0.045)	(0.024)				
Education * Father Some College	-0.006	0.010	-0.004	0.039	-0.000				
	(0.017)	(0.021)	(0.022)	(0.045)	(0.024)				
Father's Education						0.429**	0.055**	0.476**	0.057**
						(0.037)	(0.006)	(0.022)	(0.014)
Cohort1930						0.707	0.232*	1.151*	-0.131
						(0.648)	(0.100)	(0.385)	(0.242)
Cohort1940						1.070	0.087		
						(0.648)	(0.100)		
Cohort1930*Father's Education						-0.025	-0.014	-0.057	0.005
						(0.052)	(0.008)	(0.031)	(0.020)
Cohort1940*Father's Education						-0.041	-0.010		
						(0.052)	(0.008)		
Constant	8.104**	8.302**	8.298**	7.783**	7.566**	7.898**	8.733**	7.029**	8.118**
	(0.158)	(0.187)	(0.198)	(0.423)	(0.227)	(0.458)	(0.071)	(0.272)	(0.171)
Observations	30	30	30	18	18	33	33	10	10
R-squared	0.892	0.826	0.750	0.696	0.851	0.933	0.891	0.993	0.864

Standard errors in parentheses
** p<0.01, * p<0.05

Table A.7: Returns to Schooling by Parental Group and IM Statistics in OCG

Notes: Documents similar returns to schooling across father's education groups, and children's schooling and log earnings CEFs in father's education, together suggesting that educational mobility gains likely increased income mobility gains. Displays various regressions on OCG1962 and OCG1973 data. Columns (1)-(5) assess whether the returns to schooling differ by parental background. Columns (6)-(9) assess whether changes in mobility can be observed directly in OCG data. All regresssions restricted to whites. Underlying data collapsed prior to regression to child's education by father's education by cohort cells in columns (1)-(5) and to father's education by cohort level in columns (6)-(9). Sample weights used in collapse but not in regressions. Note father's education can take on any integer values between 7 and 17 in OCG1973 sample, but can only take on values of 8, 10, 12, 14, 16 in OCG1962. Columns (1)-(5) regress log of earnings on years of schooling, father's education, and the interaction of these two variables, separately for each 10-year birth cohort and OCG data set. Columns (6) and (7) regress education and log of earnings, respectively, on father's education, 10-year cohort dummies, and the interaction of these two variables in the OCG73 data. Columns (8) and (9) repeat the exercise in columns (6) and (7), respectively, on the OCG62 data.

	White				Black			
	Parental Income		Parental Education		Parental Income		Parental Education	
Year	% Missing	% Zero	% Missing	% Zero	% Missing	% Zero	% Missing	% Zero
1940	11.7	40.1	3.8	6.1	13.1	41.3	5.2	11.9
1960	5.4	26.5	0.2	2.1	6.5	26.4	0.5	4.7
1970	4.2	16.1	4.8	0.5	5.4	21.0	8.3	2.3
1980	16.2	12.8	5.7	0.3	22.9	23.3	10.2	1.2
1990	18.3	9.7	2.7	0.5	24.2	18.7	4.2	1.2
2000	19.3	11.4	2.8	0.5	27.3	18.6	4.7	0.7

Table A.8: Percent of Children with Missing and Zero Values for Parental Characteristics, by Year and Race

Notes: Sample restricted to children living with parents age 26-29 in census data.

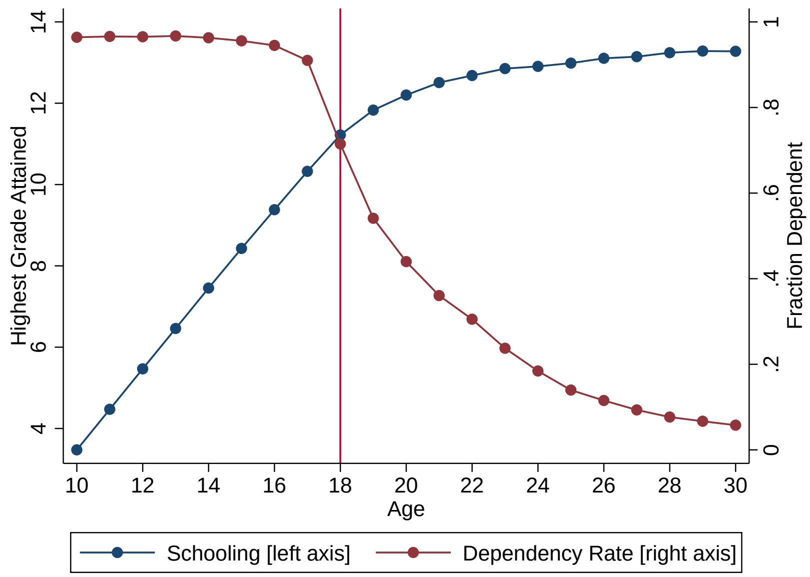
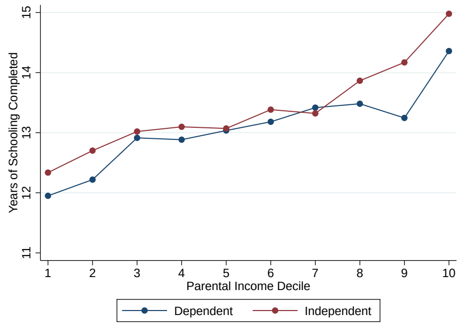
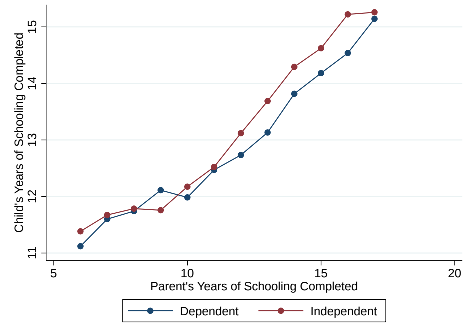


