Intergenerational Mobility in American History: Accounting for Race and Measurement Error*

Zachary Ward†

Baylor University

November 2019

Abstract: A large body of evidence suggests that intergenerational mobility in the United States has declined over the past 150 years. However, research that finds high relative mobility in America's past is based on data with few or no black families, and therefore does not account for the limited opportunities available for African Americans. Moreover, historical studies often measure the father's economic status with error, which biases estimates towards greater mobility. Using new early 20th century data, I show that the persistence of economic status from father to son is over twice as strong after accounting for racial disparities and for measurement error. After addressing these two issues, I estimate that relative mobility has increased over the 20th century. The results imply that there is greater equality of opportunity today than in the early 20th century, mostly because opportunity was never that equal.

Keywords: intergenerational mobility, measurement error, persistence

JEL Codes: J62, N31, N32

*Thanks to seminar participants at Baylor University and the University of Michigan. I also thank conference participants at the 2019 Economic Demography Workshop and 2019 Economic History Association Annual Meeting. I have benefitted from comments from Martha Bailey, Hoyt Bleakley, Leah Boustan, Connor Cole, Bill Collins, James Fenske, Andy Ferrara, Elisa Jácome, Ed Kosack, Elaine Liu, Amber McKinney, Shariq Mohammed, Greg Niemesh, Santiago Pérez, Steve Trejo and Marianne Wanamaker. Thanks to Katherine Eriksson for providing help for testing the robustness to alternative linking methods, and to Lee Alston for helping me to gain access to the historical census files. All errors are my own.

†email: Zach.A.Ward@gmail.com. Department of Economics, 1621 S 3rd St, Hankamer School of Business, Baylor University, Waco, TX, 76710.

I. Introduction

One of the reasons why America tolerates inequality is the belief that opportunity is available for everyone, whether they grew up rich or poor. However, a large body of research documents a decline in intergenerational mobility since the mid-19th century, suggesting that the United States is turning into a rigid and class-based society (e.g., Ferrie, 2005; Long and Ferrie, 2013a; Feigenbaum, 2018; Parman, 2011; Song et al., 2019).[1] Yet such a decrease in mobility may be surprising given the major institutional changes over the 20th century that have aimed to curb inequality and improve opportunity, such as the extension of civil rights protections, desegregation of schools, and the rise of a progressive income tax system. Was there truly greater equality of opportunity in the past?

In this paper, I challenge the idea that opportunity was more equal in America's past. One reason why others find high relative mobility is that they include few, if any, African Americans in the data. While this may seem odd, it is due to data limitations: since research often starts to measure mobility in 1850 – before emancipation – most black families are unobservable (e.g., Ferrie, 2005; Long and Ferrie 2013; Olivetti and Paserman, 2015; Song et al., 2019).[2] Others take advantage of income data from Iowa; however, Iowa was over 99 percent white in the early 20th century (e.g., Feigenbaum, 2018; Parman, 2011).[3] Since most black families are not in the historical data, studies also drop black families in later decades (including well after emancipation) in order to make comparisons over time consistent; therefore, the documented decline in relative mobility is actually a decline in *white* mobility. Yet, it is well known that black families remained at the bottom of the income distribution for decades following the Civil War, which suggests that historical opportunity was not that equal for the overall population (Collins and Wanamaker, 2017; Hertz, 2005; Margo, 2016).

[1] The evidence for the decline comes from comparing the same mobility measure over time, such as the Altham statistic for occupational mobility (Ferrie, 2005; Long and Ferrie, 2013), the IGE or rank-rank slope for income mobility (Feigenbaum, 2018; Parman, 2011), or the rank-rank slope for an occupational score (Song et al., 2019). However, there is debate over the trend in relative mobility over the 19th and 20th centuries (see Hout and Guest, 2013). Chetty et al. (2017) show a downward trend in "absolute income mobility" from the mid to late 20th century, which differs from relative mobility.

[2] Those who were emancipated are not linkable, partially because they were not enumerated by name in 1850 or 1860. Also see Pérez (2019) for comparative mobility estimates across countries in the 19th century.

[3] Feigenbaum (2015) also estimates income mobility by linking a 1918-1919 Bureau of Labor Statistics survey to the 1940 Census. However, he drops black families since they are not geographically representative of the population.

