Mobility for All: Representative Intergenerational Mobility Estimates over the Twentieth Century

Elisa Jácome

Northwestern University and National Bureau of Economic Research

Ilyana Kuziemko

Princeton University and National Bureau of Economic Research

Suresh Naidu

Columbia University and National Bureau of Economic Research

We estimate long-run trends in intergenerational relative mobility for representative samples of the US-born population. Harmonizing all surveys that include father's occupation and own family income, we develop a mobility measure that allows for the inclusion of nonwhite individuals and women for the 1910s–1970s birth cohorts. We show that mobility increases between the 1910s and 1940s cohorts and that the decline of Black-white income gaps explains about half of this rise. We also find that excluding Black Americans, particularly women, considerably overstates the level of mobility for twentieth-century birth cohorts while simultaneously understating its increase between the 1910s and 1940s.

We thank Diva Barisone, Madhavi Jha, Paola Villa-Paro, and especially Ahna Pearson for excellent research assistance. Sandy Black, Leah Boustan, Jaerim Choi, Jonathan Davis, Ellora Derenoncourt, James Feigenbaum, Nathan Hendren, Tom Hertz, Chi Hyun Kim, Trevon Logan, Robert Margo, Bhash Mazumder, Jake Mortenson, Chris Muller, Shu Shen, and Zachary Ward have provided invaluable data and feedback at various stages of

Electronically published November 14, 2024

Journal of Political Economy, volume 133, number 1, January 2025.

https://doi.org/10.1086/732527

I. Introduction

Intergenerational relative mobility—how tied an individual's place in the income distribution is to her parents' place in the income distribution while she was growing up—has long been an object of interest, especially in the United States. While analysis of modern data shows that the United States is less mobile than its rich peers (Jantti et al. 2006; Bratsberg et al. 2007), much less is known about trends in US mobility over the twentieth century.[1]

The main contribution of this paper is simple: we present, to the best of our knowledge, the first estimates of long-run intergenerational relative mobility trends for a representative sample of US-born individuals. In particular, we show mobility estimates for children born in the 1910s–1970s.[2] As table 1 shows, a handful of papers have made important contributions to our understanding of long-run trends in intergenerational relative mobility, typically relating occupational standing of one generation to the next (Occ.–Occ. mobility). However, for data reasons, they include only subsets (and typically advantaged subsets) of the population. Song et al. (2020) shows mobility of occupational prestige from 1830 to 1980 but only for white men. Using a clever synthetic panel strategy based on the status information conveyed by first names, Olivetti and Paserman (2015) can compare occupational mobility between fathers and sons with that of fathers and sons-in-law but only for white men and married white

this project. We are grateful to Magne Mogstad and six anonymous referees for editorial guidance. We thank seminar and conference participants at Allied Social Science Associations, University of California, Berkeley, the Bissell-Heyd Symposium at the University of Toronto, University of Cologne, Georgetown University, University of Gothenburg, the Inter-American Development Bank, Iowa State University, McMaster University, University of Michigan, National Bureau of Economic Research Development of the American Economy Working Group and Economics of Mobility Conference, Northwestern University, University of Notre Dame, New York University (NYU) Wagner Graduate School of Public Service, NYU Stern School of Business, Opportunity Insights, the Research Network on Intergenerational Mobility, Princeton University, Stanford University, Stockholm Swedish Institute for Social Research, University of Toronto, University of Colorado Denver, University of California, Merced, University College London, University of Southern California, University of Illinois Urbana-Champaign, University of Warwick, Wharton School of the University of Pennsylvania, and Yale University. This paper was edited by Magne Mogstad.

[1] Song et al. (2020, 251) write that "evidence of long-term trends in intergenerational mobility is largely absent." Similarly, Mazumder (2018, 225–26) writes, "One active topic of research that has not yet been resolved is whether there have been major changes in intergenerational mobility in the United States over time."

[2] Note that we do not examine intergenerational absolute mobility, which captures the probability that a child's income as an adult surpasses her parents' income (in real US dollars) while she was a child. Specifically, we do not have detailed income measures for the parents' generation (and instead rely on measures of predicted income, detailed in sec. III), so we cannot accurately calculate the share of adult children earning more than their parents. For recent work on intergenerational absolute mobility, see Chetty et al. (2017), Berman (2018, 2022), and Manduca et al. (2024).

TABLE 1
SELECT REVIEW OF INTERGENERATIONAL MOBILITY PAPERS USING US DATA

		Income/Status Proxy			
Paper	Cohorts (1)	Parent(s) (2)	Child (3)	Links (4)	Sample (5)
Ward 2023	1850–1910	Occupation × race × region	Occupation × race × region	Match	All males
Collins and Wanamaker 2022	1880–1970	Occupation × race × region	Occupation × race × region	Match and retrospective	All males
Song et al. 2020	1830–1980	Occupation	Occupation	Match and retrospective	White males
Abramitzky et al. 2021	1860–1900	Occupation × region	Occupation × region	Match	White males
Long and Ferrie 2013	1840, 1930	Occupation	Occupation	Match and retrospective	White males
Olivetti and Paserman 2015	1840–1910	Occupation	Occupation	Synthetic panel	White males, married females
Feigenbaum 2018	1900	Income	Income	Match	Iowa males
Feigenbaum 2015	1900–1910	Income	Income	Match	Urban males
Card et al. 2018	1920	Education	Education	Same household	Representative
Bowles 1972	1930	Income	Income	Retrospective	Current Population Survey males
Mazumder 2015	1950–70	Income	Income	Panel data	Representative
Davis and Mazumder 2022	1950–60	Income	Income	Panel data	Representative
Chetty et al. 2014a	1980–91	Income	Income	Claim dependent	Representative
Chetty et al. 2020	1978–83	Income	Income	Claim dependent	Representative
This paper	1910–70	Occupation × race × South	Income	Retrospective	Representative

NOTE.—Since many papers do not explicitly consider birth cohorts, col. 1 refers to the birth decade(s) that most of the sample comes from, given the age restrictions used in the paper. In col. 4, “Match” refers to matching across datasets (e.g., census matching by name, age, and state of birth); “Synthetic panel” refers to matching based on characteristics but not individual identity; “Claim dependent” refers to matching by whether the parent ever claims the child as a dependent to the IRS; “Retrospective” refers to adult children being asked retrospectively about the characteristics of their parents (e.g., occupation and education). Abramitzky et al. (2021) also considers modern cohorts corresponding to those in Chetty et al. (2020).

women. Collins and Wanamaker (2022) and Ward (2023) include Black Americans but only men.[3]

We begin by locating (to the best of our knowledge) all surveys that ask individuals their current family income as well as their race, father's occupation, and region of birth or childhood. Instead of relying on the traditional Occ.–Occ. mobility measure (which complicates looking at women, as few formally worked after marriage in the historical period), we relate the family income reported by prime-age adults in these surveys with their predicted family income during childhood.[4] We directly observe contemporaneous family income of the adult child, as it is a question asked in many surveys and—unlike own occupation—can be answered by male as well as female respondents. Moreover, unlike occupation alone, self-reported family income naturally reflects income gaps by race or other characteristics. Similar to a two-sample instrumental variable approach, we predict a respondent's childhood income by using their race, region, and father's occupation and calculating the average family income conditional on these characteristics among households with children in the census or other auxiliary data sources (from as close as possible to the respondent's tenth birthday).[5]

Our main finding is that both the intergenerational elasticity (IGE) and rank-rank correlation fell (meaning that mobility rose) between the 1910s and 1940s birth cohorts. The IGE (rank-rank correlation) falls from 0.75 (0.37) for those born in the 1910s to 0.42 (0.25) for those born in the 1940s. Between the 1940s and 1970s birth cohorts, the IGE measure drifts upward again, while there is little change in the rank-rank correlation. Both because the trends after the 1950s are more sensitive to the measure (IGE or rank-rank) and because alternative data sources and methodologies are available for these more modern birth cohorts, we mainly focus on the 1910s–1940s cohorts.

Importantly, we do not claim to have estimated causal effects of childhood income on adult income, which would require us to identify sources of exogenous variation in parental income in each of our birth decades. We view our results as descriptive. Moreover, given the variety

[3] As table 1 shows, many important papers do include representative samples but for either more modern birth cohorts only (Solon 1992; Chetty et al. 2014b, 2020; Mazumder 2018) or short snapshots of time (Card, Domnisoru, and Taylor 2018; Massey and Rothbaum 2020).

[4] Note from table 1 that Collins and Wanamaker (2022) and Ward (2023) both take an income score to income score approach (in essence, predicting income using information on occupation, race, and region for both generations). This approach remains problematic for women, however, for the same reason as Occ.–Occ. measures: a woman's own occupation is both endogenous to marriage and not highly predictive of economic well-being in the historical period.

[5] Björklund and Jäntti (1997), Bloise, Brunori, and Piraino (2021), and others use a two-sample instrumental variable approach to estimate intergenerational income mobility, but this approach has been less common in the US context given limited historical microdata.

of approaches authors have taken to overcome various data limitations inherent in estimating historical mobility, we do not claim to have estimated the IGE or rank-rank from this period (e.g., comparing the magnitudes of our results with those of an Occ.–Occ. estimation should be done with caution). Rather, the main goal of our various robustness checks is to show that the biases of our approach relative to the ideal (and infeasible) ordinary least squares (OLS) regression of child on parental income are not changing over time in a manner that would produce our result as an artifact.

To the best of our knowledge, the significant increase in US intergenerational relative mobility from the 1910s to the 1940s birth cohorts that we find is novel in the literature. Much of the existing historical literature emphasizes either rising persistence or stability (e.g., Olivetti and Paserman 2015; Song et al. 2020; Davis and Mazumder 2022), especially in the US context. Our uncovering a period of significantly rising mobility is due to two factors. First, the 1910s–1940s birth cohorts have been understudied, as they are born too recently for census-based linking but also too long ago for study using data from the Internal Revenue Service (IRS) or Panel Study of Income Dynamics (PSID).[6] Second, we show that the more traditional statistic—the Occ.–Occ. mobility estimate for white men—misses much of the rise in mobility for these cohorts. Indeed, we show that using additional inputs beyond father's occupation improves the prediction of childhood income and offers additional insights on the evolution of mobility than the traditional Occ.–Occ. mobility measure. This period of rising mobility in the United States complements recent studies (Pekkarinen, Salvanes, and Sarvimäki 2017; Karlson and Landersø 2021; Nybom and Stuhler 2023) documenting increases in mobility in Denmark, Sweden, and Norway following educational reforms in the mid-twentieth century.

In the second part of the paper, we focus on subgroups (mostly the four subgroups defined by Black/white race and male/female sex) and, in particular, how movements of these subgroups contributed to (or slowed) the increase in mobility from the 1910s to 1940s cohorts. Like inequality measures and unlike means, the full-population IGE (or rank-rank) slope is not a weighted average of subgroup slopes. In particular, in societies with two very unequal and endogamous subgroups (a description that applies to white and Black Americans over much of the twentieth century), between-group differences in mean incomes play a major role in determining overall relative mobility for any given birth cohort, and changes in those mean differences play an important role in determining trends in overall mobility.

[6] As of today, the most recent census that can be linked is that of 1940, when the 1910s birth cohorts would be only in their twenties and the 1940s cohorts not yet born.

We show that between the 1910s–1920s and the 1940s–1950s birth cohorts, Black Americans exhibit significant (though still partial) convergence to white individuals in both (predicted) childhood income and adult income. White individuals also enjoy income growth in real terms (though slower than Black individuals), and their IGE and rank-rank slopes become flatter (meaning that within the white population, parental income matters less in predicting own adult family income). Our decomposition, applying Hertz (2008) to our historical data, shows that the Black-white convergence in mean income accounts for half of the rise in overall mobility (and the flattening of the white slope accounts for the remainder). This result is quite striking given that Black Americans are a relatively small share of the US population—roughly 12% for much of our sample period. But because they are drawn from an extremely low part of the childhood and adult income distribution and in our historical period did not intermarry with white individuals, changes in their average income exert great statistical influence on the overall regression line.

In this paper, we pay particular attention to Black women. First, because of data limitations, there has been almost no work on historical intergenerational mobility that includes this group. Second, we show that because Black women tend to grow up in the bottom of the income distribution (as do their male counterparts) and in our historical period are the lowest-income group as adults (substantially poorer than Black men), they play an especially outsize role in increasing full-population intergenerational persistence measures. As just one example, in 1920, the IGE increases from 0.51 to 0.59 (with nonoverlapping confidence intervals) when Black women (only 6% of the population) are added to the rest of the sample. Excluding even this small share of the population overstates early twentieth-century US mobility considerably. At the same time, we show that excluding Black Americans or even only Black women significantly reduces the rise in mobility in the first half of the twentieth century and, as we note, is one reason that past studies of this period focusing on white men have not found large declines in persistence.

We close the paper with a brief analysis of what role modern racial income gaps play in explaining low levels of US mobility relative to rich peers. Decomposition exercises show that modern levels of racial inequality set a very high lower bar on US intergenerational persistence: in order for the United States to attain an IGE of 0.20 (roughly that in Denmark) while maintaining current racial differences in income, US within-race IGEs would have to fall below 0.05, a remarkably high and likely unattainable level of within-race mobility.

The remainder of the paper is organized as follows. In section II, we describe the various datasets we use. In section III, we describe our methodology, in particular, the adult-family-income-to-predicted-childhood-income mobility measure (Self-reported inc.–Predicted inc.). Section IV presents our results for the full representative population, and section V probes the robustness of these results. Section VI presents a decomposition of the full-population mobility measures and then decomposes the rise in mobility into differential mobility by race and gender. Section VII concludes.

II. Data

In this section, we briefly describe the datasets that we use in this paper and share summary statistics. More details can be found in appendix E (apps. A–E are available online).

A. Datasets and Sampling Rules

We have located, to the best of our knowledge, all surveys that ask respondents their current family income, their fathers' occupation while they were growing up (with sufficient detail), their race, and the region of the country where the respondent was born or grew up (at least to the level of the South vs. other regions). We end up locating 15 different surveys, with details on all of them provided in appendix E. Most readers will be familiar with some (e.g., the General Social Survey or the American National Election Survey), but others are not as well known (e.g., the National Survey of Black Americans or Americans View Their Mental Health).[7]

We restrict attention to US-born men and women ages 30–50 in order to ensure that we are measuring life cycle earnings as closely as possible.[8] Because advantaged children spend on average more time in formal education, their earnings tend to be disproportionately depressed in the late twenties relative to their prime-age earnings, so measuring the adult child's income at these ages may lead to downward bias of persistence

[7] In some cases, the data we use are in fact panel datasets that follow individuals and families over time (e.g., the PSID and the National Longitudinal Surveys of Mature Women and Older Men) and have often been used to measure mobility for more modern periods. To remain consistent within our methodology, however, we do not use the panel components of these datasets. In the first wave, these panel datasets often ask the adult respondent questions about their own childhood, and it is this linkage that we use to predict the respondent's family income in childhood.