Figure I: Schooling and Dependency Status by Age in 1980

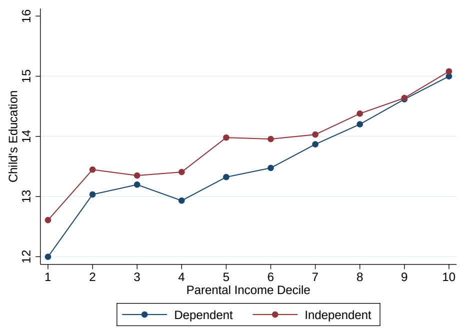
Notes: Red line plots fraction of native-born white children living with parents by age in 1980. Blue line plots average schooling of native-born children by age in 1980.



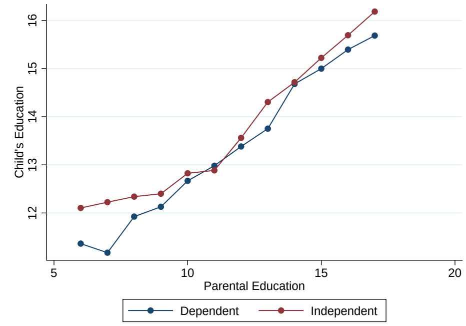
(a) Parental Income, PSID



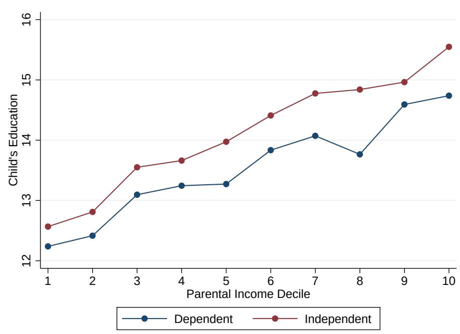
(b) Parental Education, PSID



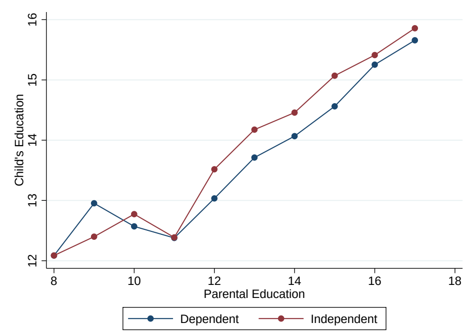
(c) Parental Income, NLSY79



(d) Parental Education, NLSY79



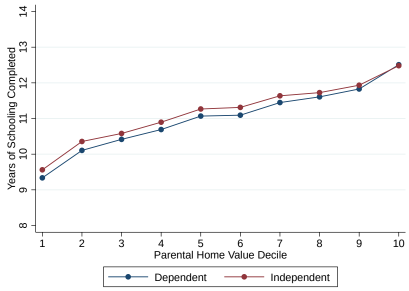
(e) Parental Income, NLSY97



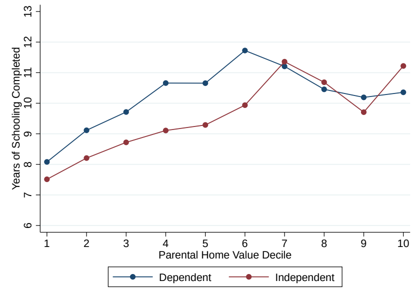
(f) Parental Education, NLSY97

Figure II: Highest Grade Attained at Ages 26-29 by Parental Characteristics at Age 17

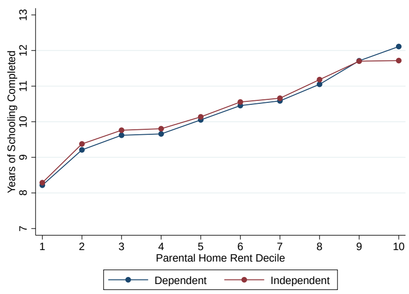
Notes: Figures based on data from PSID, NLSY79, and NLSY97 pooling years 1968-2011, 1994-2011, and 1997-2011 respectively. Parental characteristics measured when children are age 17, or earlier in adolescence if not observed at age 17. Children's schooling at ages 26-29 set to missing when reported as zero. Children with zero parental income at age 17 excluded from parental income figures. Income deciles calculated separately by year. Parental education defined as average of mother's and father's education, or education of single parent when necessary. Children with parental education in bottom 2% of distribution excluded from parental education figures. Sample weights used in all calculations.