A second reason why prior work overstates historical mobility is measurement error. Historical studies often use a single observation of the father's occupation or income to proxy for his lifetime status, which can be problematic if there are transitory shocks (e.g., Solon, 1992). While such error may not be important for occupational-based measures (which are more common in historical research), error could exist due to intragenerational occupational mobility or mistakes made by respondents, enumerators or digitizers (Kambourov and Manovskii, 2008; Mazumder and Acosta, 2015; Ward, 2019a).[4] Regardless of where errors come from, they cause the estimated association between the father and son to appear weaker than it truly is, falsely implying high mobility (Clark, 2014). A standard way to address this problem is to use multiple father observations to better proxy for his permanent status; however, this approach is seldom used due to the high cost of linking historical censuses prior to 1940 (see Ward (2019a) for an exception).[5] Therefore, it could be that the estimated decline in mobility is partially due to a decline in measurement error.

Using a new linked sample from the early 20th century, I show that estimates of relative mobility are completely revised after accounting for race and measurement error (see Figure 1). (Note that I focus on *relative* mobility estimates and not *absolute* mobility estimates, like those in Chetty et al. (2017).[6]) After imputing income by occupation, race and region, I show with 1910-1940 linked census data that adding black families to a white sample increases the intergenerational elasticity coefficient (or IGE) by 44 percent. This revised estimate implies that instead of roughly one-third of economic gaps transmitting from father to son (0.37), about one-half did (0.53). Second, after linking the 1910 fathers to themselves in the 1900 and 1920 censuses, I show that going from one snapshot of the father to averaging three observations increases the elasticity by 33 percent (from 0.53 to 0.71).[7] (I use the same sample when using one or three father observations, so the increase in IGE is not due to a change in sample composition.) With this fix

[4] See Bailey et al. (2019) for a discussion of a different form of measurement error due to false links.

[5] Ward (2019a) shows that measurement error is important for understanding why ethnic occupational gaps converged slowly during the Age of Mass Migration. I expand upon this result by estimating the importance of measurement error for the entire population, for a variety of mobility measures, and testing the classical measurement error assumption. Also see Ferrie et al. (2016) for the importance of error in the education variable.

[6] Absolute income mobility focuses on the growth in economic status across generations, while relative mobility focuses on how inequality is transmitted across generations.

[7] I only use multiple observations of the father and not the son since classical measurement error in the son's outcome does not bias the IGE. However, non-classical measurement error might in the form of life-cycle bias in the form of life-cycle bias (Haider and Solon, 2006; Grawe, 2006; Nybom and Stuhler, 2017). I measure the son's outcomes at the midpoint of his lifecycle to minimize life-cycle bias.

for measurement error, the results imply that seventh-tenths of economic gaps across families transmitted from father to son, instead of only one-third – in short, relative mobility was considerably lower than previously estimated with linked father-son data.

Moreover, averaging three father observations may still not perfectly capture his permanent economic status. A simple assumption is that the data are subject to classical measurement error, which is surprisingly consistent with patterns in the data. Given classical measurement error, eliminating noise leads to a "true" father-son elasticity of 0.84, which is 2.3 times higher than the baseline estimate of 0.37. An alternative fix for measurement error is to instrument one father observation with a second (Altonji and Dunn, 1991; Modalsli and Vosters, 2019), which yields a similar IGE estimate of 0.82. If one iterates the 0.82 coefficient across generations, then initial gaps take ten generations to converge to about 15 percent – eight generations or about 200 years longer than the estimate with one father observation and white families of 0.37.[8]

The main limitation of the results is that, due to a lack of information on income prior to 1940, economic gaps across families are measured with *imputed* income rather than *actual* income. Therefore, one should not directly compare the historical IGEs (0.82-0.84) to modern-day estimates (~0.50). Another limitation is linking error: a potential reason why I find that the father's occupation is poorly correlated across censuses is due to false positives (Bailey et al., 2019). To reduce false positives, I use conservative linking methods from Feigenbaum (2016) and Abramitzky et al., (2012).[9] I also show that the father's occupation is poorly correlated across censuses even when limiting my sample to the highest quality links, suggesting that linking error is not driving the result that measurement error matters.

I finish by re-estimating the trend of intergenerational elasticity over the 20th century and find, in contrast to recent historical work, that relative mobility has increased over time. Specifically, I compare my early 20th century estimates to estimates from 1968-1997 based on the PSID. To make the comparison consistent over time, I impute income by occupation, race and location, include black and white families, and use multiple observations of the father. Specifically,

[8] That is 0.84^{11}=0.15, 0.82^{10}=0.14 and 0.37^{2}=0.14. Extrapolating a one-generation IGE in such a way is problematic (Stuhler, 2012), but the point remains that accounting for racial disparities and measurement error substantially revises our understanding of relative mobility in America's past.