[8] We restrict the sample to US-born men and women because we want to ensure that our measures of childhood income—which are derived from US sources—are relatively accurate approximations of income in the parental generation. The share of adult children who are excluded because of this restriction is relatively small: the share of adults ages 30–50 who were born outside of the United States ranges from 5% to 9% in the 1950–80 censuses, which correspond to our time period of interest. We do note, however, that first-generation immigrant parents (a sizable group in this time period) would be included in the analysis as long as their children were born in the United States.

measures. Haider and Solon (2006) suggest as a rule of thumb to observe adult children as close to age 40 as possible.

All of the surveys used in this analysis ask respondents about their total family income. Many of the surveys ask respondents to report their income by choosing an interval (e.g., \$8,000–\$10,000), whereas others allow respondents to provide an exact value. To be consistent across surveys and over time, we transform the variables in the latter group to resemble those in the former group, so that our baseline measure of an adult child's family income is a categorical variable with a similar number of income bins over time. For more details on the construction of this harmonized variable, see appendix E.

Our baseline sample spans the 1910s–1970s birth cohorts and consists of respondents with nonmissing family income and with available information on race, childhood location, and father's occupation (used to predict parental income, as described in sec. III). In the earliest cohorts in our sample, the share of children living without fathers is very small. Later in the paper, we present various robustness checks to assess sensitivity of the more modern results to various assumptions about missing fathers (i.e., including nonworking or retired fathers or using information about the mother's occupation).

In many cases, the data collection for these surveys was explicitly meant to be representative and provides survey weights to correct deviations due to sampling error. In those cases, we use the provided sampling weights. Of course, some of these surveys target one sex (e.g., the National Fertility Survey) or one race (e.g., the National Survey of Black Americans) and so are clearly not representative of the full US-born population. In the early cohorts, we also have a substantially lower share of women in our data relative to the general population. For this reason, we will always reweight the pooled dataset so that each cohort has weighted shares for white women, white men, Black women, and Black men that match the corresponding shares in the census (Ruggles et al. 2021).[9] In appendix B, we show that our main results barely change under other weighting schemes, including not weighting at all.

B. Summary Statistics

Panel A of table 2 shows summary statistics of the fathers of the respondents in our main dataset, separately by decade of birth. In this table, we do not weight at all so that readers can get a sense of the raw data.

[9] We focus on only individuals whose race is classified as white or Black. Individuals of other races account for tiny shares of the surveys' samples in these historical time periods (1% or less of the sample in the pre-1950 cohorts). The decomposition in sec. VI also highlights that groups with very small population shares are unlikely to affect the full-population measures of persistence.

TABLE 2
Summary Statistics by Birth Decade

	1910s	1920s	1930s	1940s	1950s	1960s	1970s
	A. Father Demographics						
Foreign born	.22	.17	.11	.05	.04	.03	.05
High school educated	.17	.19	.26	.45	.60	.70	.81
College educated	.04	.04	.06	.09	.16	.20	.26
Farming occupation	.37	.29	.24	.15	.09	.05	.03
	B. Respondent Demographics						
Female	.12	.33	.45	.44	.57	.52	.56
Age	45.89	41.52	36.95	38.50	38.05	38.45	38.92
Black	.12	.13	.15	.15	.18	.16	.23
High school educated	.50	.61	.71	.85	.90	.91	.91
College educated	.10	.14	.16	.28	.28	.30	.39
Moved regions	.21	.21	.22	.24	.22	.21	.22
Union member (males)	.31	.31	.29	.28	.22	.17	.12
Veteran (males)	. . .	.77	.59	.46	.21	.14	.14
	C. Parental Income						
Predicted income (1950 US$)	2,340	2,575	3,292	5,373	7,686	9,104	9,482
Missing income	.13	.15	.16	.15	.16	.24	.21
Rank	45.37	45.44	45.50	45.93	45.61	46.83	44.91
	D. Respondent Income						
Family income (1950 US$)	5,506	6,803	7,292	7,900	7,617	7,878	8,521
Missing income	.15	.10	.06	.06	.08	.08	.04
Bottom coded	.08	.05	.03	.02	.04	.05	.06
Top coded	.07	.07	.08	.12	.12	.09	.08
Family income rank	49.00	48.23	47.09	46.31	46.38	47.71	45.95
Observations	5,207	13,328	12,446	11,589	10,951	6,611	3,119

Note.—The table combines 15 different surveys, which are described in sec. II and in further detail in app. E. All of the shares in this table are unweighted and are based on the baseline sample of respondents ages 30–50 (i.e., those with nonmissing family income and predicted parental family income). The two exceptions are the rows for missing income, which consider all US-born respondents ages 30–50 in the 15 surveys. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). For characteristics that are unavailable in every survey (e.g., father's educational attainment), the average is computed using only the baseline sample respondents in the surveys with the available information. When considering union membership and veteran status, we restrict the sample to male respondents. "Bottom coded" and "Top coded" refer to the share of individuals who had family income values in the bottom or top bins, respectively.

The decline of agriculture as a dominant occupation for fathers is readily apparent for children in the 1910s–1950s birth cohorts, falling from over one-third to less than one-tenth. We do not have father's education in every survey, but the table shares summary statistics from those surveys that do include father's education. In our earliest birth cohorts, the fathers in our data are born in the last few decades of the nineteenth century and thus grew up before the high school movement, which is reflected in their low levels of secondary education. Less than 20% of the fathers of our 1910s and 1920s birth cohorts graduated from high school. College graduation was a rarity for these fathers, and as late as the 1950s birth cohort, less than one in six of respondents have fathers who completed college.

Summary statistics for the adult children (i.e., the survey respondents) appear in panel B of table 2. The age of respondents is relatively similar and always close to 40, as we would expect from our 30–50 age restriction. In contrast to past historical work on US mobility—which either excludes nonwhite individuals or uses linkage techniques that significantly undersample nonwhite individuals—our samples have coverage of Black individuals very close to their population shares even before weighting.

A number of well-known trends among the children are apparent in our data. The rise in educational attainment from the 1910s to the 1950s birth cohorts is striking and consistent with Goldin and Katz (2010). High school attainment increases from one-half to 90%, and college graduation rates nearly triple from 10% to 28%. The increase in education from one generation to the next is massive as well: for the 1910s–1930s birth cohorts, the likelihood our survey respondents graduate from high school is triple that of their fathers.[10]

Table A.1 (tables A.1–E.7 are available online) separates our data (unweighted, as in table 2) by time period, race, and sex and compares survey respondents with the relevant population in the census. As before, we see that in all periods and separately for men and women, our data are very close to representative on race (roughly 10%–15% of the sample). In fact, one of the only variables in which there are small discrepancies between our raw survey data and the census data is education in the earliest birth cohorts (we later show robustness to using weights that adjust for these differences). Otherwise, our raw survey data are remarkably similar to the census in terms of age, the share living or originating from the South (an especially important variable for Black respondents), and marital patterns.

III. Methodology

With ideal data, we would regress log permanent household income of the adult child on log permanent income of her household while she

[10] Another marked trend for the adult children in our data is the decline in veteran status (which the table reports only for men in surveys that asked about veteran status). While over 70% of men in our 1920s cohort report military service, by the 1950s, cohort military service has become relatively rare. Finally, another noticeable trend is union membership: while it holds steady for our early cohorts, it begins a steady decline with the 1950s cohort, consistent with Farber et al. (2021).

was growing up. As is well understood in the US historical mobility literature, such a regression is not feasible, so sections III.A–III.E describe the approach we follow instead.

A. Specifications

With ideal data, we would estimate changes in intergenerational mobility over time using the classic log-log specification (Becker and Tomes 1979):

$$\log(y_{ic}) = \beta_c^{\mathrm{OLS}} \log(y_{ic}^{\mathrm{p}}) + \epsilon_{ic}, \tag{1}$$

where y_{ic} is the permanent household income of respondent i born in cohort c, y_{ic}^{p} is permanent family income of respondent i's parents, and ϵ_{ic} is the error term. Here, the coefficient β_c^{OLS} is the IGE, and it is a descriptive coefficient that does not take on causal interpretation. Estimating this equation separately by birth cohort would allow us to see how β_c^{OLS} changes across cohorts c.

Because our surveys do not include information about parental income, estimating this ideal β^{OLS} is not feasible. The preferred approach is thus a two-sample two-stage least squares (TS2SLS; Inoue and Solon 2010) estimation, using auxiliary data sources as well as information about the respondents' upbringing to predict their log parental income. This empirical strategy has been a common approach in the intergenerational mobility literature (see, e.g., Björklund and Jäntti 1997; Aaronson and Mazumder 2008; Olivetti and Paserman 2015; Bloise, Brunori, and Piraino 2021). Because the 15 surveys in our baseline sample include information about the respondent's race, childhood location, and father's occupation, we can use these variables in auxiliary datasets (e.g., US census microdata) to predict the log income of individuals with those same characteristics.

Our surveys include respondents born between the 1910s and 1970s birth cohorts, so implementing a TS2SLS strategy requires microdata that span this time period. Nevertheless, as is well known in the US economic history literature, there are limited sources of microdata that include income measures prior to the 1940 census, especially for representative samples of the population. Given that we do not have microdata to predict parental income for the 1910s birth cohort, we instead implement a modified TS2SLS strategy. In section V, we present numerous robustness checks, including standard TS2SLS estimates.

We use available sources of microdata as well as historical records of income from the early twentieth century (described in sec. III.B) to calculate average income conditional on occupation, race, and location. We then apply a log transformation and use these imputed measures of log parental income as our right-hand-side variable.[11] This imputation approach is frequently used in US economic history papers, including Collins and Wanamaker (2022) and Ward (2023). Our baseline specification is thus

$$\log(y_{ic}) = \beta_c \widetilde{\log(y^{\mathrm{p}}_{ic})} + \epsilon_{ic}. \tag{2}$$

In appendix D, we more formally compare our estimated $\hat{\beta}_c$ and the ideal β_c^{OLS}, but in this section and in section III.B, we focus on simply describing our estimation procedure.

As an alternative to the IGE, we follow the rank-rank approach in Chetty et al. (2014a).[12] The rank of the adult child, Rank_{ic}, is the rank of predicted family income among all adult children born in the same year. Similarly, the rank of the parents, $\mathrm{Rank}^{\mathrm{p}}_{ic}$, is the percentile (based on predicted parental income) among all parents having a child born in the same year. The mapping of child's rank to parental rank (the copula of the joint distribution) tends to be linear and can handle zeros, which may be missed in the (logarithmic) IGE specification. Chetty et al. (2014a) focus on this specification:

$$\mathrm{Rank}_{ic} = \gamma_c \widetilde{\mathrm{Rank}}^{\mathrm{p}}_{ic} + \delta_{sy} + \epsilon_{ic}. \tag{3}$$

In this estimation, γ_c is an estimate of the rank-rank correlation for cohort c. Again, we estimate this equation separately by birth decade.[13]

While the rank-rank measure has become a fixture of the intergenerational mobility literature, it is unwieldy for decomposing changes in the full-population mobility measure into subgroup-specific changes, as changing a subgroup's mobility will affect the ranks of the whole population. As decompositions of changes in mobility along the lines of race and gender is a key focus of our paper, we also show results for the intergenerational correlation (IGC), which is the same as the IGE but standardizes the log income of children and parents by the mean and standard deviation. The IGC thus measures a positional mobility concept, similar to the rank-rank

[11] The difference between the TS2SLS and imputation approaches is the order of prediction vs. log transformation of parental income. TS2SLS predicts average log parental income using microdata in the first stage of the estimation, whereas the imputation approach calculates average income for each cell and then applies the log transformation. In sec. V.C, we present results from a levels-based specification that avoids these issues as well as TS2SLS estimates of our main specifications, though we cannot extend the analysis as far back in time with this approach.

[12] Discussions of the relative merits of different measures of mobility can be found in Fields and Ok (1999), Deutscher and Mazumder (2023), and Ray and Genicot (2023).

[13] Note that unlike the log specification, one cannot implement a TS2SLS procedure with ranked income. It is, of course, possible to construct predicted rank income for fathers in auxiliary data (i.e., estimate a first stage). However, computing the average rank for each cell implies compressing the rank distribution, so that in the second stage, the distribution of ranked parental income is no longer uniform and its variance affects the level of the rank-rank correlation.

correlation, while retaining the tractability of the IGE specification. Appendix D discusses the relationship between these three concepts as well as a specification that utilizes levels of income for both generations.[14]

B. *Predicting Parental Income*

IPUMS provides 1950-based occupational income scores, which calculate the median total income of people (pooling men and women) in each occupation in the 1950 census. These income scores have been used to approximate the income of individuals in earlier (or later) censuses who have the same occupations. Our approach to constructing income predictions is similar in spirit to that of the IPUMS occscore variable, but we differ in four notable ways.

First, not all of our surveys have father's occupation categories that are as detailed as those in the census. Across all of our surveys, we can harmonize occupations into 28 categories. We thus build and use crosswalks that map the occupations in our surveys into these 28 categories. These coarsened bins include broad occupations, like doctors, clerical workers, craftsmen, and farm laborers; the full list is in appendix E.

Second, when constructing measures of predicted parental income, we limit the samples whenever possible to men ages 30–50 who are living with a biological child younger than age 18 (these men are almost always living with a wife as well). This sample restriction should better proxy household income of fathers with a given occupation, which is the population of interest when we try to predict income during the respondent's childhood.

Third, we calculate the average household income (summing across all working adults in the household) by father's occupation, race (Black vs. white), and region (South vs. elsewhere). We follow recent papers (Saavedra and Twinam 2020; Abramitzky et al. 2021; Collins and Wanamaker 2022; Ward 2023) that utilize characteristics beyond occupation to improve measures of predicted income. Given widespread discrimination and occupational segregation, using occupational scores computed from pooled Black and white populations will substantially mismeasure childhood incomes. Similarly, the South is far poorer than other regions during our sample period, so pooling across all regions throws out valuable information, especially for Black respondents who are vastly overrepresented in the region. The choice of South versus elsewhere for the construction of predicted income is motivated by the fact that this level of detail is present in every survey. However, in section V.A, we check the

[14] We show that TS2SLS and OLS imputation-based estimates are numerically equivalent to a levels-on-levels specification and that one can transform the β^{levels} coefficient using the income distributions to approximate the IGE and rank-rank correlation.

robustness of the main results to using the four census regions as predictors of income for the subsample of respondents for whom we have this level of information.