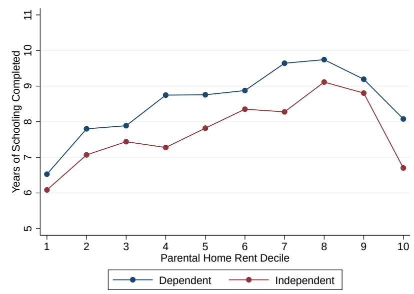
(a) By Parental Home Value, Whites



(b) By Parental Home Value, Blacks



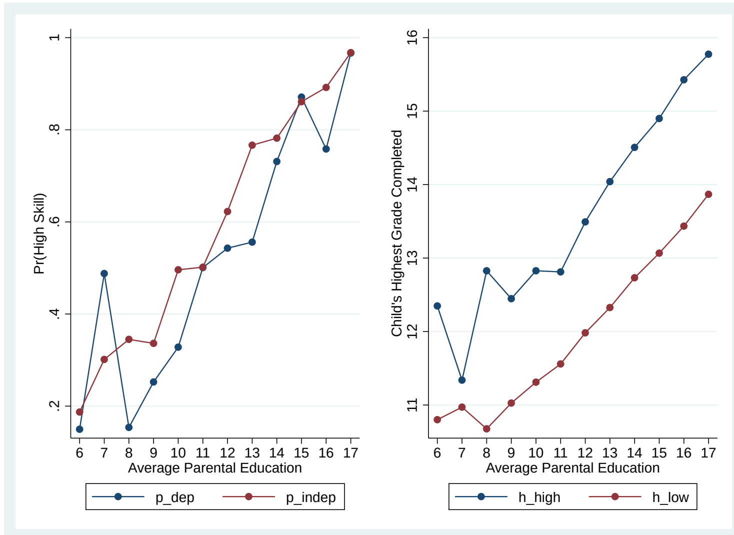
(c) By Parental Rent, Whites



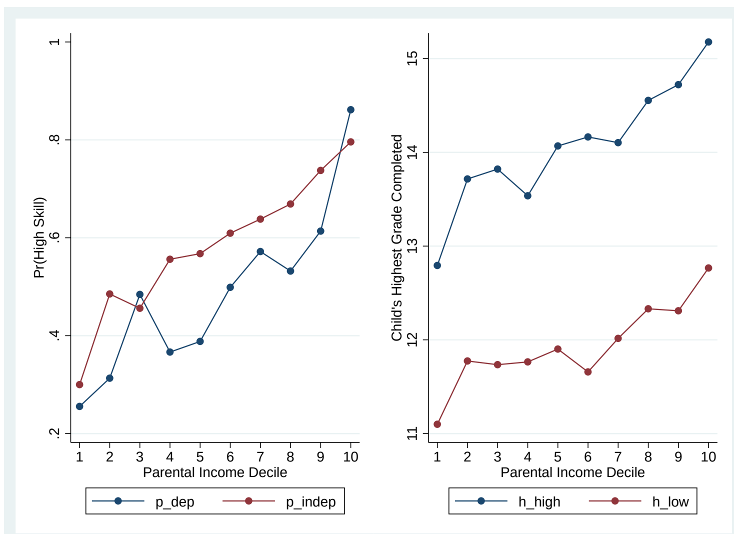
(d) By Parental Rent, Blacks

Figure III: Final Schooling at Ages 24-27 in 1940 by Parental Group Status

Notes: Figures plot highest grade attained for ages 24-27 by parental home value or rent deciles based on matched 1930-1940 census data. Families with zero rent and earnings in 1930 excluded. Deciles calculated on full population of parents with any children age 10-17 in 1930, including all non-farm owner-occupied or renter-occupied units, weighting by number of children.



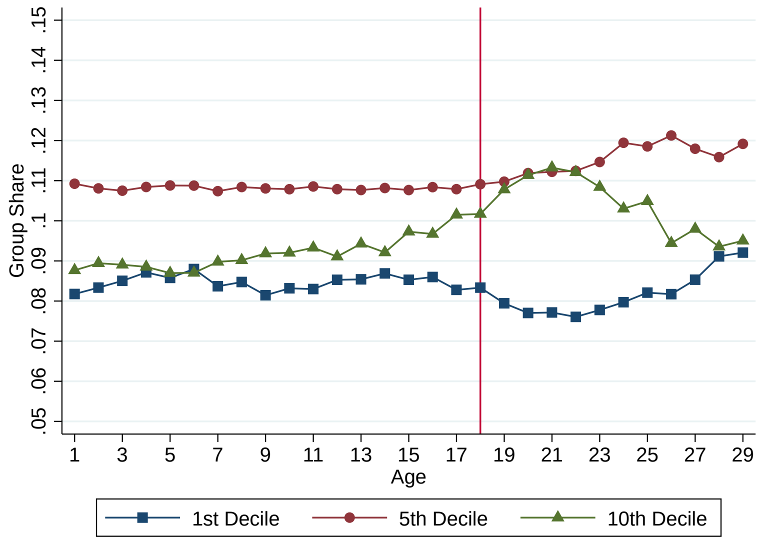
(a) Parental Education



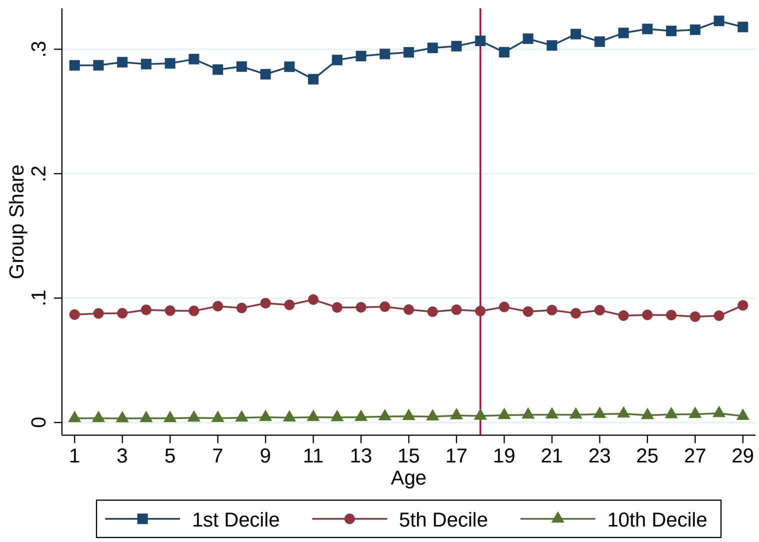
(b) Parental Income

Figure IV: Ability Composition and Educational Attainment by Parental Group Status