[9] For the Feigenbaum (2016) method, I increase the meta-parameter that sets the precision of the model in order to reduce false positives. For the Abramitzky et al. (2019) method, I only keep links that are unique and exact within a 5-year birth window.

I estimate that the imputed-income IGE has fallen from about 0.82 to 0.57.[10] This result only appears after accounting for race and measurement error; otherwise, I would have estimated an increase in white persistence, similar to others using linked father-son data (e.g., Feigenbaum, 2018; Long and Ferrie, 2013). I also find a decrease in persistence when using rank-rank measures or limiting the sample to black or white families; for example, the white IGE is estimated to drop from 0.69 to 0.51. Overall, the results suggest that there is greater equality of opportunity today than in the early 20th century, mostly because opportunity was never that equal.

Low mobility in the early 20th century is consistent with the "Great Gatsby" curve since the early 20th century also had high levels of wealth and income inequality (Corak, 2013; Goldin and Katz, 2008; Piketty and Saez, 2003; Saez and Zucman, 2016). Future linked studies may find that relative mobility improved following the compression of the income distribution between 1940 and 1950 (Aaronson and Mazumder, 2008; Goldin and Margo, 1992). At the same time, note that the long-run trend in relative mobility may differ from the long-run trend in absolute mobility. For instance, Chetty et al. (2017) document that the fraction of children who earn more than their parents has been declining for birth cohorts since the 1940s. Therefore, while opportunity may be more equal today than in the past, the growth in income across generations may have slowed down.

The results contribute to the fast-growing literature on historical mobility.[11] In particular, the results raise the possibility that measurement error varies across source, time and space, which could then bias inference in comparative mobility research. For example, I show that the historical IGE was more strongly attenuated than the modern one, perhaps because the early 20th century data are of lower quality than the PSID or because there was substantially higher intragenerational mobility in the past. I also show that if one does not address measurement error in both datasets, then one would incorrectly estimate the trend in mobility. These results suggest that studies that compare mobility across historical datasets should try to account for variation in measurement error with multiple father observations.

[10] The (weighted) black share of the data is similar over the early 20th century (0.099) and the late 20th century (0.094), suggesting that the relative share of the population is not driving the change in relative mobility.

[11] For examples of historical mobility research using linked data, see Abramitzky et al. (2019), Ager et al., (2019), Bailey et al. (2019), Collins and Wanamaker (2017), Connor (2018), Craig et al. (2019), Dupraz and Ferrara (2019), Feigenbaum (2015), Feigenbaum (2018), Ferrie (2005), Grusky (1986), Guest et al., (1989), Kosack and Ward (2019), Karbownik and Wray (2019), Long and Ferrie (2013), Long and Ferrie (2018), Modalsli (2017), Pérez (2017), Pérez (2019), Song et al. (2019), Tan (2018), Ward (2019a) and Ward (2019b).

Finally, the results partially reconcile a tension between mobility estimates from linked father-son data and estimates from rare surname data (Clark, 2014). Clark argues that the true level of mobility is between 0.7 and 0.8, which aligns with some estimates from this paper.[12] The results in this paper agree with Clark's argument that linked father-son studies suffer from attenuation bias. However, while Clark claims that error arises because observables, such as occupational status, fail to capture one's underlying level of social competence, I estimate low mobility when using multiple occupation observations. Therefore, it may be that part of the difference in persistence rates across Clark and earlier linked work was due to measurement error from data entry or transitory shocks, rather than due to multiple measures failing to capture underlying social competence.[13] At the same time, I also find that relative mobility improved over the 20th century, which contrasts with Clark's argument that mobility is constant over time.

II. Measuring intergenerational mobility

There are many ways to measure intergenerational mobility, but I primarily focus on the intergenerational elasticity coefficient (IGE). The standard IGE comes from regressing the son's log permanent income $(y_{i,s})$ on the father's log permanent income $(y_{i,f})$:

$$y_{i,s} = \beta_0 + \beta_1 y_{i,f} + \varepsilon_{i,s} \tag{1}$$

The coefficient of interest β_1 predicts how much of an income gap across two fathers persists to their sons. Thus, a high-mobility economy has a β_1 near zero while a low-mobility economy has a β_1 near one.