Finally, instead of relying on only the 1950 census, we use multiple datasets spanning the twentieth century to approximate parental income on the basis of when the survey respondent was growing up. Specifically, we use income information from the 1901 Cost of Living Survey, the full-count sample of the 1940 census, as well as the 1960–90 censuses (Ruggles et al. 2021).[15] We combine our data sources so that families are assigned measures of predicted income that come from the data sources closest in time to when the respondent is 10 years old. That is, the 1910s–1920s cohorts are assigned measures of predicted income that are weighted averages of the 1901- and 1940-based predictions, the 1930s–1940s cohorts are assigned measures that are weighted averages of the 1940- and 1960-based predictions, and the 1950s–1970s birth cohorts are similarly assigned income predictions that are weighted averages of measures constructed using the 1960–90 censuses.

One feature of historical measurement of occupational incomes is that farm income is notoriously difficult to impute, as it is both highly volatile (being subject to weather and price shocks) as well as difficult to measure (as comprehensive measurement of agricultural costs is difficult to capture). More than other occupations, farmers have also declined in relative status over the first half of the twentieth century; using data from Iowa, Feigenbaum (2018) shows that farmer families have median household income in 1915 but are at the 10th percentile by 1950, so their status in one decade cannot proxy for their status earlier or later.

For our earliest cohorts, we follow the approach in Goldenweiser (1916) and Abramitzky, Boustan, and Eriksson (2012) and use the 1900 census of agriculture to calculate farmers' net earnings. In our calculations, we allow for variation at the race × South level and take into account the share of individuals in each group who are not farm owners (i.e., part owners, or cash or share tenants). Moreover, because the 1940 census income variable excludes income from self-employment, which includes most farmers, we supplement the 1940-based predictions with the 1936 Expenditure Survey of the Bureau of Labor Statistics, which includes family income for farmers and the self-employed. To our knowledge, this data source is the earliest microdata to include total family income for these categories. More detail on this additional data source for farmers and the self-employed is available in appendix E. One reassuring comparison is that in our data, white

[15] Similar to Collins and Wanamaker (2022), we do not use the 1950 census to construct measures of predicted income, as only sample line respondents are asked about their income. For our purposes, this smaller sample size means that we are unable to calculate the average income for 11 occupation × race × South cells pertaining to Black fathers. Moreover, the sample line restriction makes it impossible to calculate average household income.

respondents born in 1910–29 outside of the South to farmer fathers are estimated to be growing up around the 37th–47th percentiles of the childhood income distribution, consistent with results in Feigenbaum (2018) for Iowa.

While we show robustness to many modifications of this prediction methodology in section V, the measure described in this section serves as our baseline approach for predicting childhood income, as we can calculate it for the respondents in all 15 of our surveys. Panel C of table 2 displays summary statistics related to predicted parental income.

C. Assessing the Accuracy of Predicted Parental Income

There are at least two challenges in predicting parental income using our methodology. First, adult children may not accurately recall their father's occupation. Second, even if recall is perfect, the way in which we assign parental income to survey respondents—on the basis of occupation, race, and region cells in auxiliary datasets—may not be reliable for predicting income (or the predictive power of the cells may change differentially over time).

Appendix C provides greater detail on the accuracy of adult children's recall, but we summarize some key results here. First, we show that predicted childhood incomes calculated for male and female respondents are indistinguishable (as we would expect, given that there is no documented evidence of sex selection in the United States in our historical period and thus boys and girls on average grow up in the same families) and show no differential trends over time. This equivalence by sex across decades holds for the full sample as well as for white and Black respondents separately (figs. C.1, C.2; figs. A.1–E.1 are available online). Second, figure C.3 shows that in surveys like the National Longitudinal Surveys (NLS), where multiple siblings from the same household are sampled, siblings' recall of their father's occupation is extremely highly correlated.

Third, we can perform a direct evaluation in modern data, as the PSID (beginning in 1997) asks household heads to recall their father's occupation, and in many cases, we directly observe the fathers of these respondents in earlier waves of the surveys (i.e., the 1960s and 1970s) when they are asked to report their own occupations. Over 80% of these household heads report an occupation that the father also reports, and the most common mistakes are small and understandable (e.g., one party reporting "craftsmen," and the other reporting "operatives"). Indeed, the correlation between logged predicted income based on father's self-reported occupation and those based on the child's recall is 0.83, and the relationship is very linear across the entire support of father's predicted income (so we do not see, e.g., that the children of the lowest-status dads tend to overstate their father's occupational status or that children of the highest-status dads tend to understate it; see fig. C.5). Further, the coefficient from a regression of the 5-year average of log father's income on our income prediction (using the retrospective question) results in a coefficient very close to 1, suggesting that our retrospective measures of predicted income are quite closely correlated with father's actual permanent income. While these data pertain to birth cohorts more modern than our years of interest, it is nonetheless reassuring that recall appears highly reliable.

Finally, if our surveys of adult children are representative and the respondents' recall is accurate, then the fathers described by our respondents and the fathers in the census when these respondents were growing up should be drawn from the same underlying population and thus appear similar on observables. Indeed, we show that the average (predicted) parental income as well as the types of occupations reported by our survey respondents are similar to the occupations of actual fathers in the census (see fig. C.4; tables C.2, C.3). These comparisons help alleviate concerns that children tend to inflate the status of their father when they are asked to recall their upbringing and that the implied distribution of parental income in the surveys will not correspond to the distribution of parental income in the same time period. Importantly, these exercises suggest that respondents' recall was not improving or deteriorating over time in a way that would drive the mobility trends we uncover.

As noted, even if recall is perfect, the predictions may be so noisy as to convey little information. Another concern—especially for earlier cohorts for whom we cannot use the census to predict childhood income—is that the auxiliary datasets are not representative. As a check on these concerns, we show that our predicted childhood income tracks known trends in overall inequality over the twentieth century. Figure A.1 shows the Gini coefficient as well as the ratio of top 10 to bottom 50 based on our predicted childhood income measure, separately by respondents' birth decades, as well as the analogous statistics of national income throughout this time period from the World Inequality Database (WID). Recall that comparisons of levels are not helpful, since by construction our measures will miss all household inequality arising from within father occupation × race × South cell variation. But for both the Gini and the ratio of top 10 to bottom 50, inequality measures based on our predicted childhood income and those based on the WID data track each other remarkably well. Figure A.1*c* plots the Black-white income gap for the past 150 years (based on series compiled by Margo 2016) and shows that our prediction of childhood income captures Black-white convergence at midcentury.

D. *Comparing Our Approach with an Ideal OLS Coefficient*

The evidence in section III.C reassures us that the fathers in the auxiliary datasets and the fathers of our survey respondents are drawn from the same or very similar underlying populations. However, even under this assumption, our imputation approach—which is akin to a TS2SLS estimation—produces known biases relative to the target OLS coefficient. In appendix section D.1.1, we show that $\beta_c^{\mathrm{TS2SLS}}$ can be expressed as a function of the ideal β_c^{OLS} and two bias terms: a prediction error term and an exclusion restriction violation term. The sign of this second bias is generally believed to be positive (Zimmerman 1992), as missing factors that positively influence parental income conditional on the set of instruments are likely to also positively influence the adult child's income conditional on parental income.

As our main claim in the paper is that intergenerational persistence is declining between the 1910s and 1940s cohorts, the key concern with our approach is that these two biases may be changing across cohorts in a manner that produces a decline in our mobility measures, whereas the true measure of mobility actually trends differently. In section V and appendix section D.1.2, we present a variety of evidence against this concern.[16]

E. Comparison to Past Measures of Parental Income

Data limitations have long plagued the study of mobility in the United States, and our approach is no exception. We briefly review the main approaches in the literature, highlighting their advantages and disadvantages to better put our approach and results in context.

1. Papers Using Historical Data

The census provides identified data for individuals in the 1940 and earlier censuses (and is in the process of releasing the 1950 census). Recent papers have used linking algorithms to find the same individual across censuses on the basis of their name, year of birth, and place of birth. This approach faces several challenges, the most important in our context being that, except for white men, linking rates are poor.[17] Most obviously, the linked sample is not representative by sex, as women during this period almost all changed their names upon marriage. To date, all published

[16] To highlight just one example here, we show that for the 1940s and 1950s cohorts—when we can use the NLS and PSID to directly observe parental income, instead of having to predict it—the imputation approach and the ideal OLS approach move together in changes (and, especially in the rank-rank specification, are very close in levels as well). While we cannot directly verify that the two move together in changes over the 1910s–1940s cohorts, we are reassured that they covary in the period of overlap.

[17] See Ferrie (1996) for an important and early contribution to this literature. There is an active literature on the correct linking methodology and the preferred tolerance for rates of falsely matching and missing true matches (see, e.g., Abramitzky et al. 2019; Bailey et al. 2020). Matching methodologies are still in flux, and best practices will likely evolve as machine-learning techniques improve.

mobility papers using census linking drop all women.[18] While in principle Black men are linkable, in practice match rates are very low for them. For example, an important contribution of Ward (2023) is the inclusion of Black men, but his linked sample is only 2% Black before those observations are upweighted. Similarly, Collins and Wanamaker (2022) are able to find reliable adult matches for 3% and 5% of Black children in the 1880 and 1900 censuses, respectively. Moreover, Black Americans—and particularly Black men—are systematically undercounted in censuses even before any linking is performed.[19] Even beyond gender and race, certain types of names are very hard to link with precision (such as very common names, like John Smith, or long foreign names that might have changed over time).

Our approach circumvents many of the challenges associated with linking. In most cases, the link to the father's occupation and other childhood characteristics are merely included as questions answered by the adult child respondent in the survey. Most of our surveys aim to be nationally representative (see table E.1), so the percentage of Black respondents in our (unweighted) data is very close to that in the full US population, even for our earliest cohorts.

That said, there are important subgroups that may be missed even in our surveys. Given our focus on representativeness of the US population, especially by race, the fact that incarcerated or otherwise institutionalized people are unlikely to complete the surveys in our sample may bias our estimates of intergenerational mobility. Figure A.2 shows the share of individuals ages 30–50 who are institutionalized (e.g., in correctional facilities or mental hospitals), separately by subgroup and cohort. The stark increase in the Black male incarceration rate for cohorts born since the 1960s is clear in the census data. But there is little differential trend for Black male institutionalization for individuals born prior to 1960, which are the cohorts that are the focus of our study.

Beyond individuals being linked across time, another challenge for historical work on mobility is the lack of individual or family income data until the 1940 census. Most historical US mobility research focuses on the occupational status of the father (as we do, though we adjust it along additional dimensions) and often the son as well. Relative to a single snapshot of parental or father's income, which is a very noisy proxy for average childhood income and thus leads to severe attenuation bias (Solon 1992), a single snapshot of father's occupation may have the advantage of being more stable over time. But a single observation of a father's occupation has noise from two sources, as Ward (2023) recently highlights. First, fathers change

[18] Recent papers studying historical intergenerational mobility have begun to include women via linking, including Bailey and Lin (2022), Althoff, Gray, and Reichardt (2023), Buckles et al. (2023), and Eriksson et al. (2023).

[19] O'Hare (2019) calculates that the net undercount rate for the Black population has gone from 8.4% in 1940 to 2.5% in 2010.

occupations from year to year, especially when occupations are measured at the three-digit level that is often used in this literature. While this attenuation bias is likely smaller than that from year-to-year changes in family income, it could still be substantial. Ward (2023) shows that mobility estimates using father's occupation as observed in a single census year substantially overestimate mobility relative to those that use multiple observations across different censuses. Second, census takers appear to record occupations with substantial error, at least in the historical period.[20]

We do not observe fathers for just one (or two) census snapshots but rather observe them in the recollections of their adult children during their prime-age years. In that sense, we avoid the problem that census researchers face of potentially observing the father in a particularly unrepresentative year in terms of his occupation. By contrast, it seems natural to assume that the adult child would remember the father's main occupation over her entire childhood, so the retrospective nature of our data likely aids in identifying the chief occupation of the father.

2. Papers Using More Modern Data

The PSID and NLS datasets have many advantages for modeling intergenerational mobility (papers that use these data to estimate mobility for the 1950s through 1970s cohorts include Mazumder [2015], Bratberg et al. [2017], and Davis and Mazumder [2022]). First, they tend to have multiple observations of father or family income while the child is growing up, alleviating concerns about attenuation bias. Second, they have been fielded over decades, so the children can now be observed in their prime-age years, alleviating concerns about life cycle bias. However, it is difficult for long panels such as these to avoid attrition, which typically results in nonrepresentative samples, as the most disadvantaged respondents prove harder to track over time and across generations.[21]

[20] As Ward (2023) details, in a special case when a recensus was required in St. Louis in 1880, one-third of occupations were reported differently only 5 months later, despite the reference date for the occupation being unchanged. Given the challenges of linking, researchers have turned to creative solutions. We noted in the introduction the synthetic panel approached used by Olivetti and Paserman (2015) to study white men and married white women. To the extent that children stay in their parents' households as adults, then household surveys like the census allow researchers to observe both child and parents without needing to link, an insight that Hilger (2015) and Card, Domnisoru, and Taylor (2018) have used to study intergenerational mobility with respect to education. But this approach works only for periods in which most children have completed their education while living with their parents and of course does not provide a workable solution when the outcome of interest is the adult child's family income, as few children remain with their parents during their prime-age years.

[21] Schoeni and Wiemers (2015) show that the patterns of attrition by parent and child income result in biased estimates of intergenerational mobility. Indeed, as we show in table A.2, individuals for whom we observe 5 years or even 10 years of childhood household income in the PSID have fathers who are more likely to be white and much more educated than the general population of fathers.

Chetty et al. (2014b) pioneered the use of administrative data, available since the 1990s, to study US mobility. These data obviate the need for linking (the observations have identification numbers) and are much less susceptible to attrition and attenuation bias, as many years of income of both parents and children are available. Even with these administrative data, there are numerous challenges for mobility research. First, roughly 7% of children cannot be linked to parents for various reasons (in our main sample, for the 1910s–1940s cohorts, roughly twice that share of children are missing information for father's occupation, and we show robustness to various adjustments in the appendix). Second, to date, only the 1980s cohort can be studied (as they are young enough to have lived with their parents in the 1990s when IRS data became available and old enough to be observed today in prime earning years), and even for these individuals, early-childhood income is not observed.[22] Therefore, these data cannot track changes in mobility over decades.