Notes: Figure shows that conditions required for parallel trends appear reasonable in data from NLSY79. The left figure in panel (a) displays the probability that children are above-median AFQT separately for dependents and independents by average parental educational attainment. The right figure in panel (a) displays final schooling for above- and below-median AFQT children at ages 26-29 by parental educational attainment. Parental education defined as average of mother and father, or education of one parent if education of both parents not available. Panel (b) repeats panel (a) but with respect to parental income deciles. Parental income deciles defined using sum of mother's and father's earnings when child age 17.



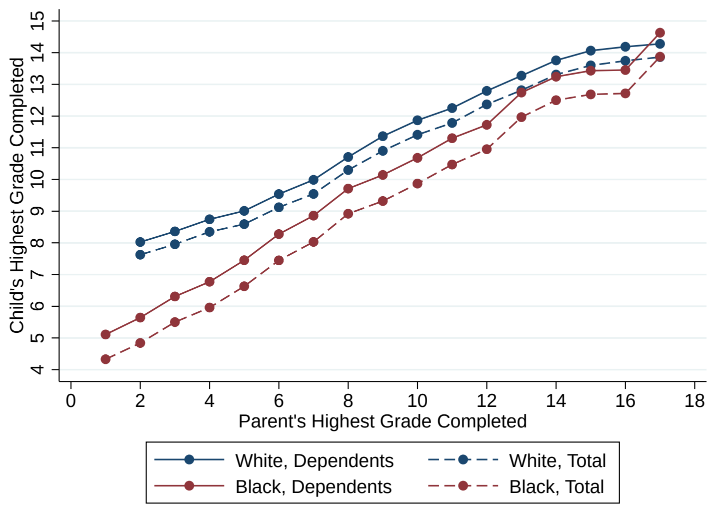
(a) Whites



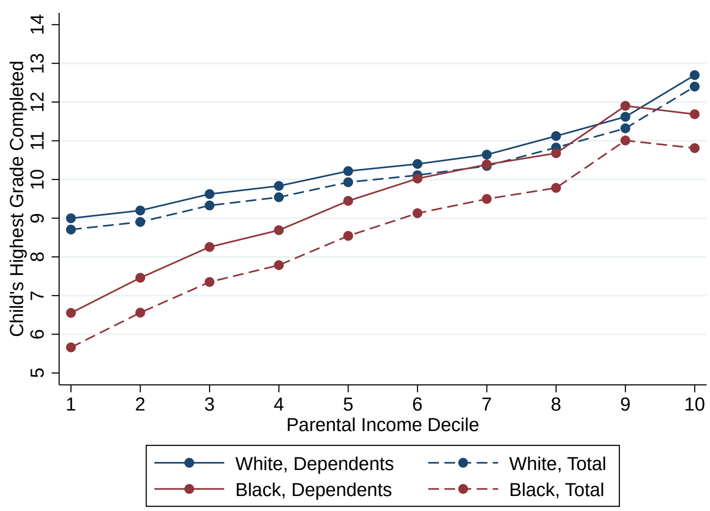
(b) Blacks

Figure V: Share of Dependent Children in Parental Income Decile by Age, 1940

Notes: Figures plot frequencies for white native-born children living with parents by age and race in 1940 100% IPUMS data sample.



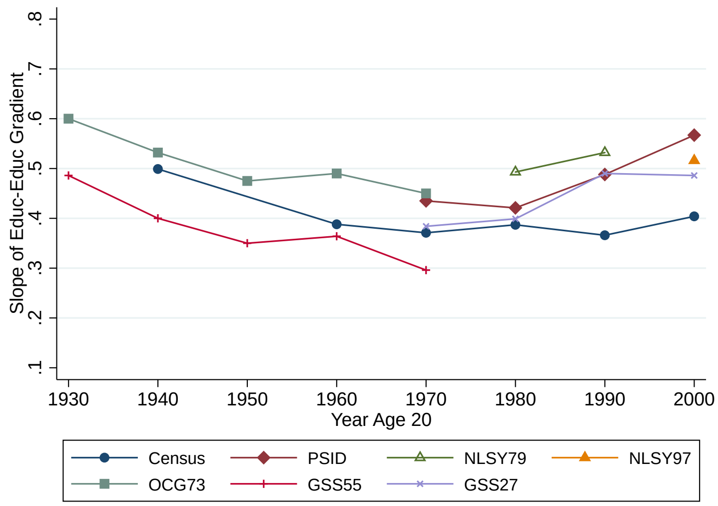
(a) Education CEF in Parental Education, Before and After Adjustment