While the above regression should be for permanent income, permanent income is often unobserved. Instead, many in the early literature used only one income observation to proxy for permanent income; however, as is now well known today, one snapshot poorly measures permanent income due to transitory fluctuations (e.g., Solon, 1992; Mazumder, 2005). Such error attenuates β_1 toward zero and falsely implies high mobility. Under the assumption of classical measurement error where the parent's income varies from permanent income by random noise

[12] The point estimates for the IGE are lower than found in Olivetti and Paserman (2015), who use a grouped estimated based on one's first name. Theoretically, measurement error should be averaged out for their estimates as well. However, Olivetti and Paserman argue that their methodology is better suited for estimating the trend rather than the level of mobility due to other effects of one's first name on outcomes.

[13] There are other reasons why mobility estimates differ across grouped and linked datasets (Olivetti and Paserman, 2015; Torche and Corvalan, 2018).

$(y_{i,f} = y^*_{i,f} + v_{i,f})$, then attenuation bias falls when averaging the father's income more times (T):

$$plim\ \widehat{\beta_{avg}} = \beta_1 \frac{var(y^*_{i,f})}{var(y^*_{i,f}) + {var(v_{i,f})}\big/{T}} \tag{2}$$

Modern-day studies with annual data often use long-run averages of ten or fifteen father observations; however, historical data rarely go beyond $T = 1$ due to the high costs of obtaining more information via linking censuses. Note that this model focuses on measurement error in the father's outcome and not the son since classical error in the son's income does not bias estimates of the IGE.[14,15] However, other measures of mobility (e.g., the intergenerational correlation, rank correlation or transition matrices) are influenced by error in the son's outcome (Nybom and Stuhler, 2017).

Since this model is framed in the context of transitory income shocks, measurement error may not be a problem for occupational categories or imputed income – the most common measures of "permanent status" in historical data. If transitory shocks do not cause people to switch occupations (or self-report a change in occupation), then it may be that one observation of the father's occupation does well to capture permanent status. However, the empirical evidence shows that attenuation bias is important for occupational-based measures in both modern-day and historical data (Mazumder and Acosta, 2015; Ward, 2019a). Instead of transitory shocks, measurement error could also come from errors in the data, such as from data entry. For example, this type of error was found in the PSID due to inconsistent coding of occupations (Kambourov and Manovskii, 2008) and in the 1940 census for education (Ferrie et al., 2016).

III. Data

To test how mobility estimates change when accounting for race and measurement error, I need linked father-son data that include African Americans and have multiple observations of the father's occupation. The data structure combines three different links: one of the child in the 1910

[14] See Haider and Solon (2006) and Nybom and Stuhler (2017) on non-classical measurement error based on the point of the son's life-cycle.

[15] See Mazumder (2005) for a model where the error term is correlated across subsequent observations. Since my data are from decennial censuses, I am less concerned about auto-correlated measurement error. For example, Haider (2001) finds that less than 15 percent of a transitory income shock persists after three years.

census to his adult outcome in the 1940 census, one of the father in 1910 to himself in the 1900 census, and one of the father in 1910 to himself in the 1920 census. Importantly, I link both black and white individuals across all censuses. However, the data do not include Asians or Native Americans (who make up about 0.8 percent of the 1940 population). The data also do not contain females since they may change surname between censuses.

Each link (1910-1940; 1910-1900; 1910-1920) is made mostly following the machine-learning method described by Feigenbaum (2016).[16] For example, with the 1910-1940 link, a sample of 2,000 US-born white and 2,000 US-born black children aged 0 to 14 are drawn from the 1910 census. These children are hand-linked to the 1940 census based on first name, last name, race, year of birth and birthplace. Then the hand-linking process is modeled with probits (separately by race) to predict the best link, and the resulting model is then applied to the full link between the 1910 and 1940 censuses. I follow the same methodology for the fathers when linking them to the 1920 or 1900 censuses (see Appendix B). Note that the results are robust to using an alternative iterative strategy where I keep unique links in a 5-year-of-birth window (i.e., plus or minus two years; see Abramitzky et al., 2019).

The data structure leads to sons being in the middle of their lifecycle in 1940 (i.e., 30-44 years old). Capturing sons at the midpoint is important for reducing measurement error based on the son's age or "lifecycle" bias (Grawe, 2006; Haider and Solon, 2006). I further restrict the sample so that fathers are observed throughout the main part of their lifecycle (i.e., between 20-37 in 1900, 30-47 in 1910 or 40-57 in 1920). I also only keep father-son links that have occupation observations in all years.[17] After these restrictions, the average age of the father is 38.2, and the average age of the son is 36.6. The final sample contains 394,864 sons linked to 326,359 fathers.