Relative to these data sources, our approach allows us to reach further back in history (though not as far back as census linking—as in Olivetti and Paserman [2015], Song et al. [2020], Collins and Wanamaker [2022], or Ward [2023]—because the types of surveys we use become common only in the 1940s, so they will not capture nineteenth-century cohorts at prime age). However, relative to IRS data, our sample sizes are orders of magnitude smaller, preventing us from breaking the data into neighborhoods or single percentiles, as in Chetty et al. (2014a) and Chetty and Hendren (2018a, 2018b).

IV. Results for Representative Samples

A. *Main Results*

The first series of figure 1 shows the IGE for survey respondents over time, pooling across surveys and applying our baseline population-adjusted weights. We show the IGE separately by decade of birth and report the corresponding estimates in table A.3. Between the 1910s and 1940s birth cohorts, the IGE falls markedly from roughly 0.75 to 0.42. We then see an increase in this measure in subsequent birth cohorts, so that the IGE appears to take on a U shape over time.

The second series shows the results from the rank-rank specification. As is typically found in other papers, our rank-rank coefficients are lower in magnitude than our IGEs: it begins the sample period around 0.37 and declines to a low of 0.25 for the 1940s birth cohorts. While the IGE and

[22] See Heckman, Pinto, and Savelyev (2013), Uguccioni (2021), as well as references therein for evidence that early-childhood resources are especially important to later-life outcomes.

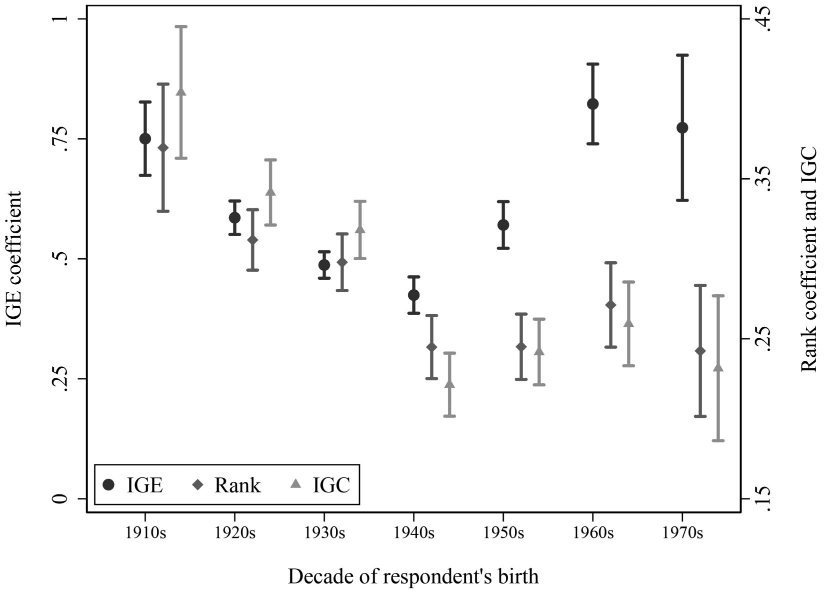

FIG. 1.—IGE and rank-rank measures by birth decade. The figure combines 15 different surveys, which are described in section II and in further detail in appendix E. The IGE and rank-rank estimates are based on the baseline sample of respondents ages 30–50 using equations (2) and (3). The IGC is equal to IGE × (σ_{y^p}/σ_y), and σ_y and σ_{y^p} are the standard deviations of adult children's and parental logged income, respectively. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). We use sample weights where provided and further reweight each birth cohort (i.e., decade) so that they have representative race × sex shares.

rank-rank track each other in changes very closely between the 1910s and 1940s cohorts, the IGE drifts upward for more modern cohorts, while the rank-rank correlation stays relatively flat.[23]

For several reasons, we focus on the decline in the IGE and rank-rank measures that occurs from the 1910s to the midcentury birth cohorts instead of the subsequent rise in the IGE or the stabilization of the rank-rank correlation thereafter. First, as we noted in our discussion of table 2, the share of data with missing information about fathers increases over time, so levels and trends of mobility estimates toward the latter part of our sample period

[23] Figure A.3 plots the estimates separately for each survey in order to give readers a sense of which surveys contribute to each decade's estimate and their relative magnitudes. Given the focus on representativeness, we exclude surveys whose respondents are only one race or not representative of the 30–50 age group.

might reflect sample selection.[24] Second, beginning in the 1960s, data sources with income information for both generations (i.e., modern panel data such as the PSID and later on linked administrative IRS data) become increasingly available.[25] The availability of parental income data is particularly important in these more modern cohorts given rising residual wage inequality since the 1970s (Lemieux 2006), which would likely increase the degree of bias in our estimates through incorrect predictions and omitted variable bias.

Figure 2 shows the decline in intergenerational persistence between the 1910–19 cohorts and the 1940–49 cohorts as binscatter figures. Figure 2*A* shows the change in the IGE relationship: a large shift rightward and upward (reflecting real income growth for both generations) as well as a significant flattening of the slope (because the upward shift is especially large among individuals growing up with less family income). Figure 2*B* shows that the decline in the rank-rank is also large and precisely estimated. Given that by construction there can never be an overall increase in parents' or children's ranks (their average must always be 50), we see only a flattening of the slope.[26] Table A.4 quantifies the decline between the 1910s and 1940s birth cohorts, showing that the IGE (rank-rank correlation) falls roughly 0.007 (0.004) percentage points per year in the 1910s–1940s period.

A natural question that figure 1 raises is why the IGE increases in more recent decades, while the rank-rank correlation stays relatively constant. The diverging paths of these measures can be explained by the fact that, holding the copula fixed, the IGE will rise with the ratio of children's to parents' inequality, whereas by definition, the rank-rank will not. Specifically, using our baseline approach, the variance of log parental income declines over time, implying that the IGE will increase over time even if the covariance of log income across generations is relatively unchanged. In contrast, the variance of ranked parental income is fixed by construction,

[24] Figure A.4 plots the variance of logged (predicted) parental income in the baseline sample. To the extent that the lower variance in the 1950s–1970s cohorts partially stems from sample selection, then the IGE estimates for these later cohorts will be biased upward. Indeed, the robustness checks in sec. V.D show that once we incorporate respondents who provide information about their mothers' occupations (when fathers' occupations are missing), the magnitudes of the IGE in this later time period are significantly reduced.

[25] Davis and Mazumder (2022) find an increase in persistence between the 1950s and 1960s cohorts (we find an increase in the IGE and a modest rise in the rank-rank correlation for the same cohorts). We view their findings as consistent with ours in terms of implying relatively high levels of mobility for cohorts born in the middle of the twentieth century.

[26] While caution is warranted in terms of comparing the levels of our rank-rank estimates (which use predicted parental income) with those from modern administrative data (which use actual income data averaged over several years from the parents), we use the modern estimates as rough benchmarks to assess the importance of the changes. The rank-rank slope we find for the 1910s–1920s cohorts is roughly equal to the modern United States (see Chetty et al. 2014a), whereas the slopes we find for midcentury cohorts are close to the modern estimates in Canada (Connolly, Corak, and Haeck 2019) and Denmark (Helsø 2021).

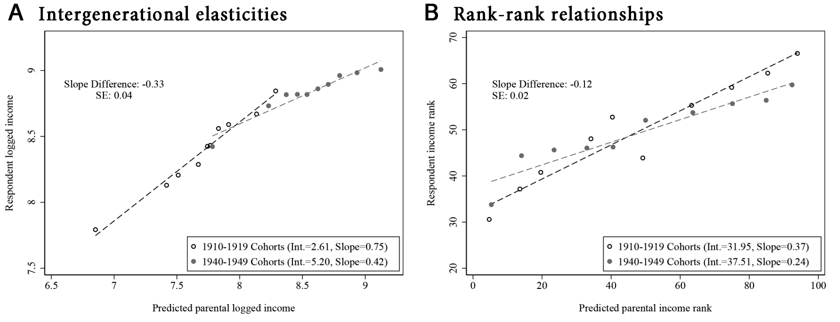

FIG. 2.—Binscatter depictions of decline in intergenerational persistence. Data are from 15 different surveys, described in section II and in further detail in appendix E. The IGE and rank-rank are based on the baseline sample of respondents ages 30–50. The estimated slope difference and its standard error come from regressions similar to equations (2) and (3) but which allow the slope and intercept to differ by cohort. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). We use sample weights where provided and further reweight each birth cohort (i.e., decade) so that they have representative race × sex shares.

so changes in the rank-rank correlation will reflect only changes in the covariance of ranked income across generations. Indeed, the third series of figure 1 illustrates that when we instead compare the rank-rank correlation to the IGC—where IGC $=$ IGE $\times (\sigma_{y^P}/\sigma_y)$ and σ_y and σ_{y^P} are the standard deviations of adult children's and parental logged income, respectively—the trends in the two measures coincide throughout the twentieth century.

Overall, we believe that the decline in intergenerational relative persistence from the 1910s to the 1940s cohorts is a novel finding, though there have been hints of it in past work. Using a dynamic Occ.-Occ. mobility approach for white men (where status is fixed for all men within an occupation-decade but occupational status can change over time), Song et al. (2020, 253) find a modest decline for those born around 1946–55, which we also find as roughly the nadir of our IGE and rank-rank series. They write, "We consider the deviation of the 1950 birth cohort best interpreted as suggestive." Similarly, including white and Black men, Ward (2023) finds that mobility is significantly lower in 1910 than in 1960 (though he does not have data for the intervening years), again consistent with our results for representative samples.

B. *Comparison to Occupational Mobility*

We adopted a Self-reported inc.–Predicted inc. approach to better include women and nonwhite individuals, but a natural question is how our results compare with the more traditional Occ.–Occ. measures. Note that we can perform this comparison for only men. In figure A.5, we show that the standard Occ.–Occ. approach using the census occscore variable shows only limited decline in intergenerational persistence (fig. A.5*a* includes all men, and fig. A.5*b* includes just white men). Similarly, there is little decline when using an Occ.–Predicted inc. approach (second series). The third and fourth series show results for Predicted inc.–Predicted inc. and Self-reported inc.–Predicted inc., and in both cases a large decline in intergenerational persistence appears. The differences between the later (third and fourth) and earlier (first and second) series suggest that an important part of the rise in male-only mobility comes from within-occupation upgrading of men with low-status fathers. As such, while the motivation for adjusting the Occ.–Occ. measure was in large part to include women, the adjustments also provide new insights for male-only mobility during this period and further show why the decline in persistence was harder to detect with more traditional, occupation-based mobility measures.

As an alternative way to see the difference between our approach and the more traditional occupation-based measures, we can study how our full-sample mobility estimates change as we transition from using only father's occupation to incorporating information about respondents' race and geography in the income prediction. (Note that we use reported family income as the dependent variable in this exercise in order to include women.) The first series of figure A.6 shows the occupation-only estimates, confirming a decline in persistence between the 1910s and 1940s cohorts. Incorporating race and Southern residence into the prediction of childhood income accelerates the 1910–40 decline in the IGE (statistically significant at $p < .01$). Its effect on the rank-rank correlation is also visually evident, and in fact using only father's occupation to predict childhood income would have reduced the overall decline by roughly one-third.[27] This figure thus shows that using a richer set of predictors than merely father's occupation increases in a statistically significant manner (and, for the rank-rank, economically significant) the estimated decline in intergenerational persistence from the 1910s to 1940s cohorts.

In summary, we have so far provided evidence of a significant decline in IGE and rank-rank persistence measures between the 1910s and 1940s birth cohorts. Importantly, these results reflect samples that are representative

[27] In particular, as we add predictors of parental income, both the variance of logged parental income as well as the covariance of logged income across generations increase, so that the trends in the IGE remain relatively unchanged. By contrast, the covariance of ranked income across generations increases, so that the rank-rank correlation always increases as we include predictors beyond occupation. Table A.5 shows the R^2 from regressing logged income on the predictors. The table highlights that occupation is certainly an important predictor of income, but incorporating race, region, and education all improve the power of our measures to predict household income.

of the full US-born population, including women and nonwhite respondents. In section V, we show robustness of this result to what we consider to be the most central concerns.

V. Robustness of the Full-Population Result

We divide our robustness checks into three main concerns: measurement error of childhood predicted income, life cycle bias, and econometric challenges related to the two-stage estimation. Section V.D summarizes robustness checks that do not fit into these main categories.

While more details on all of these results are provided in various appendixes, table 3 summarizes more succinctly how the main result—the decline in the IGE and rank-rank correlation between the 1910s and 1940s cohorts—holds up after changing methodological choices.

A. Measurement Error in Predicted Childhood Income

As noted earlier, a key challenge for our approach is measurement error in estimating the respondent's parental income during her childhood. This measurement error can arise from several sources, which we address in turn below.

1. Recall Bias

Section III.B already provided evidence that the recall of father's occupation appears reasonable—men and women give the same answers on average, and the answers given match the occupational mix of actual fathers in the census during the period in which the respondent grew up. We also performed a direct validation using the PSID, where we can observe the father reporting his own occupation and then decades later observe the adult child's recollection of that occupation. Figure B.1 shows that if the types and frequencies of recall errors made in the PSID were made in all of our other surveys, our main result of declining intergenerational persistence between the 1910s and 1940s cohorts would still hold.