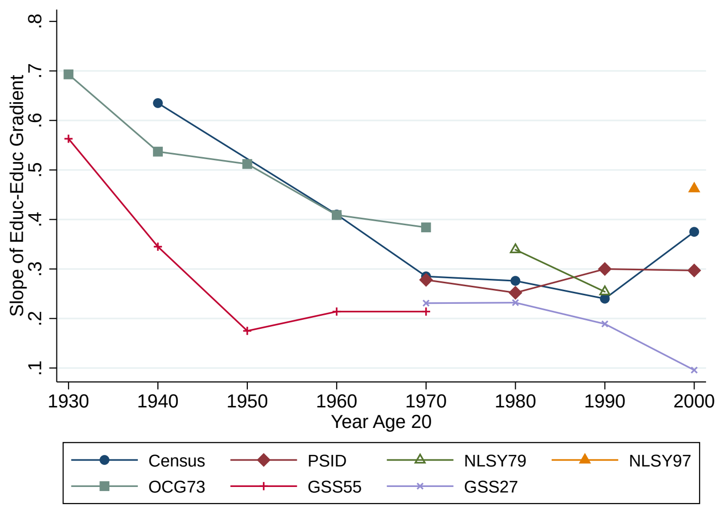
(b) Education CEF in Parental Income Decile, Before and After Adjustment

Figure VI: Final Schooling Attainment at ages 26-29 by Parental Group Status, 1940

Notes: Figure plots estimated final schooling pooling separate estimates for ages 26-29, using the correction for independent children described in the text. Uncorrected estimates restrict to dependent children who can be linked with parents directly.



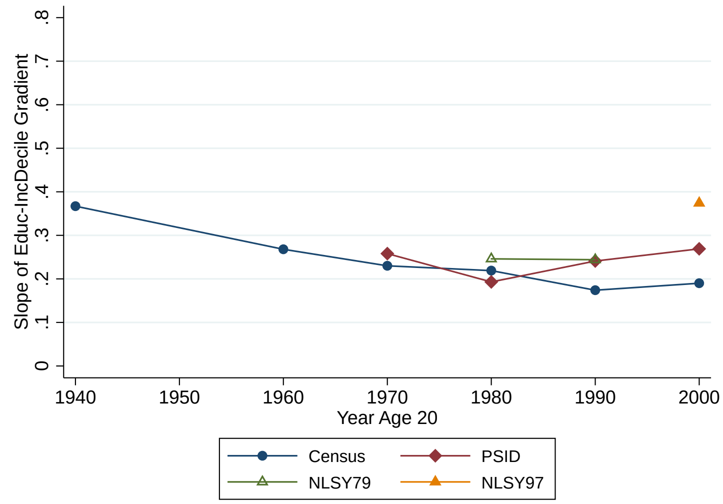
(a) Whites



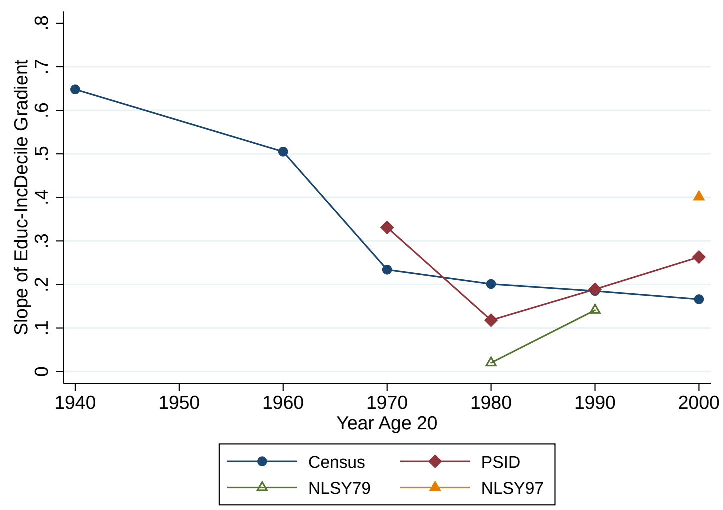
(b) Blacks

Figure VII: Educational Mobility in Parental Education, 1930-2000

Notes: Figure plots slopes from regression of child education on parental education by year. Child and parental education defined and adjusted for independentts as described in text. "Year" in census denotes year cohorts turn ages 26-29. "Year" in PSID, NLSY79, and NLSY97 denotes decades, e.g. "1980" reflects cohorts turning 26-29 in 1980-89. "Year" in OCG73, GSS55 and GSS27 defined as year cohorts would have turned 20-29. "GSS55" and "GSS27" refer to cohorts in the GSS age 55-65 and 27-37, respectively, over the years 1972-2012. All estimates use sample weights and exclude bottom 2% of parental education distribution.