The benefits of these data come with the costs of having a select sample of sons and fathers. The final sample is 4.6 percent of the possible children to link from 1910 to 1940. The low linking rate is due to the sample being triple-linked. Since each census link captures about 25 percent of

[16] There are a few departures from the Feigenbaum (2016) method to make linking more feasible between two complete-count censuses. Primarily, I additionally block on the first letter of the last name, the first letter of the first name and race. See Appendix B. Some of these links were already created in prior work (e.g., the 1910-1940 link in Kosack and Ward (2018) and the 1910-1920 link in Ward (2019a)).

[17] The 1900 and 1920 preliminary full-count censuses have occupations which are not yet classified. I clean these codes by assigning them the most common occupation code based on their occupation string. Specifically, I use the most common code in the 1920 Census and then in the 1930 Census. The results are not sensitive to cleaning these codes or dropping them from the data.

the linkable population, successfully finding people across all four censuses is less likely.[18] Since a successful link is not random but depends on the uniqueness of one's name, race, age and birth state combination, the sample is not be representative of the population (see Appendix Table B5). To address selection into the sample, I reweight it to match the 1940 population's characteristics on age, high school education, race, and census region.[19] The most important part of the weighting process is that black sons are given five times the weight of white sons since the linking rate is much lower for African Americans. While the representativeness of the sample is concerning, it does not affect the argument that accounting for race and measurement error matter for IGE estimates since I always use the same sample. However, an unrepresentative sample could bias the trend in IGE over time; however, I find large differences between the early and late 20th century IGE (0.82 v. 0.57) that the trend is unlikely to be due to a change in representativeness.

Imputed Income.

I prefer to estimate intergenerational income mobility; however, income data are not available. Therefore, I rely on imputed income measures. A commonly used measure is the *occscore* variable from IPUMS, which is the median income in the 1950 Census by 3-digit occupational codes. (I often refer to this score as the "1950 occupational score.") My preferred way to impute income is based on Collins and Wanamaker (2017) where income not only varies by occupation, but also by race and region of residence. That is, I impute income based on the average wage income for wage workers in the 1940 census in each occupation, race and region cell (the "1940 income score"). Self-employed earnings are also imputed using information from the 1960 Census.[20]

The 1940 income score has several advantages over the 1950 occupational score. Differentiating income by race and region addresses the significant racial and regional income gaps in the early 20th century (Margo, 2016; Mitchener and McLean, 1999). Further, the 1940 score imputes perquisites for farmers and farm laborers, comes before the compression of the

[18] One reason why the linking rate is low is that I set conservative parameters to reduce false positives. I set the tuning parameters such that 10 percent of the linked dataset is expected to be a false positive, as estimated when fitting the probit model to the training data. False positives would lead to additional measurement error in my mobility estimates. However, it appears that false links from the Feigenbaum method do not strongly affect the IGE since the false links have similar characteristics as the true links (Bailey et al., 2019).
[19] The qualitative results are similar when using inverse proportional weights (Bailey et al., 2019).
[20] See Appendix C of Kosack and Ward (2018) for further detail on the creation of the income score.

income distribution between 1940 and 1950, and allows the score to vary by farmer owners and tenants (based on the home ownership variable).[21] This last reason allows me to capture important gaps within farmers that are ignored by the 1950 occupational score. (For example, this income score places farm owners at the 39th percentile in 1910 and tenant farmers at the 23rd percentile, on average). I also show how alternative measures of status besides these 1940 and 1950 scores are influenced by racial disparities and measurement error. Yet since I am missing key information on productivity within occupation, one should not compare income score mobility to income mobility (Inwood et al., 2019; Saavedra and Twinam, 2019).

The first descriptive statistics

The first key statistic from this new sample, shown in Table 1, is that the black-white gap in income scores was large and persisted from father to son. Specifically, black income scores were about 80 log points (or 56 percent) lower than white income scores – for both fathers and sons. Similarly, the gap in percentile ranks remained at about 45 percentiles. Since black sons ended at a low rank like their fathers, there was little upward rank mobility for this group, as shown in detail by Collins and Wanamaker (2017).

Table 1 also shows that my preferred income score captures the actual black-white income gap much better than occupational scores. For instance, Margo (2016) estimates that the actual income gap was 62 percent in 1940. My preferred score estimates a 56 percent gap, while the 1950 occupational score measures a gap of 28 percent. Therefore, racial gaps in income are muted when using only occupation to impute status.