2. Unobserved Within-Cell Variance

Our baseline approach assigns each respondent a childhood income based on the mean family income in a father occupation × race × South cell from the appropriate census or other auxiliary dataset, and it thus ignores within-cell variation. To the extent that some within-cell variation in a single census year is merely transitory, excluding within-cell variation will better approximate permanent average childhood income. But to the extent that within-cell variation reflects systematic income differences

TABLE 3
DIFFERENCES BETWEEN 1910 AND 1940 IGE USING VARIOUS ADJUSTMENTS

	IGE			RANK-RANK		
	Ratio (1)	Difference (2)	Racial Convergence (3)	Ratio (4)	Difference (5)	Racial Convergence (6)
Baseline	.57	.33	.19	.66	.12	.05
	(.04)	(.04)		(.05)	(.02)	
	A. Alternative Parental Income Measures					
Using Collins and Wanamaker (2022) farm fix	.58	.30	.16	.63	.14	.05
	(.04)	(.04)		(.04)	(.02)	
Dropping farmers	.57	.30	.18	.68	.10	.06
	(.06)	(.07)		(.07)	(.03)	
Only occupation in prediction	.57	.31	.03	.72	.09	.00
	(.05)	(.06)		(.05)	(.02)	
Using IPUMS occscore	.62	.21	.01	.66	.10	.00
	(.05)	(.04)		(.05)	(.02)	
Using father's income	.54	.31	.17	.63	.14	.05
	(.04)	(.04)		(.04)	(.02)	
Using nearest census	.63	.27	.16	.64	.14	.04
	(.04)	(.04)		(.04)	(.02)	
Alternative weights in prediction	.57	.32	.18	.67	.12	.05
	(.04)	(.04)		(.05)	(.02)	

	B. Including Missing Income					
Includes nonworking fathers	.56	.33	.20	.67	.12	.05
	(.04)	(.04)		(.05)	(.02)	
Using mother's occupation	.54	.33	.19	.66	.13	.05
	(.04)	(.04)		(.05)	(.02)	
Includes mothers and nonworking fathers	.54	.33	.20	.67	.12	.05
	(.04)	(.04)		(.05)	(.02)	
	C. Additional Predictors					
Using father's education	.65	.20	.09	.67	.12	.05
	(.05)	(.04)		(.05)	(.02)	
Using childhood region	.54	.33	.20	.60	.15	.06
	(.04)	(.04)		(.05)	(.02)	
	D. Different Age Ranges					
30–45	.53	.38	.20	.64	.14	.05
	(.04)	(.06)		(.06)	(.03)	
35–50	.65	.26	.17	.70	.11	.05
	(.05)	(.05)		(.05)	(.02)	

NOTE.—Columns 1 and 4 report the ratio of the IGE and rank-rank correlation estimates, respectively, for the 1940s cohort relative to the estimates for the 1910s cohort. Columns 2 and 5 report the difference in the IGE and in the rank-rank correlation, respectively, between the 1910s and 1940s cohorts. Columns 3 and 6 report the difference in the third term of the decomposition (i.e., the third term of eq. [4], denoting average income differences by race) between the 1910s and 1940s cohorts. "Using Collins and Wanamaker (2022) farm fix" refers to using the adjustments to farmer and self-employed income from Collins and Wanamaker (2022). "Using father's income" refers to using the father's personal rather than household income. "Using nearest census" refers to using the nearest census to childhood (i.e., 1940 with 1936 adjustments for the 1910s–1930s birth cohorts and 1950–80 for the 1940s–1970s cohorts, respectively). "Alternative weights in prediction" refers to using the number of children a father has when constructing average predicted parental income.

missed by father occupation × race × South, our measure of predicted childhood income will bias us away—in an a priori unclear direction—from the persistence measure of interest.

We begin to address this concern by reestimating predicted childhood income on a subset of our data that includes more information on childhood background, namely, father's education and detailed childhood region. A priori, father's education is one of the most likely factors to create systematic deviation from our father occupation × race × South–based mean family income. Indeed, adding information about father's education to our standard approach significantly increases predictive power (e.g., in 1960, the R^2 rises from 0.29 to 0.33; see table A.5). Figures B.2 and B.4 show that when we improve our childhood income measures with important predictors, the trends in mobility remain unchanged, providing some reassurance that systematic, unobserved within father occupation × race × South cell variation in income is not driving our results.[28]

We now take a different approach to assessing the extent of potential bias due to unobserved within-cell variance. Essentially, we ask, even if we assume that all within-cell variance reflects true permanent differences in childhood income, can we still detect a decline in intergenerational persistence between the 1910s and 1940s cohorts? For each father occupation × race × South cell, we observe the actual family income values of all observations in that cell (i.e., in microdata from the appropriate census or 1936 Expenditure Survey). We thus reestimate the IGE using both a multiple imputation estimation (see, e.g., Rubin 1987; Little and Rubin 2019) as well as direct draws from the empirical distribution of all observed family income values (fig. B.5). We find that even when we make maximal assumptions—that all within-cell variation reflects permanent variation in childhood income—we find a decline in intergenerational persistence between the 1910s and 1940s cohorts.

3. Farmer Income

Our baseline measure of parental income acknowledges the difficulty in estimating farmer (and self-employed) income in the first half of the twentieth century using conventional survey or census data. We therefore use the 1900 census of agriculture (for farmers) as well as the 1936 Expenditure Survey (for farmers and self-employed) given the limitations of the

[28] Note that incorporating this information into the imputation implicitly tests the robustness of the two-sample approach, as different predictions will emit different prediction error and exclusion restriction violation bias terms. In app. D, we also show that the 1910s–1940s decline in persistence is robust to reducing the number of variables used to impute parental income. As our estimated coefficient is a function of the true (unobserved) OLS target parameter and the two bias terms, the fact that we consistently find a decline from the 1910s to 1940s cohorts—despite the bias terms changing with each variation of the predictions—suggests that a decline in the true target parameter is driving the estimated decline.

1940 census for these groups. In figure B.6, we show that our main result is unchanged when imputing farmer and self-employed income using an alternative approach that follows Collins and Wanamaker (2022) and when dropping farmers from the sample.

B. Life Cycle Bias

Various papers in this literature have noted that using current income to proxy for adult children's lifetime earnings may bias estimates of mobility (see, e.g., Haider and Solon 2006; Lee and Solon 2009; Nybom and Stuhler 2016). Recall that we already restrict the sample to ages 30–50 to limit life cycle effects. However, figure B.7 shows the robustness of the main result to alternative specifications and sample restrictions that attempt to further minimize this life cycle bias (e.g., including polynomials in adult children's age and restricting the sample to older respondents whose total family income may be better approximations of their lifetime earnings).

C. Robustness to Econometric Approach

1. Functional Form

One concern with the empirical approach is that we rely on the log or rank transformations for estimating relative mobility. Following Dahl and Lochner (2012) and Løken, Mogstad, and Wiswall (2012), in table D.1, we instead consider levels of income for both the survey respondents and their parents. This table confirms the weakening relationship between income across generations, with the main decline occurring between the 1910s and 1940s cohorts.[29] In appendix D, we also show that the coefficient from the levels-on-levels regression can be transformed using the first and second moments of the parent and child marginal income distributions to generate close approximations of the IGE and rank-rank measures, confirming that the nonlinear transformations embedded in the latter two measures are not driving the rise in mobility (fig. D.7).

2. Connection to TS2SLS

As noted in section III.A, the baseline empirical approach is similar in spirit to a TS2SLS approach. In appendix D, we implement the TS2SLS

[29] Table D.2 estimates quadratic specifications using levels of income, finding that the slope of the relationship at the 25th percentile of the parental income distribution also flattens between the 1910s and 1940s cohorts. Tables D.3 and D.4 show analogous results for the IGE and rank-rank correlation, confirming a decline in persistence between the 1910s and 1940s birth cohorts for individuals throughout the (predicted) parental distribution (i.e., at the mean and at the 10th and 90th percentiles).

approach using the nearest source of microdata (i.e., the 1936 Expenditure Survey and the 1940–80 censuses) to predict parental logged income. Because of the lack of first-stage microdata for the 1910s cohorts, we cannot replicate the entire 1910s–1940s persistence decline in this exercise.

Figure D.4*a* shows that when we refrain from using nonlinear transformations and instead use levels of income for both generations, the TS2SLS estimator and the OLS estimator using imputed averages are numerically identical. The levels-on-levels specification exhibits a strong decline from the 1920s to the 1940s.[30]

In figure D.4*b*, we implement TS2SLS using the log-log functional form, which is not numerically identical to the OLS imputation approach. The former uses the most contemporaneous source of microdata to predict logged income in the parental generation, whereas the latter computes the average predicted income for each cell and then applies the log transformation.[31] The third series in this figure displays robust TS2SLS standard errors (Pacini and Windmeijer 2016; Choi, Gu, and Shen 2018). This panel highlights that although the levels differ slightly—resulting from the different moment in which the log transformation is applied—the mobility trends are very consistent with our baseline results (a U shape with a 1940s nadir), and in fact the TS2SLS approach displays a somewhat more marked 1920s–1940s decline.

In figure D.1, we present results showing robustness to varying the set of parental income predictors. We show that the 1910–40 decline in persistence holds using any subset of instruments. Further, any subset of instruments that includes occupation yields very similar results, and it is only when race and region are used without occupation—a case where the exclusion restriction is much more likely to be violated—that mobility estimates diverge from our baseline estimates.

D. Other Robustness Checks

We also conduct a variety of other exercises in appendix B that check the robustness of the 1910s–1940s decline to other sampling and specification

[30] In fact, the one difference between the levels-based specification and our baseline results is that for the former, the persistence decline continues through the 1950s birth cohort before plateauing and then reversing, whereas the nadir using our baseline methodology occurs for the 1940s cohort. The divergence between the levels, logs, and rank specifications in 1950 comes from different standardizations of income. See app. D for a full treatment.

[31] As noted in sec. III, an analogous approach for estimating the rank-rank correlation via TS2SLS is not desirable. Instead, we show an approximation of the rank-rank correlation based on the coefficient from a levels specification (i.e., on the basis of the normal distribution, the rank-rank measure is $\beta^{\mathrm{RR}} = (6/\pi)\arcsin(\beta^{\mathrm{IGC}}/2)$, where β^{IGC} is the IGC calculated from β^{levels}). We show that this approximation is quite close to the estimates in the main text.

choices. Notably, we incorporate into the sample respondents whose fathers were present but not working (e.g., retired) as well as respondents who provided information about their mother's occupation. We also consider the sensitivity of the results to alternative weighting schemes, to including survey-year fixed effects, and to changes in household size (i.e., adjusting measures of income using reported household size).

As a final plausibility check on our main result, we examine a different outcome variable for the adult children: education. While the exact return to education varies over time, on average, more educated individuals have significantly higher earnings and family income. Thus, it would be somewhat surprising if the predictive power of parental income on children's education did not fall given that its predictive power over adult family income did. In figure A.7, we estimate variants of equations (2) and (3) where we put the adult child's self-reported years of schooling as the outcome variable (available in all of our datasets). The figure shows that the relationship between father's predicted income and respondent's educational attainment declines sharply between the 1910s and 1950s birth cohorts.[32] Figure A.7*b* uses the subset of 12 surveys for which we have father's education and shows that the declining education-on-predicted-income correlation is mirrored by a declining education-education correlation.[33] Recall that father's education is not used as a predictor of father's income, so the observed decline in education-education correlations over time is an independent check on our main IGE and rank-rank results showing declines in persistence over the first half of the century. Moreover, while we have tried to address concerns about using auxiliary data in a two-step process in section V.C, the education-education results are further reassuring because no first-stage prediction is required.

VI. Decomposing the Rise in Mobility

In this section, we show how to decompose the overall IGE and rank-rank relationships into factors related to subgroups, building on Hertz (2008). We then use this decomposition to show how much changes in mobility or

[32] Using data from the modern period, Landersø and Heckman (2017) has questioned whether mobility is truly lower in places such as Scandinavia than in the United States, because when education of the adult child is the outcome of interest instead of earnings or income, mobility measures in the United States and Scandinavia look more similar. In our analysis, both family income and years of education appear to have a decreasing dependence on predicted childhood income over the first half of the twentieth century. Figure A.8 illustrates these changes using binscatter figures, highlighting that this weakening relationship is largely driven by the rapid increase in respondents' high school completion in the bottom half of the income distribution rather than the later rise in college completion.

[33] These patterns mirror the rise and subsequent decline in relative educational mobility documented in Hilger (2015), which restricts attention to adult children living with their parents at the time of the census.

income among subgroups—particularly by race and gender—explain the overall rise of mobility over the first half of the twentieth century.

A. *Decomposing the IGE and Rank-Rank Slopes*

Consider any partition of the full sample, emitting subgroups $g \in G$ with subgroup g's share of the total sample given by p_g. Further, let β_g^{IGE} be equal to β from estimating equation (2) on subgroup g.

From the OLS formula and the law of total covariance, the whole-population IGE is given by

$$\beta^{\mathrm{IGE}} = \underbrace{\sum_{g \in G} p_g \frac{\mathrm{Var}(y^{\mathrm{P}} \mid g)}{\mathrm{Var}(y^{\mathrm{P}})} \beta_g^{\mathrm{IGE}}}_{\text{Weighted average of subgroup slopes}} + \underbrace{\frac{\mathrm{Cov}(\mathrm{E}[y \mid g], \mathrm{E}[y^{\mathrm{P}} \mid g])}{\mathrm{Var}(y^{\mathrm{P}})}}_{\text{Between-group covariance of subgroup averages}}. \tag{4}$$

The formula remains unchanged for β^{IGC} save for setting $\mathrm{Var}(y^{\mathrm{P}}) = 1$, if the y and y^{P} are considered to be standardized (demeaned and divided by standard deviation) versions of logged income. A slight modification gives a similar (and more novel) expression for the whole-population rank-rank slope γ^{RR}, equal to γ from equation (3). If we assume that both the parental and adult children's ranked incomes have a uniform distribution, the same application of the law of total covariance gives

$$\begin{aligned} \gamma^{\mathrm{RR}} &= \sum_{g \in G} p_g \frac{\mathrm{Var}(\mathrm{Rank}^{\mathrm{P}} \mid g)}{\mathrm{Var}(\mathrm{Rank}^{\mathrm{P}})} \gamma_g^{\mathrm{RR}} + \frac{\mathrm{Cov}(\mathrm{E}[\mathrm{Rank} \mid g], E[\mathrm{Rank}^{\mathrm{P}} \mid g])}{\mathrm{Var}(\mathrm{Rank}^{\mathrm{P}})} \\ &= 12 \times \left(\sum_g p_g \mathrm{Var}(\mathrm{Rank}^{\mathrm{P}} \mid g) \gamma_g^{\mathrm{RR}} + \sum_g p_g \mathrm{E}[\mathrm{Rank}^{\mathrm{P}} \mid g] \mathrm{E}[\mathrm{Rank} \mid g] - 0.25 \right). \end{aligned} \tag{5}$$

To ease intuition and to focus on one of the key applications for our paper, we rewrite the IGE decomposition for two groups, namely, white (W) and Black (B) respondents:

$$\begin{aligned} \beta^{\mathrm{IGE}} = {} & p_W \frac{\mathrm{Var}(y^{\mathrm{P}} \mid W)}{\mathrm{Var}(y^{\mathrm{P}})} \beta_W^{\mathrm{IGE}} + (1 - p_W) \frac{\mathrm{Var}(y^{\mathrm{P}} \mid B)}{\mathrm{Var}(y^{\mathrm{P}})} \beta_B^{\mathrm{IGE}} \\ & + \frac{p_W \mathrm{E}[y^{\mathrm{P}} \mid W] \cdot \mathrm{E}[y \mid W] + (1 - p_W) \mathrm{E}[y^{\mathrm{P}} \mid B] \cdot \mathrm{E}[y \mid B] - \mathrm{E}[y^{\mathrm{P}}] \mathrm{E}[y]}{\mathrm{Var}(y^{\mathrm{P}})}. \end{aligned} \tag{6}$$

The decomposition helps clarify two points. First, because population shares act as weights in the first two terms of equation (6), changes in the within-group IGE of the large majority group, β_W^{IGE}, will—perhaps not surprisingly—affect the full-population IGE. By the same logic, while the Black-only slope, β_B^{IGE}, may be of interest in other applications, it will not play a large role in determining the overall slope.