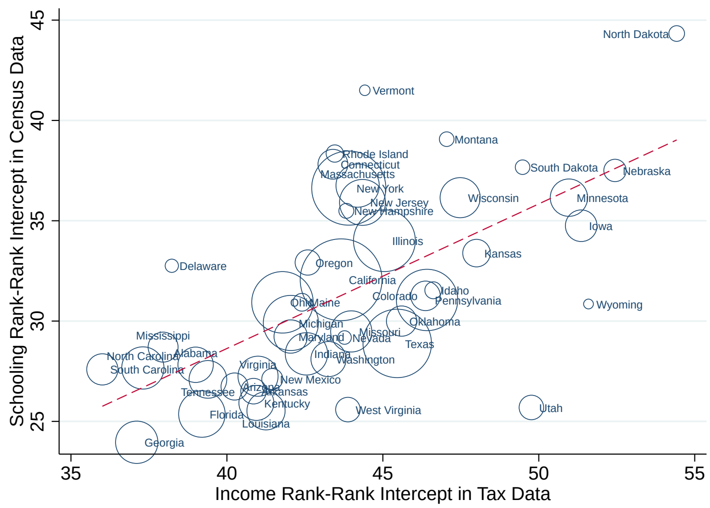
(a) Whites



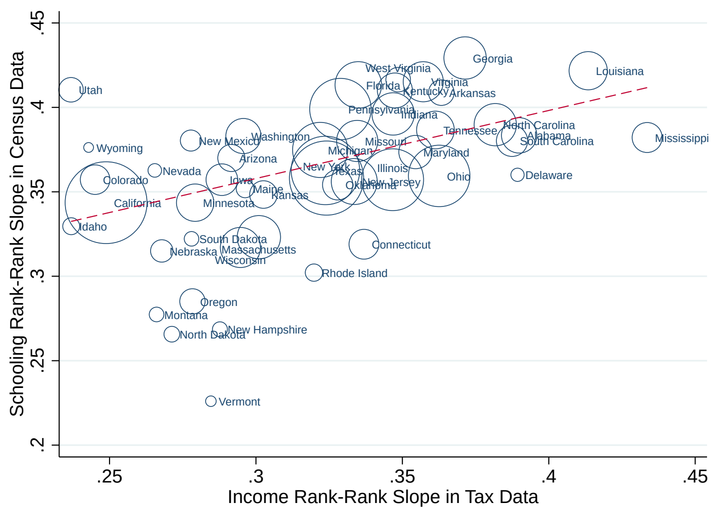
(b) Blacks

Figure VIII: Educational Mobility in Parental Income Decile, 1940-2000

Notes: Figure plots slopes from regression of child education on population parental income decile by year. Child education and parental income decile defined as described in text. Census estimates adjusted for independents as described in text. "Year" in census defined as year available cohorts turn ages 26-29. "Year" in PSID, NLSY79, and NLSY97 defined as decades, e.g. "1980" reflects cohorts of children turning 26-29 during the years 1980-89. All estimates make use of sample weights and exclude zero incomes.



(a) Absolute Upward Mobility



(b) Relative Mobility

Figure IX: Comparison of IM Estimates by State in Census and Tax Data

Notes: Panel (A) plots intercepts from regression of child education rank on parental education rank in census data with adjustment for independents, against intercepts from regression of child income rank on parental income rank in U.S. population tax records. Panel (B) plots slopes from the same regressions. All races pooled. Education measured as highest grade attained. Education ranks computed on national sample for each age and year separately with midpoints of rank intervals assigned to mass points. Children's education measured at ages 26-29. Points weighted by estimated total number of children age 26-29 in census. Census regressions pool data from 1980, 1990 and 2000 censuses. Description of income rank-rank mobility estimates available in Chetty et al. (2014a).

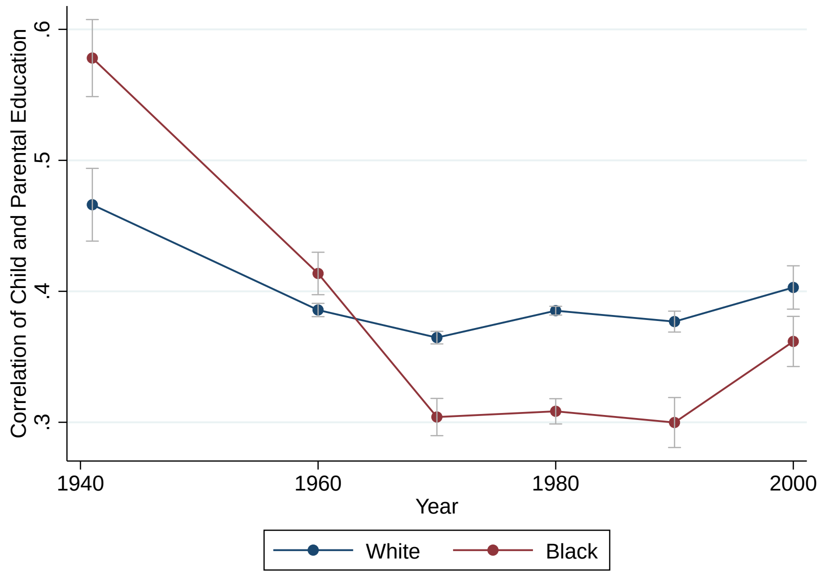


Figure A.1: Child-Parent Educational Correlations by Race, 1940-2000

Notes: Correlations constructed as elasticities multiplied by ratio of standard deviation of parental average education over standard deviation of child average education. Standard errors assume Moulton error structure.