The second key statistic from the sample is that the father's occupation was weakly correlated across censuses (see Table 2). For example, the correlation of the log income score in 1910 and log income score in 1920 is 0.67 (after removing life-cycle effects).[22] This weak correlation may be partially due to false positives, but limiting the sample the top 20 percent of linking scores also leads to a weak correlation of 0.70. As another measure, only 48 percent of fathers in 1910 had the same (3-digit) occupation in 1920.[23] If one snapshot accurately captured

[21] Farmer and farm laborer perquisites are imputed based on information from a 1939 USDA report, as discussed by Collins and Wanamaker (2017).
[22] I remove life-cycle effects for both the father and the sons after controlling for a quartic in age.
[23] When using the Song et al. (2019) percentile rank measure, the 1910-1920 correlation is similarly low at 0.51.

the father's long-run status, then these correlations should be closer to one. While my focus is on the early 20th century, these low correlations are not limited to this period. If one uses the 1870-1880 IPUMS Linked Representative Sample, then the correlation of 1950 log occupational income was 0.62 – close to the 0.59 estimate in my 1910-1920 data.[24] Given these low correlations, studies which use one snapshot of the father likely overstate mobility throughout American history.

IV. Estimating mobility when accounting for race

In this section, I show that accounting for racial disparities strongly influences estimates of relative mobility. That is, I show how the IGE changes when going from a sample of only white families to a sample of black and white families. While not listed in Equation (1), I account for life-cycle effects with a quartic in age, though the results are not sensitive to this control.[25]

The IGE for the common historical sample of one father snapshot and white families is 0.37 (see Table 3). This result suggests that about one-third of an income score gap between two white fathers is expected to persist to their sons. The elasticity when restricting the sample to black families is also low (0.30), indicating higher relative mobility for black families than for white families (0.37). These estimates are lower than modern-day actual income IGEs at around 0.50. Therefore, a naïve comparison over time suggests that mobility was higher 100 years ago. However, this comparison should not be made: besides the fact that modern-day studies use actual income and not imputed income, modern-day estimates also include black and white families and address attenuation bias with multiple observations of the father's outcome.

Low within-race elasticities do not imply between-race convergence. When one plots the IGEs for black and white families, as in Figure 2, there is a wide mobility gap. Conditional on the father's income score, black sons are estimated to end up with 53 log-point (or 41 percent) lower income scores than white sons. This result suggests that the lack of racial convergence is not due

[24] To maintain comparability with the early 20th century sample, I limit the IPUMS Linked Representative Sample to 30-47 year old males in 1870 who have an occupation in both 1870 and 1880. I use the sample that links males.

[25] That is, I first regress the father's outcome on a quartic of the father's age and son's age, and then use the residuals. Similarly, I regress the son's outcome on a quartic of the father's age and son's age. Point estimates are not that sensitive to age controls. Note that while the recent literature has used rank-based measures partially because they satisfy the linear specification (e.g., Chetty et al., 2014), the log-log association for my data does appear to be linear (see Figure A1). The rank-rank association also appears to be linear (see Figure A2).

to an inheritance of poverty but due to other factors outside of the father's status, a pattern which is discussed in more detail by Collins and Wanamaker (2017).

Since black sons ended lower than white sons conditional on parental status, the white-only elasticity fails to measure equality of opportunity for the whole population. Pooling black and white families together increases the population-level IGE from 0.37 to 0.53 – an increase of 44 percent. Intuitively, since most black fathers and sons were low in the distribution, adding black families drops the white log-log intercept and steepens the slope. The increase of 44 percent is a large change. For example, Olivetti and Paserman's (2015) result that intergenerational persistence increased between 1870 and 1940 is due to a smaller 27 percent increase in the IGE. Therefore, adding black families to the sample changes our understanding of intergenerational mobility for the overall population in the early 20th century. Of course, the IGEs that are separately estimated by race more accurately predict the convergence of economic gaps across families, but the standard way to describe a country's overall rate of mobility, like the modern-day 0.47 estimate from Corak (2013), pools all races and ethnicities together.