Second, and less obviously, the decomposition highlights the important role of between-group differences in parental income y^p in determining the full-population IGE. To see this point, assume for the moment that W and B are two distinct subgroups but are drawn independently from the same distribution of parental income y^p. In this special case of no between-group differences in parental income, $\beta^{\mathrm{IGE}} = p_W \beta_W^{\mathrm{IGE}} + (1 - p_W)\beta_B^{\mathrm{IGE}}$, or in other words, the full-population IGE is the average of the two subgroup IGE slopes weighted by the subgroup share of the total population. This result holds regardless of the adult childhood outcomes (e.g., even if the mean adult income y of group B is well below that of group W).[34] However, if there exist large differences in parental income between the two groups (as there are for Black and white Americans), then this third term will be heavily weighted and will play a key role in determining the full-population IGE.[35]

B. *Decomposing Mobility by Race and Gender*

In this section, we show visually the mappings of (predicted) parental income to adult children's incomes, separately by the race and gender of the respondent, and how these mappings change over time. As we will split our data into small race × sex subgroups in the following analyses, to gain power and precision we will typically compare the low-mobility 1910s–1920s birth cohorts to the high-mobility 1940s–1950s cohorts.

1. Decomposing Mobility by Race

Given the discussion in section VI.A, we expect that the between-group component will prove important for a decomposition along racial subgroups, given that Black Americans grow up with far less parental income than white individuals in our period (as well as today). For this reason, we show results as binscatter graphs, as it is important to examine the means of childhood and adult income by group and over time.

Figure 3 shows Black and white mobility for the earlier less mobile cohorts compared with the later more mobile cohorts (IGE in fig. 3*A* and rank-rank correlation in fig. 3*B*). Perhaps the most striking aspect of the graph is how little overlap there is in the support of the Black versus white distributions: Black and white childhood income overlaps only modestly,

[34] In the less extreme case in which the two groups have the same average parental income but different variances, then the third term still cancels out, and the full-population IGE is a weighted average of the subgroup IGEs, where the weights are a function of population shares and (conditional and unconditional) variances of parental income.

[35] This decomposition thus highlights that income changes in the parental generation will affect the subsequent generation's level of mobility, a point explored in greater detail in Nybom and Stuhler (2023).

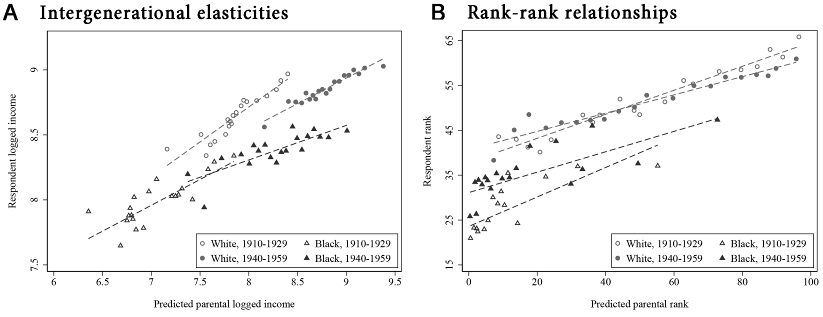

FIG. 3.—Mobility by race, 1910s–1920s versus 1940s–1950s. The figure combines 15 different surveys, which are described in section II and in further detail in appendix E. The IGE and rank-rank are based on the baseline sample of respondents ages 30–50. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). We use sample weights where provided and further reweight each birth cohort (i.e., decade) so that they have representative race × sex shares. To account for pooling of cohorts, specifications include birth decade fixed effects.

especially in the early period. In the rank-rank figure, almost no white respondents grow up in the bottom 10% of predicted childhood income, and few Black respondents grow up above the 30th percentile, so the overlap of the two groups mostly occurs over an interval of approximately 20 percentiles.[36]

Another notable result is the significant progress Black respondents make relative to their white counterparts in both the parents' and children's generations. In the IGE graph, both the Black and white regression lines shift rightward, denoting substantial average real income growth during these respondents' childhood but more so for Black Americans. The rank-rank graph cannot capture average real income growth given its zero-sum nature, so the overall support is fixed between zero and 100. The catch-up of Black adult income here is striking. A Black child in the earlier cohorts growing up at the 15th percentile (which we choose as a point of maximal overlap between Black and white children) would be predicted to have an adult family income at the 28.5th percentile compared with the 41.9th percentile for a similarly situated white child. But

[36] One feature of our small data is that the vast differences between how Black and white children grow up are readily apparent in the support of these figures: with full-population administrative data, one can capture the tiny number of Black children who grew up in rich families and thus extend the regression lines over the entire 0–100 domain of parental income rank. But even today, prime-age Black adults are vastly underrepresented in the upper parts of the parental income distribution while growing up. The tiny share of Black children in the upper ranks of parental income distribution even in modern data can be seen in the appendix figures of Chetty et al. (2020).

for midcentury cohorts, Black children born at the 15th percentile are predicted to appear at the 34.5th percentile as adults compared with the 43.7th percentile for white individuals (closing the gap with their white counterparts from around 13.4 to 9.2 percentile ranks).

While we have so far focused on Black-white convergence, the regression lines depicting white-only mobility also change over this period. In both the IGE and the rank-rank estimates, the slopes flatten significantly. The rank-rank slope falls from 0.27 to 0.20. With white individuals as the large majority group, the flattening of the mobility slope among white individuals will have an important effect on the overall full-cohort IGE and rank-rank estimates (while the Black-only slope also flattens over time in both graphs, given that this component is weighted by a small population share, the effect on overall mobility will be very small).

2. By Gender

A major motivation for our family-income-to-predicted-childhood-income mobility concept is that it enables us to perform intergenerational mobility estimation including women. The decomposition in section VI.A suggests that the key elements of a decomposition of mobility by gender will differ from that by race. Because women and men grow up on average in the same households in the United States, the between-group component of equation (6) should be close to zero, and thus the full-population IGE is well approximated by the simple mean of the within-gender IGE slopes (as each sex is roughly half of the population). Put differently, the male-only IGE will be a biased measure of the full-population IGE only if the female slope is significantly different than the male slope, and differences in adult income means between the two groups will not matter.

Of course, a priori, there is no reason to assume that the mobility slopes of men and women will coincide. For example, marriage patterns could differ by parental income, and they will tend to matter more for women's family income than for men's, especially in the historical period when most married women did not work.

Figure 4 (as well as tables A.6 and A.7) compares male and female slopes over time instead of using binscatter graphs, as between-group mean childhood income differences are trivial. For both measures and for all birth decades, persistence measures for women are greater than or equal to those for men. The male-female gap appears to be relatively stable over time, especially for the 1910s–1940s cohorts.[37]

[37] As noted in sec. II, some of our datasets include only women (e.g., the National Longitudinal Surveys of Mature or Younger Women) or only men (the Occupational Changes in a Generation datasets), so a possible concern is that the differences in mobility by sex are an artifact of using different datasets. In fig. A.9, we show robustness to restricting the baseline sample to datasets that include both men and women (roughly 47% of the baseline sample).

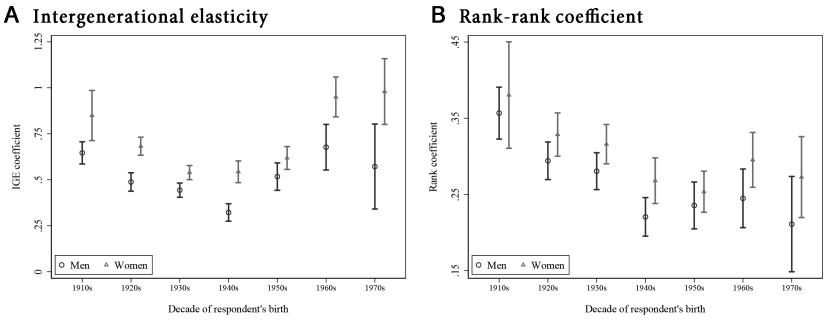

FIG. 4.—IGE and rank-rank measures by birth decade, by sex. The figure combines 15 different surveys, which are described in section II and in further detail in appendix E. The IGE and rank-rank are based on the baseline sample of respondents ages 30–50. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). We use sample weights where provided and further reweight each birth cohort (i.e., decade) so that they have representative race × sex shares.

Why does women's adult family income depend more on their parents' income than is the case for men? To answer this question, we turn again to differences by race.

3. By Race and Gender

We now consider differences by race separately for men and for women. In particular, figures 5 and 6 further break down the by-race results in figure 3 by gender. Figure 5 shows that among men, Black Americans closed much of the mobility gap with white individuals by midcentury (of course, as the supports of the regression lines make clear, Black men still grew up in far poorer households, so their average adult income in either logs or ranks is still much lower than that of white individuals). By midcentury, there is considerable overlap in adult outcomes between Black and white men born to similarly advantaged parents. For example, in the more mobile midcentury cohorts, Black men born at the 15th percentile are predicted to appear at the 37.6th percentile as adults, just slightly below their white counterparts at the 43.8th percentile. This 6.2 percentile point gap is 10.4 points in the earlier cohorts, with Black men born at the 15th percentile predicted to appear at the 32.4th percentile as adults compared with the 42.8th percentile for their white counterparts.

Figure 6 paints a different picture for women. First, when we compare figures 5 and 6, it is clear that Black adult women are simply poorer than their male counterparts. Their entire regression line is below that of Black men. In the early cohorts, a Black woman born at the 15th percentile is

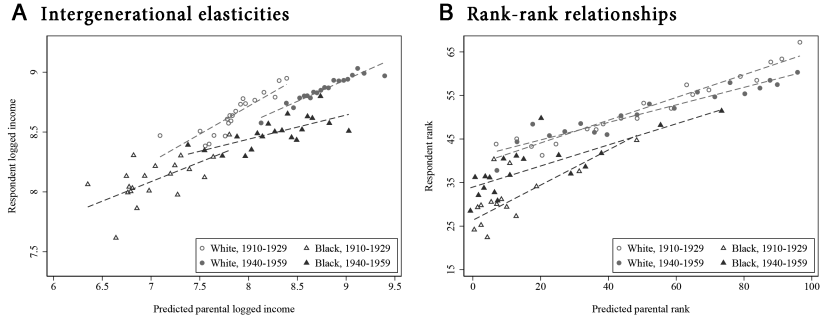

FIG. 5.—Mobility by race for men, 1910s–1920s versus 1940s–1950s. The figure combines 15 different surveys, which are described in section II and in further detail in appendix E. The IGE and rank-rank are based on the baseline sample of respondents ages 30–50. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). We use sample weights where provided and further reweight each birth cohort (i.e., decade) so that they have representative race × sex shares. To account for pooling of cohorts, specifications include birth decade fixed effects.

predicted to barely climb upward at all (an expected adult family income percentile rank of 25.2 compared with 41.0 for a similarly situated white woman). While Black women make progress over time, even at midcentury the corresponding prediction is only the 31.9th percentile (compared with 43.7 for white women).

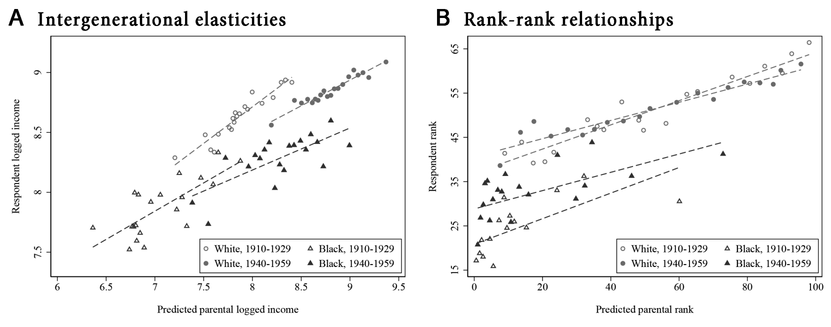

FIG. 6.—Mobility by race for women, 1910s–1920s versus 1940s–1950s. The figure combines 15 different surveys, which are described in section II and in further detail in appendix E. The IGE and rank-rank are based on the baseline sample of respondents ages 30–50. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). We use sample weights where provided and further reweight each birth cohort (i.e., decade) so that they have representative race × sex shares. To account for pooling of cohorts, specifications include birth decade fixed effects.

Thus, for midcentury cohorts, while the racial mobility gap at the 15th percentile for men is down to 6.2 percentiles (from 10.4), it remains at 11.7 (down from 15.8) for women. While Black women make considerable progress over time, given their low starting point, even in the most mobile midcentury period, a Black girl is predicted to grow up to be significantly poorer than any other group born to similar circumstances.[38]

By contrast, white boys and girls both grow up with the same childhood income, but conditional on their place in the childhood income distribution, they also enjoy similar family income as adults. Indeed, comparing figures 5 and 6 shows that the white-only mobility slopes are nearly identical for men and women. For the rank-rank correlation, the male and female slopes are both 0.27 in the early period and 0.20 in the later period. That white men and women's family incomes were equally tied to the status of their fathers in an era when most married white women did not work suggests that they were marrying individuals very similar in earnings to their brothers.

In summary, the higher IGE and rank-rank persistence measures for women relative to men in figure 4 are not driven by white individuals. Instead, the fact that Black women do poorly relative to Black men in adulthood pulls down the overall female mobility regression line for the lowest percentiles of parental income and results in a steeper slope for full-population female mobility relative to male mobility throughout much of the twentieth century.

C. *Decomposing the Decline in Intergenerational Persistence*

As already discussed, the full-population persistence slope is approximately equal to the (simple) mean of the male-only and female-only slopes. Because the gap between those two slopes is quite stable between the 1910 and 1940 cohorts (shown in fig. 4), a decomposition by sex is unlikely to help us explain the decline in full-population persistence over this period. So we consider the decomposition by race instead.

Returning to figure 3 with the decomposition in mind allows us to assess the effects of the various movements in the by-race IGE and rank-rank mappings. Figure 3 depicts a number of different changes over time, some of which will increase mobility (the income growth for Black respondents, the flattening slope for the white majority), some of which will reduce mobility (the income growth for white individuals), and some of which should

[38] Note that the lack of gender gaps by family income among white respondents and the large gaps (favoring men) among Black respondents is apparent in the basic summary statistics shown in table A.1, both in our surveys and in the census.

have minimal effect (the flattening of the Black-only slope). The decomposition can quantify the various contributions.