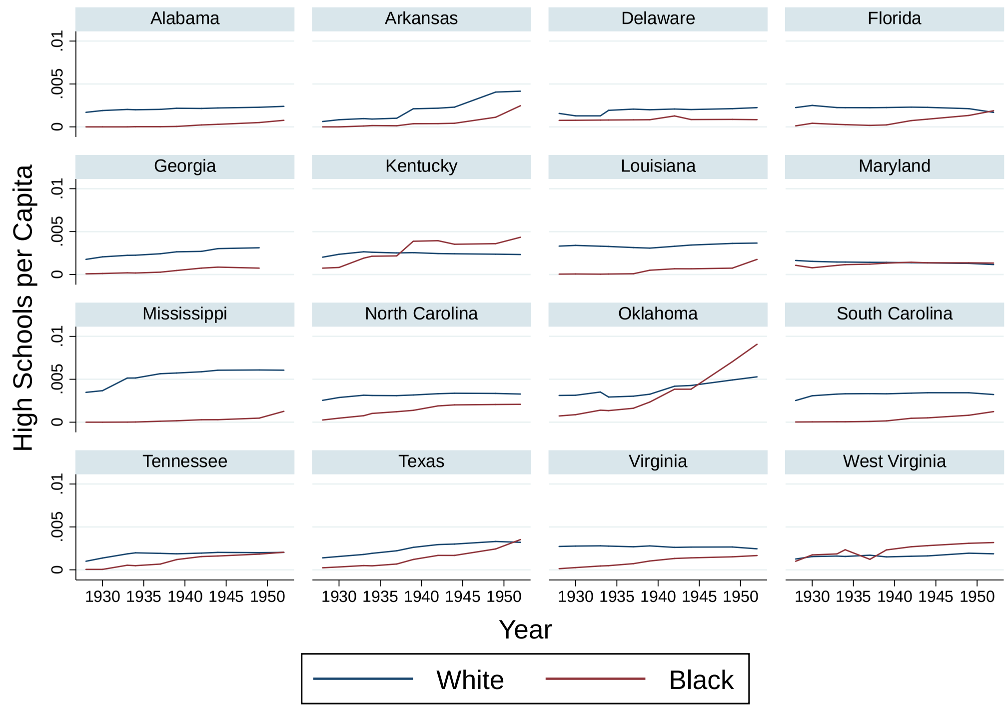


Figure A.2: Black and White High Schools per Capita in U.S. South by State, 1928-1952

Notes: Figure plots number of black and white high schools divided by number of black and white children, respectively, ages 14-17 in each state.

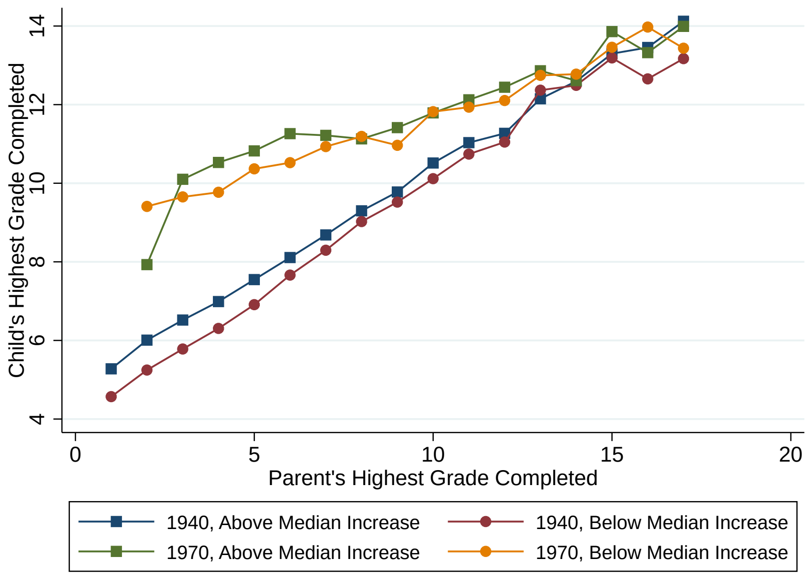
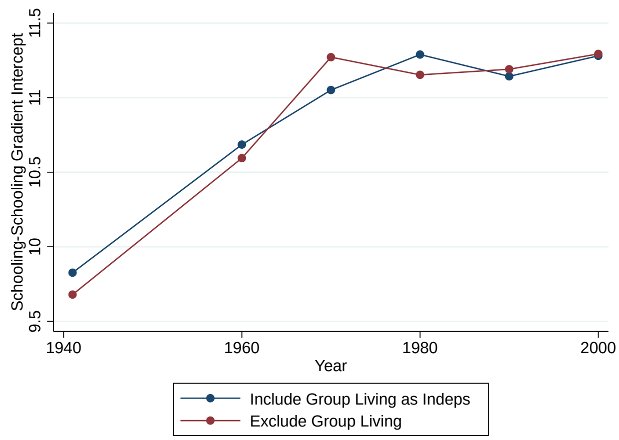
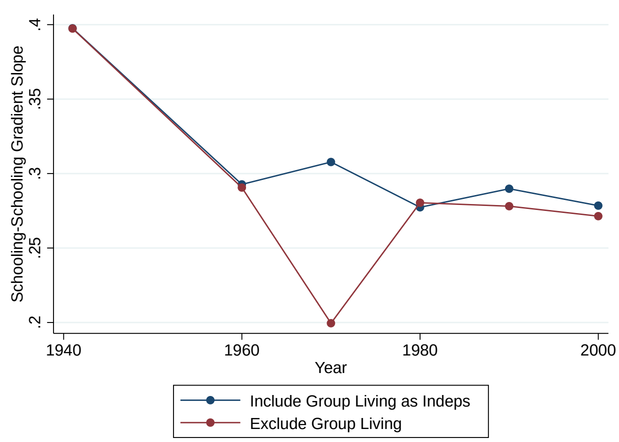


Figure A.3: Black Mobility in Southern States with Big vs. Small Increases in Black High School Density

Notes: Figure plots highest grade completed at age 20-29 in Southern states with above- and below-median increases in black public high schools per capita. High schools per capita measured in 1928 and 1952.