The large increase to the IGE after adding black families depends on using a score that allows income to vary by race. If one uses a score where everyone in a given occupation earns the same amount, such as the 1950 occupational score, then the IGE increases by far less after adding black families (44 percent v. 8 percent increase, see Panel B of Table 3). To determine how much racial and regional disparities within occupation matter, I adjust the 1950 occupational score to reflect these gaps.[26] After making racial and regional adjustments, the IGE increases by 36 percent when adding black families (0.36 to 0.49) (see Panel C). Therefore, it is not only important to include black families in the sample, but the income score should also capture the actual black-white gap in income.

Rather than using IGEs to measure intergenerational mobility, one could instead use rank-rank measures (e.g., Chetty et al., 2014). After percentile ranking the fathers in 1910 and the sons in 1940 (within each of their birth cohorts), the rank-rank slope increases when going from a white-

[26] To adjust for race and region, I take the 1940 income score and calculate how the average income in each occupation/race/region cell differs from the average for the national occupation cell. For example, white blacksmiths in the northeast earned 115 percent of the national average for blacksmiths. I then multiply the *occscore* variable by this percent. I similarly adjust the 1950 score to differentiate farm owners from farm tenants, since this was also done in the 1940 score.

only sample to a pooled black and white sample (from 0.39 to 0.50). Therefore, no matter whether one prefers a log-log or percentile rank specification, equality of opportunity is substantially mismeasured if one discounts the limited opportunities available for African Americans.

V. Estimating mobility when accounting for measurement error

Time-averaged estimates of mobility

In this section, I turn from the importance of racial disparities to the importance of measurement error. So far, I have only estimated the IGE with a single-linked dataset from 1910 to 1940, where the father is observed only once. Now I show the importance of measurement error by estimating how the IGE changes from using a single snapshot to averaging the father's outcome from the 1900, 1910 and 1920 Censuses.

For the preferred 1940 income score, going from one to an average of two father observations increases the IGE from 0.53 to 0.65, or by 23 percent (see Figure 3). This result is exactly as expected given measurement error in the father's income score. An increase in the IGE when going from one to an average of two father observations is also consistent with prior evidence for a sample of immigrant descendants in the early 20th century (Ward, 2019a).

If one goes further and uses the average of three father observations, the IGE increases from 0.65 to 0.71.[27] Since it is commonly thought that transitory fluctuations in occupation are not that strong (e.g., Zimmerman, 1992), the fact that the IGE further changes may indicate that measurement error is due to data quality issues, such as errors in reporting, enumeration, or digitization. While it is unclear where the error comes from, the estimate suggests that instead of one-third of initial gaps transmitting from father to son (as was estimated with the one-father white-only IGE), seven-tenths did – an increase of 92 percent. For context on the size of this change, Corak (2013) reports that the difference in IGEs between the modern-day Sweden and United States is 74 percent (0.27 v. 0.47).

The influence of measurement error does not just apply to the pooled IGE, but also if one limits the sample to white or black families (see Panels B and C of Table 4). For instance, the white IGE using three father observations is 0.55, which is 47 percent higher than the one-observation

[27] If one instead uses the maximum log income score between 1900, 1910 and 1920, the IGE is 0.61.

estimate of 0.37. The black IGE also increases, from 0.30 to 0.52 – a larger increase of 72 percent, perhaps due to more error when recording black occupations or higher intragenerational mobility. Either way, multiple father observations also lead to higher persistence estimates within race.

In contrast to the IGE, rank-rank slopes appear to be less affected by measurement error in the father's outcome (see right-hand columns of Table 4). The rank-rank association for the 1940 income score increases by only 11 percent when averaging three father observations (0.50 to 0.56), less than the 44 percent increase for the IGE. (Note that I average the father's score and then rank them.) A key reason why measurement error attenuates the IGE is that measurement error adds extra "bad" variation to the scores. "Bad" variation is not added to ranks in the same way since the percentile rank transformation fixes the variation because of the uniform distribution.[28] However, the father's true rank is narrowed on after averaging observations, which does lead to a stronger rank-rank slope. While the smaller bias to the rank-rank slope may suggest that one should prefer rank-rank measures of mobility to the IGE, rank-rank mobility is also biased by error in the son's outcome, unlike the IGE (Nybom and Stuhler, 2017). Therefore, the small bias to the rank-rank slope when averaging more father observations may be misleading about the overall bias from the father and son. In the next section, I will create rank-rank estimates using a method that aims to eliminate the noise component for both the father and son.