We begin by considering the role of Black-white income convergence over the first half of the twentieth century for the increase in the IGE. Figure 7 shows (second set of bars) that if Black individuals had instead experienced the same real income growth as white individuals during adulthood (without changing the slopes for either group or the averages or variances of parental incomes), then 57% of the IGE decline would not have been realized. Thus, Black respondents' catch-up to white individuals in income levels over this period explains a large share of the total decline in persistence, despite Black Americans being only a small share of the population.

The flattening of the white slope also plays a major role in the decline of the IGE: had it retained its 1910s level and otherwise allowing all other factors to move as they actually did between the 1910s and 1940s cohorts,

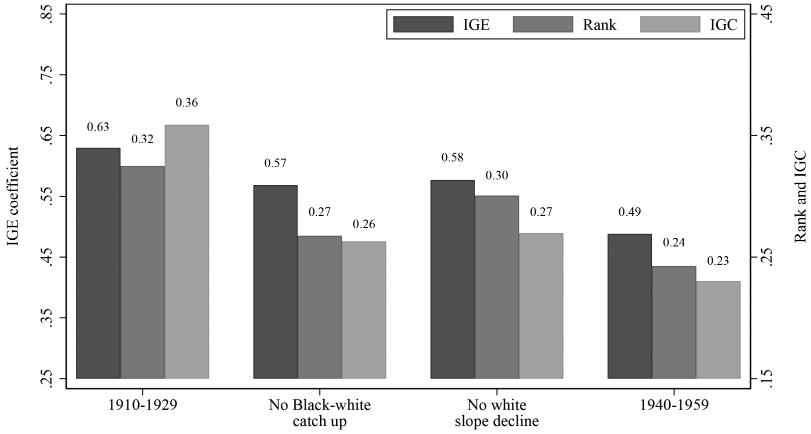

FIG. 7.—Decomposing rise in mobility from 1910s–1920s to 1940s–1950s. The figure combines 15 different surveys, which are described in section II and in further detail in appendix E. The figure shows the contribution of different components of the decomposition in section VI to the increase in intergenerational mobility that occurred between the 1910s–1920s cohorts and the 1940s–1950s cohorts. Specifically, the figure shows the contribution of the reduction in the white-only persistence measure and the contribution of the between-group convergence in income levels. "No Black-white catch up" refers to the assumption that Black respondents had the same income growth as white respondents in log points in the adult children's generation. "No white slope decline" refers to white individuals in the 1940s birth cohorts having the same slope as the 1910s cohorts (without altering the average incomes of white and Black adult children). IGE (dark gray bars) and IGC (light gray bars) display the intergenerational elasticity and correlation, respectively, in the early and late period as well as under these two scenarios. Rank (medium gray bars) displays the rank-rank correlation in the early and late period as well as under these two scenarios (reranking individuals after altering their logged incomes to reflect both scenarios). To account for pooling of cohorts, specifications include birth decade fixed effects.

then the IGE would have fallen only 5 points instead of 14 (0.58 vs. 0.49). This result also emphasizes the importance of the Black-white income convergence: even though white individuals experienced no increase in mobility, the convergence in average income across races still yields more than one-third of the decline in persistence.

We repeat this analysis for the rank-rank correlation as well as the IGC. Separating the decline of the rank-rank measures into within- and between-group components is slightly complicated by the fact that either changing the slope of white respondents' income or altering Black mean income to grow at the same lower rate as white income will mechanically change the ranked income of individuals in the other group. Hence, for the rank-rank measure, we account for the effects of the whole distribution by first conducting the IGE counterfactual, reranking adult children in the counterfactual distribution, and then estimating γ^{RR} in the counterfactual late period. For the IGC, we compute β^{IGE} and multiply it by σ_{y^p}/σ_y, where σ_y and σ_{y^p} are the standard deviations of counterfactual adult children's and actual parental logged income, respectively. Results for the IGC and the rank-rank are similar both qualitatively and quantitatively. As with the IGE, both Black-white catch-up and the flattening of the white-only slope each explain a sizable portion of the decline.

In summary, while the exact shares are sensitive to the use of the IGE, the IGC, or the rank-rank, in all cases we find that Black-white income convergence and the flattening of the white-only slope are the key changes that drove the decline in full-population intergenerational persistence in the first half of the twentieth century.

D. *Convergence in Racial Income Gaps Using Census Data*

The analysis above suggests that convergence in white-Black means—the third term of the decomposition—is a major factor in the decline in overall intergenerational persistence. One implication is that we can calculate this component of mobility without access to data that link adult children to their parents. Figure A.12 thus plots the third term (i.e., the between-group term) of the IGE and rank-rank decomposition using (unlinked) census data as an additional robustness check for the full-population mobility decline from section IV. The same U shape appears when considering logged income, as in the IGE in figure 1. Similarly, the L shape of the rank-rank correlation also emerges when using ranked income. Thus, readers who remain skeptical of adult children's recall or have other concerns about measurement error in childhood income can observe that the (large) component of mobility estimates that do not rely on linking can be replicated using completely different data than our 15 surveys.

Note that Margo (2016) and others have already documented much of this Black-white convergence, though it was not organized by birth cohorts, did not focus on fathers (and thus did not have a direct intergenerational link), and was not parameterized in the same manner so as to directly relate it to mobility decompositions.

E. Comparing Representative versus Subgroup Mobility Estimates

While we have shown which components of the decomposition play the largest roles in affecting both the levels and changes of full-population persistence measures, a separate question is how biased subgroup (e.g., white men) estimates are relative to representative estimates. If we had performed our family-income-to-childhood-income mobility estimation on, say, only white men, how biased (in levels and changes) would these estimates be relative to a representative sample?

In figure 8, we show how the mobility estimates change as we sequentially add various subgroups (as usual, the IGE is depicted in fig. 8*A* and the rank-rank in fig. 8*B*). We begin with white men (first series), the group most often studied in the existing mobility literature. In some decades, adding white women (second series) increases estimated persistence,

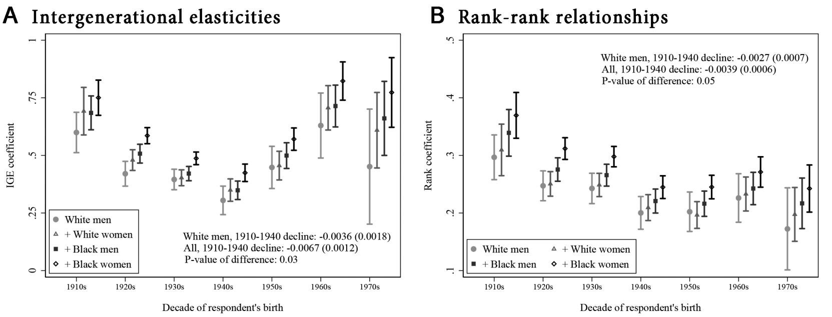

FIG. 8.—Mobility patterns during twentieth century, including underrepresented groups. The figure combines 15 different surveys, which are described in section II and in further detail in appendix E. Estimates are based on the baseline sample of respondents. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). We use sample weights where provided and further reweight each birth cohort (i.e., decade) so that they have representative race × sex shares. Each panel reports the decline between the 1910s–1940s cohorts from a specification that models the decline in the slope linearly. Specifically, we run regressions in which we interact predicted parental income (or rank) with a variable that measures the number of years between a respondent's birth and 1911 (including birth year fixed effects). The *p*-values correspond to a test of whether the two coefficients (using white men vs. representative samples) are equal using seemingly unrelated regressions.

and in other decades it reduces estimated persistence, but in all cases, confidence intervals overlap.

We then add Black respondents, first men (third series) and then women (fourth series). Both additions increase the estimated persistence measures, as we would expect from the evidence already presented. And again, as expected, the change tends to be larger once we add Black women. As they are born to families at the bottom of the distribution (like their male counterparts) and tend to remain poor as adults (more so than their male counterparts), excluding this group significantly biases measures of intergenerational persistence downward, despite being just over 5% of the population.[39]

In terms of the actual effects of using representative samples versus only white men on various point estimates, consider the 1920s cohort as an example. The white male rank-rank slope is 0.25 and does not change after adding white women. Adding Black men—just over 5% of the population—increases it an additional 3 percentage points to 0.28, and adding the similarly small group of Black women increases it to 0.31. Similarly, the IGE in this cohort rises from 0.42 for white men to 0.59 for the representative population. Excluding Black men and especially Black women paints an overly optimistic picture about the level of intergenerational mobility in the first half of the twentieth century.

Considering a representative population instead of only white men also changes our view of the evolution of mobility over this period. For white men, the IGE falls roughly 0.004 percentage points per year from 1910 to 1940 (table A.8). For the full population, it falls considerably faster over this period—0.007 points per year—and we can reject equality of these two rates at the 5% significance level. The analogous rates are 0.003 and 0.004 for the rank-rank and equality and can also be rejected at the 10% level. In summary, including only white men misses a substantial part of the decline in the slope and thus paints an overly pessimistic picture of the rise in intergenerational mobility over this same period.

VII. Discussion and Conclusion

We provide, to the best of our knowledge, the first evidence on long-run intergenerational relative mobility trends for representative samples of the US-born population. We find a robust decline in IGE and rank-rank persistence measures from the 1910s to the 1940s birth cohorts. Previous studies that have examined historical mobility have overwhelmingly focused on white men, which both overstates mobility relative to the full population—a point also made by Ward (2023) in the context of male-only mobility—but at the same time understates the rise in mobility from

[39] Figure D.6 shows that these patterns remain in the TS2SLS levels-based estimates.

the 1910s to the 1940s. Including only white men misses out on the important progress that Black Americans—particularly Black women—make relative to white individuals, which has large implications for full-population mobility given the extreme disadvantage of Black children over our sample period. In short, the United States starts the twentieth century much further from the American dream ideal of a mobile society but also improves more significantly when the full population is considered rather than only white men.

While we avoid comparing our 1910s–1970s survey data with the 1980s IRS data in levels (given that the latter data source has income information for both generations), we compare the relative positions of the four groups. Figure 9 considers individuals growing up at the 25th percentile of the income distribution, separately by race and sex for each birth cohort in our data. We include an additional data point from Chetty et al. (2020) labeled 1980s (though technically these individuals are born between 1978 and 1983). The results from the 1910s–1970s reflect findings we have already presented. For example, Black women are the poorest as adults but also show the most dramatic progress of any group. Similarly, Black-white convergence appears to peak around the 1940s (a brief moment where a Black boy and a white boy born at the 25th percentile would be predicted to end up at a similar family income rank as adults). In the 1960s and 1970s, Black-white adult income gaps for those born at the 25th percentile regain much of their earlier magnitude. The 1980s data show that Black women have continued their progress, in this case overtaking Black men (there is gender reversal among white individuals as well but much less dramatic both because the 1910s–1970s differences were always close to zero and because the female advantage in the 1980s is small). But overall, Black-white adult income gaps for those born at the 25th percentile continue to grow, continuing the trend we saw in the 1960s and 1970s in our data.

As Black-white convergence helped drive the rise in mobility over the first half of the twentieth century, it is natural to ask how important racial income gaps are today in shaping overall US intergenerational mobility. In particular, if modern-day income gaps between racial groups remain unchanged, how much would IGEs within racial groups have to fall for the overall US IGE to reach 0.20 (roughly that in Denmark; see Helsø 2021)? Using the decomposition in section VI and statistics from Chetty et al. (2020) on contemporary income distributions, we find that within-group IGEs would have to fall below 0.05, an implausibly high level of mobility.[40] Put differently, US intergenerational persistence faces a high lower bound unless major income convergence across racial groups occurs.

[40] To simplify this calculation, we assume that all racial groups would have the same within-group IGE. We use summary statistics from Chetty et al. (2020) to approximate modern

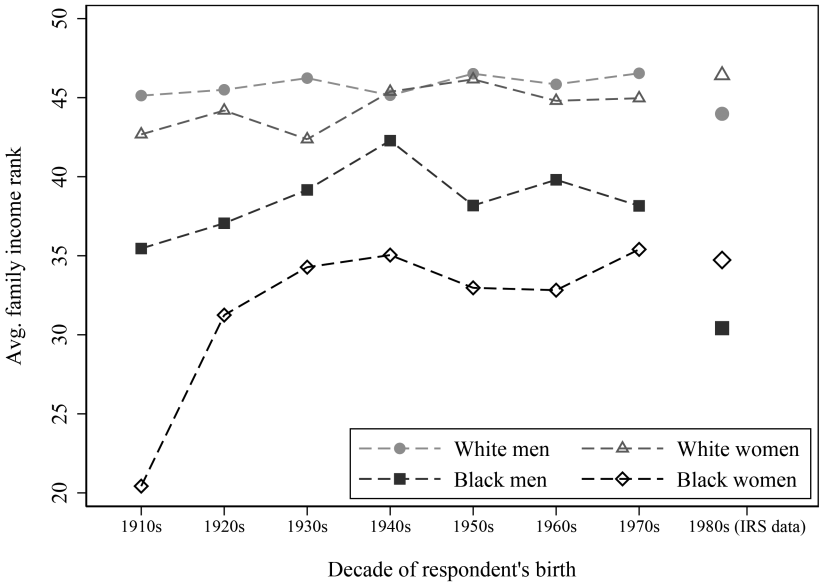

FIG. 9.—Average income rank of individuals born to 25th percentile of parental income distribution, by subgroup and birth cohort. Data for the 1910s–1970s birth cohorts combine 15 different surveys, which are described in section II and in further detail in appendix E. Data for the 1980 birth cohort are from Chetty et al. (2020; https://opportunityinsights.org/data/). The figure plots the average adult income rank for individuals growing up at the 25th percentile of the parental income distribution, separately by race, sex, and birth cohort. For survey respondents, we use equation (3) to compute the expected income rank for individuals growing up at the 25th percentile of the parental income distribution. To predict parental income, we use family income conditional on father's occupation, race, and region (South vs. elsewhere) from auxiliary data (often the census) as close as possible to the respondents' tenth birthday (see sec. III.B for more details). For the 1980 cohort, we use the average percentile rank in the national distribution of household income (measured in 2014–15) for individuals growing up at the 25th percentile of the parent household income distribution (measured in 1994–2000).

The comparison to modern data also suggest at least two areas for further work, both related to racial gaps given their centrality to overall mobility levels. First is that declining marriage rates and diverging outcomes by gender interact to produce changing patterns of mobility. The lower marriage rates of Black Americans relative to white individuals throughout our sample period (see table A.1) and continuing today permit large mobility gaps between men and women (as they are not married to each

within-racial-group income distributions. We find similar results using the rank-rank correlation and find that the between-group term accounts for over 25% of the overall rank-rank coefficient. For more details on these calculations, see app. sec. E.8.

other and thus do not mechanically share a family income). In addition to studying the implications of declining marriage rates for intergenerational mobility, future work might also examine the rise of interracial marriage—while rare during our sample period, today 18% of recently married Black Americans have a spouse of a different race.[41]

Second, any candidate explanation for the reversal of Black progress in closing the mobility gap with white individuals would need to have a large gender-specific component, given the relative progress Black women have made. Mass incarceration, a phenomenon that largely postdates our historical cohorts but has important implications for modern cohorts of Black men, and deindustrialization, which impacted Black men earlier than whites (see Wilson 1997), are two natural candidates.