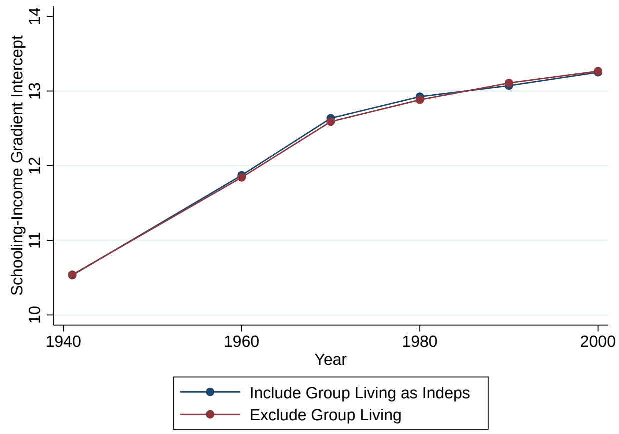
(a) Intercepts



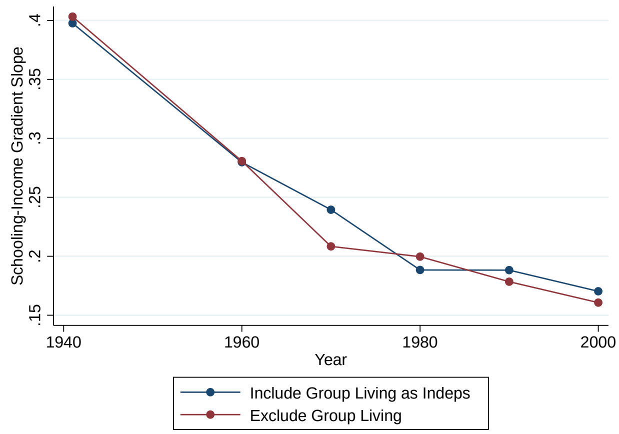
(b) Slopes

Figure A.4: Intercepts and Slopes of Schooling CEFs in Parental Education by Sample and Year

Notes: Figure documents that education CEFs in parental education are not sensitive to different ways of classifying children as independent, by year, for whites and blacks. Presents estimated intercepts and slopes from linear regressions of children's highest grade attained on parent's highest grade attained, using data grouped at the year by race by parental education level. Sample weights are used to construct cell means, and regressions on collapsed data are weighted by the square of cell size.



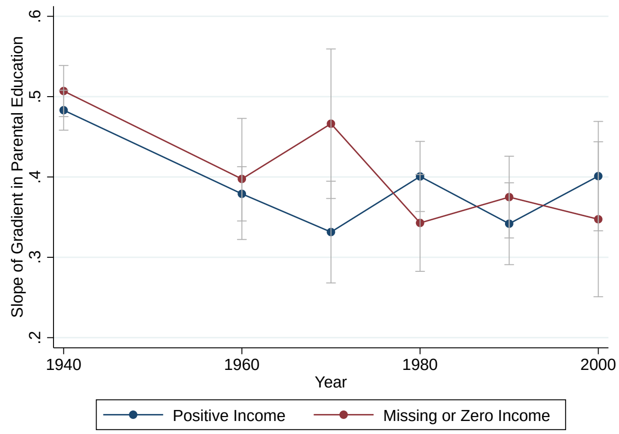
(a) Intercepts = Absolute Upward Mobility



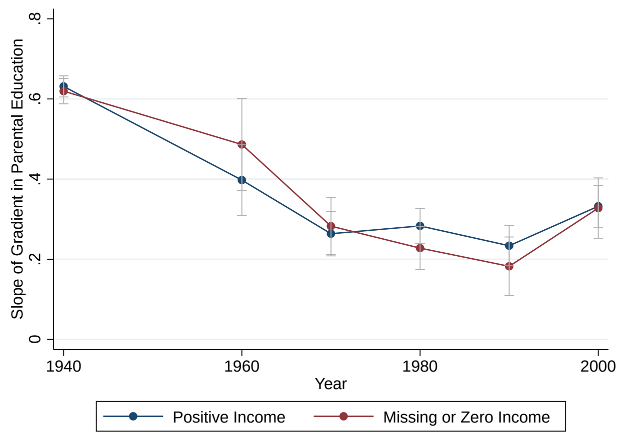
(b) Slopes = Relative Mobility

Figure A.5: Intercepts and Slopes of Schooling CEF in Parental Income Deciles by Sample and Year

Notes: Figure documents that education CEFs in parental income are not sensitive to different ways of classifying children as independent, by year, for whites and blacks. Presents estimated intercepts and slopes from linear regressions of children's highest grade attained on parental income decile, using data grouped at the year by race by parental income decile level. Sample weights are used to construct cell means, and regressions on collapsed data are weighted by the square of cell size.



(a) Whites



(b) Blacks

Figure A.6: Slopes of Schooling CEFs in Parental Education by Missing Income Status and Year

Notes: Figure documents that education elasticities are similar in families with positive household earnings and families with missing/zero household earnings, by year, for whites and blacks. Presents estimated slopes from linear regressions of children's highest grade attained on parent's highest grade attained, using data grouped at the year by race by parental education level. Sample weights are used to construct cell means, and regressions on collapsed data are weighted by the square of cell size.