We close with some final thoughts on what our paper suggests about the persistence of advantage across generations. On the one hand, the decline in intergenerational persistence over the first half of the twentieth century we document challenges scholarship that has concluded that intergenerational mobility remains relatively stable even in the face of large political and structural changes (see, e.g., Erikson and Goldthorpe 2002; Clark 2015, 2023; Olivetti and Paserman 2015; Ager, Boustan, and Eriksson 2019; Alesina et al. 2020; Song et al. 2020).[42] On the other hand, the return of early twentieth-century race-specific mobility gaps is cause for pessimism (perhaps suggesting that the midcentury convergence we document was a mere aberration).

Overall, we view the twentieth-century patterns as providing evidence that policy and institutions can increase US intergenerational mobility. The birth cohorts in our paper span the mechanization and declining importance of American agriculture, the high school movement, two World Wars, the Great Depression, the New Deal, the Great Compression, and the Civil Rights movement. Even the modern return of the race-specific mobility gaps present evidence of dynamism: Black women reversed a large gender gap that existed for at least seven decades. These documented changes across time suggest that mobility patterns are not set in stone and we hope will inspire future research to better understand the underlying institutional and policy determinants of intergenerational transmission of advantage.

[41] See https://www.pewresearch.org/social-trends/2017/05/18/1-trends-and-patterns-in-intermarriage.

[42] In contrast, recent studies of Scandinavian countries have also documented periods of rising mobility in the twentieth century following nationwide educational reforms (see, e.g., Pekkarinen, Salvanes, and Sarvimäki 2017; Karlson and Landersø 2021; Nybom and Stuhler 2023).

Data Availability

Code replicating the tables and figures in this article can be found in Jácome, Kuziemko, and Naidu (2024) in the Harvard Dataverse, https://doi.org/10.7910/DVN/DQECDA.

References

Aaronson, D., and B. Mazumder. 2008. "Intergenerational Economic Mobility in the United States, 1940 to 2000." *J. Human Res.* 43 (1): 139–72.

Abramitzky, R., L. P. Boustan, and K. Eriksson. 2012. "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *A.E.R.* 102 (5): 1832–56.

Abramitzky, R., L. P. Boustan, K. Eriksson, J. J. Feigenbaum, and S. Pérez. 2019. "Automated Linking of Historical Data." Working Paper no. 25825, NBER, Cambridge, MA.

Abramitzky, R., L. Boustan, E. Jácome, and S. Pérez. 2021. "Intergenerational Mobility of Immigrants in the United States over Two Centuries." *A.E.R.* 111 (2): 580–608.

Ager, P., L. P. Boustan, and K. Eriksson. 2019. "The Intergenerational Effects of a Large Wealth Shock: White Southerners after the Civil War." Working Paper no. 25700, NBER, Cambridge, MA.

Alesina, A. F., M. Seror, D. Y. Yang, Y. You, and W. Zeng. 2020. "Persistence despite Revolutions." Working Paper no. 27053, NBER, Cambridge, MA.

Althoff, L., H. B. Gray, and H. Reichardt. 2023. "The Missing Link: Women and Intergenerational Mobility." Working paper.

Bailey, M. J., C. Cole, M. Henderson, and C. Massey. 2020. "How Well Do Automated Linking Methods Perform? Lessons from U.S. Historical Data." *J. Econ. Literature* 58 (4): 997–1044.

Bailey, M. J., and P. Z. Lin. 2022. "Marital Matching and Women's Intergenerational Mobility in the Late 19th and Early 20th Century US." Working paper, Univ. California Los Angeles.

Becker, G., and N. Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *J.P.E.* 87 (6): 1153–89.

Berman, Y. 2018. "The Long Run Evolution of Absolute Intergenerational Mobility." Working paper.

———. 2022. "The Long-Run Evolution of Absolute Intergenerational Mobility." *American Econ. J. Appl. Econ.* 14 (3): 61–83.

Björklund, A., and M. Jäntti. 1997. "Intergenerational Income Mobility in Sweden Compared to the United States." *A.E.R.* 87 (5): 1009–18.

Bloise, F., P. Brunori, and P. Piraino. 2021. "Estimating Intergenerational Income Mobility on Sub-Optimal Data: A Machine Learning Approach." *J. Econ. Inequality* 19 (4): 643–65.

Bowles, S. 1972. "Schooling and Inequality from Generation to Generation." *J.P.E.* 80 (3): S219–S251.

Bratberg, E., J. Davis, B. Mazumder, M. Nybom, D. D. Schnitzlein, and K. Vaage. 2017. "A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US." *Scandinavian J. Econ.* 119 (1): 72–101.

Bratsberg, B., K. Røed, O. Raaum, et al. 2007. "Nonlinearities in Intergenerational Earnings Mobility: Consequences for Cross-Country Comparisons." *Econ. J.* 117 (519): C72–C92.

Buckles, K., J. Price, Z. Ward, and H. E. Wilbert. 2023. "Family Trees and Falling Apples: Historical Intergenerational Mobility Estimates for Women and Men." Working Paper no. 31918, NBER, Cambridge, MA.

Card, D., C. Domnisoru, and L. Taylor. 2018. "The Intergenerational Transmission of Human Capital: Evidence from the Golden Age of Upward Mobility." Working Paper no. 25000, NBER, Cambridge, MA.

Chetty, R., D. Grusky, M. Hell, N. Hendren, R. Manduca, and J. Narang. 2017. "The Fading American Dream: Trends in Absolute Income Mobility since 1940." *Science* 356 (6336): 398–406.

Chetty, R., and N. Hendren. 2018a. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *Q.J.E.* 133 (3): 1107–62.

———. 2018b. "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates." *Q.J.E.* 133 (3): 1163–228.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter. 2020. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *Q.J.E.* 135 (2): 711–83.

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." *Q.J.E.* 129 (4): 1553–623.

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." *A.E.R.* 104 (5): 141–47.

Choi, J., J. Gu, and S. Shen. 2018. "Weak-Instrument Robust Inference for Two-Sample Instrumental Variables Regression." *J. Appl. Econometrics* 33 (1): 109–25.

Clark, G. 2015. *The Son Also Rises: Surnames and the History of Social Mobility*. Princeton, NJ: Princeton Univ. Press.

———. 2023. "The Inheritance of Social Status: England, 1600 to 2022." *Proc. Nat. Acad. Sci. USA* 120 (27): e2300926120.

Collins, W. J., and M. H. Wanamaker. 2022. "African American Intergenerational Economic Mobility since 1880." *American Econ. J. Appl. Econ.* 14 (3): 84–117.

Connolly, M., M. Corak, and C. Haeck. 2019. "Intergenerational Mobility between and within Canada and the United States." *J. Labor Econ.* 37 (S2): S595–S641.

Dahl, G. B., and L. Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *A.E.R.* 102 (5): 1927–56.

Davis, J., and B. Mazumder. 2022. "The Decline in Intergenerational Mobility after 1980." Working paper.

Deutscher, N., and B. Mazumder. 2023. "Measuring Intergenerational Income Mobility: A Synthesis of Approaches." *J. Econ. Literature* 61 (3): 988–1036.

Erikson, R., and J. H. Goldthorpe. 2002. "Intergenerational Inequality: A Sociological Perspective." *J. Econ. Perspectives* 16 (3): 31–44.

Eriksson, K., G. Niemesh, M. Rashid, and J. Craig. 2023. "Marriage and the Intergenerational Mobility of Women: Evidence from Marriage Certificates 1850–1910." Working paper.

Farber, H. S., D. Herbst, I. Kuziemko, and S. Naidu. 2021. "Unions and Inequality over the Twentieth Century: New Evidence from Survey Data." *Q.J.E.* 136 (3): 1325–85.

Feigenbaum, J. J. 2015. "Intergenerational Mobility during the Great Depression." Working paper.

———. 2018. "Multiple Measures of Historical Intergenerational Mobility: Iowa 1915 to 1940." *Econ. J.* 128 (612): F446–F481.

Ferrie, J. P. 1996. "A New Sample of Males Linked from the Public Use Microdata Sample of the 1850 U.S. Federal Census of Population to the 1860 U.S. Federal Census Manuscript Schedules." *Hist. Methods* 29 (4): 141–56.

Fields, G. S., and E. A. Ok. 1999. "The Measurement of Income Mobility: An Introduction to the Literature." In *Handbook of Income Inequality Measurement*, edited by J. Silber, 557–98. Dordrecht: Springer.

Goldenweiser, E. A. 1916. "The Farmer's Income." *A.E.R.* 6 (1): 42–48.

Goldin, C., and L. F. Katz. 2010. *The Race between Education and Technology*. Cambridge, MA: Harvard Univ. Press.

Haider, S., and G. Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *A.E.R.* 96 (4): 1308–20.

Heckman, J., R. Pinto, and P. Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *A.E.R.* 103 (6): 2052–86.

Helsø, A.-L. 2021. "Intergenerational Income Mobility in Denmark and the United States." *Scandinavian J. Econ.* 123 (2): 508–31.

Hertz, T. 2008. "A Group-Specific Measure of Intergenerational Persistence." *Econ. Letters* 100 (3): 415–17.

Hilger, N. G. 2015. "The Great Escape: Intergenerational Mobility in the United States since 1940." Working Paper no. 21217, NBER, Cambridge, MA.

Inoue, A., and G. Solon. 2010. "Two-Sample Instrumental Variables Estimators." *Rev. Econ. and Statis.* 92 (3): 557–61.

Jácome, E., I. Kuziemko, and S. Naidu. 2024. "Replication Data for: 'Mobility for All: Representative Intergenerational Mobility Estimates over the Twentieth Century." Harvard Dataverse, https://doi.org/10.7910/DVN/DQECDA.

Jantti, M., B. Bratsberg, K. Roed, et al. 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 Discussion Paper no. 1938, Inst. Labor Econ., Bonn.

Karlson, K., and R. Landersø. 2021. "The Making and Unmaking of Opportunity: Educational Mobility in 20th Century-Denmark." IZA Discussion Paper no. 14135, Inst. Labor Econ., Bonn.

Landersø, R., and J. J. Heckman. 2017. "The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US." *Scandinavian J. Econ.* 119 (1): 178–230.

Lee, C.-I., and G. Solon. 2009. "Trends in Intergenerational Income Mobility." *Rev. Econ. and Statis.* 91 (4): 766–72.

Lemieux, T. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *A.E.R.* 96 (3): 461–98.

Little, R. J., and D. B. Rubin. 2019. *Statistical Analysis with Missing Data*. Vol. 793. New York: Wiley.

Løken, K. V., M. Mogstad, and M. Wiswall. 2012. "What Linear Estimators Miss: The Effects of Family Income on Child Outcomes." *American Econ. J. Applied Econ.* 4 (2): 1–35.

Long, J., and J. Ferrie. 2013. "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *A.E.R.* 103 (4): 1109–37.

Manduca, R., M. Hell, A. Adermon, et al. 2024. "Measuring Absolute Income Mobility: Lessons from North America and Europe." *American Econ. J. Appl. Econ.* 16 (2): 1–30.

Margo, R. A. 2016. "Obama, Katrina, and the Persistence of Racial Inequality." *J. Econ. Hist.* 76 (2): 301–41.

Massey, C., and J. Rothbaum. 2020. "The Geography of Opportunity over Time."

Mazumder, B. 2015. "Estimating the Intergenerational Elasticity and Rank Association in the Us: Overcoming the Current Limitations of Tax Data."

———. 2018. "Intergenerational Mobility in the United States: What We Have Learned from the PSID." *Ann. American Acad. Polit. and Soc. Sci.* 680 (1): 213–34.

Nybom, M., and J. Stuhler. 2016. "Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation." *J. Human Res.* 51 (1): 239–68.

———. 2023. "Interpreting Trends in Intergenerational Mobility."

O'Hare, W. P. 2019. "Census Coverage of the Black Population." In *Differential Undercounts in the US Census*, edited by W. P. O'Hare, 83–91. Cham: Springer.

Olivetti, C., and M. D. Paserman. 2015. "In the Name of the Son (and the Daughter): Intergenerational Mobility in the United States, 1850–1940." *A.E.R.* 105 (8): 2695–724.

Pacini, D., and F. Windmeijer. 2016. "Robust Inference for the Two-Sample 2SLS Estimator." *Econ. Letters* 146:50–54.

Pekkarinen, T., K. G. Salvanes, and M. Sarvimäki. 2017. "The Evolution of Social Mobility: Norway during the Twentieth Century." *Scandinavian J. Econ.* 119 (1): 5–33.

Ray, D., and G. Genicot. 2023. "Measuring Upward Mobility." *A.E.R.* 113 (11): 3044–89.

Rubin, D. B. 1987. *Multiple Imputation for Nonresponse in Surveys.* New York: Wiley.

Ruggles, S., S. Flood, S. Foster, et al. 2021. *Integrated Public Use Microdata Series: Version 11.0 [dataset].* Minneapolis, MN: IPUMS.

Saavedra, M., and T. Twinam. 2020. "A Machine Learning Approach to Improving Occupational Income Scores." *Explorations Econ. Hist.* 75:101304.

Schoeni, R. F., and E. E. Wiemers. 2015. "The Implications of Selective Attrition for Estimates of Intergenerational Elasticity of Family Income." *J. Econ. Inequality* 13 (3): 351–72.

Solon, G. 1992. "Intergenerational Income Mobility in the United States." *A.E.R.* 82 (3): 393–408.

Song, X., C. G. Massey, K. A. Rolf, J. P. Ferrie, J. L. Rothbaum, and Y. Xie. 2020. "Long-Term Decline in Intergenerational Mobility in the United States since the 1850s." *Proc. Nat. Acad. Sci. USA* 117 (1): 251–58.

Uguccioni, J. 2021. "Is Timing Everything? Parental Unemployment and Their Children's Future Earnings." Working paper, Univ. Toronto.

Ward, Z. 2023. "Intergenerational Mobility in American History: Accounting for Race and Measurement Error." *A.E.R.* 113 (12): 3213–48.

Wilson, W. J. 1997. *When Work Disappears: The World of the New Urban Poor.* New York: Knopf.

Zimmerman, D. 1992. "Regression toward Mediocrity in Economic Stature." *A.E.R.* 82 (3): 409–29.