Estimating father-son mobility based on classical measurement error or instrumental variables

Averaging three father observations revises the IGE, yet it still may be subject to measurement error. Under the assumption of classical measurement error, it is possible to project what the "true" IGE should be after eliminating noise. Before doing this projection, I can test whether the assumption is valid by comparing the actual three-father IGE to the projected three-father IGE under measurement error. Based on how the IGE changes from one to two father observations, classical measurement error predicts that the three-father-observation IGE is 0.704.[29] This prediction is surprisingly accurate: the actual estimate is 0.707. In fact, for all specifications

[28] Theoretically, the variation in percentile ranks for the [0,100] interval should be $\frac{1}{12}(100)^2$, or 833.3, no matter how many father's observations are averaged before ranking them. However, the variation of percentile ranks in my data is less than this theoretical value since many fathers have the same percentile rank due to having the same occupation, race and region.

[29] Based on Equation (2), I estimate $\hat{\beta}_{three\ obs} = \left[{\left(3\hat{\beta}_{one\ obs} \times \hat{\beta}_{two\ obs}\right)} \big/ {\left(4\hat{\beta}_{one\ obs} - \hat{\beta}_{two\ obs}\right)} \right]$.

shown in Table 5, the projected three-father IGE under classical measurement error is similar to the actual one. Therefore, while the classical measurement error assumption may seem too simplistic, it is consistent with patterns in the data.

Since the classical error assumption appears to hold, I continue to use it to eliminate the error and predict the "true" father-son elasticity. Based on this assumption, the predicted "true" father-son elasticity is 0.84 (see Table 5), or 19 percent higher than the three-father estimate.[30] A 0.84 estimate implies that only 16 percent of initial economic gaps across families disappeared by the next generation, which paints American history as highly immobile rather than highly mobile.

Rather than using the classical measurement error formula, one could instead use instrumental variables to estimate the IGE. This method instruments one father observation with another one under the assumption that the transitory components are not correlated across observations (Altonji and Dunn, 1991, Modalsli and Vosters, 2019). If one takes this approach and instruments the 1910 father observation with the 1920 father observation, then the estimated IGE is 0.82 – close to the 0.84 estimate under classical measurement error. Results are similar when switching different father years as the instrumental or endogenous variable.[31] Note the since I have multiple father observations as instruments, I can test whether the instruments are exogeneous with an overidentification test, which indeed suggests that the instruments are valid.[32]

Both the IV method and classical measurement error assumption require only two father observations rather than three, which suggests that I could relax the data requirement that fathers are double linked. If I instead keep fathers linked between 1910 and 1920, then the sample more than doubles from 394,864 to 825,251. This two-father-observation sample leads to the same result as the main sample: the IGE increases from 0.37 to 0.82 after accounting for race and measurement error with 2SLS (see Table A1). This result suggests that others who wish to estimate mobility do not need to satisfy the high data requirement of linking fathers three times. It also suggests that the

---

[30] Based on Equation (2), I estimate $\hat{\beta}_{\text{"}true\text{"}} = \left[\left(\hat{\beta}_{one\ obs} \times \hat{\beta}_{two\ obs}\right) \big/ \left(2\hat{\beta}_{one\ obs} - \hat{\beta}_{two\ obs}\right) \right]$.

[31] Estimates do not change much when using different years for the father as the endogenous and exogenous variable. For example, instrumenting the 1910 father observation with the 1900 observation yields a 0.815 estimate, instead of the 0.818 estimate when instrumenting with the 1920 father observation. Alternatively, instrumenting the 1920 observation with the 1910 one leads to a 0.804 estimate. I prefer to use the 1910 father observation as the endogenous variable since it is in the middle of the father's life-cycle.

[32] The p-value for the overidentification test (Hansen J statistic) is 0.41.

additional link does not change the representativeness of the sample (that is, after correcting for unrepresentativeness by weighting on observables). Given the similarity of results, I continue to use the three-father-observation sample for the rest of the paper.

Iterating the 0.82 estimate across multiple generations suggests that it will take ten generations to reduce initial gaps to about 15 percent, which is 8 generations longer (or roughly 200 years) than the baseline IGE of 0.37 – a large revision to historical mobility estimates based on linked father-son data. At the same time, this iterated result should be taken with a grain of salt because iterating across multiple generations is problematic (Stuhler, 2012). However, the point estimate does closely align with the surname-based evidence from Clark (2014, Chapter 3), who finds low mobility throughout American history. Note that Clark includes non-white populations in his data and averages out error when grouping by surname.

The result that mobility was low in the past is not entirely driven by the wide and persistent black-white gap.[33] When limiting the sample to only white families, the IGE is estimated at 0.69; when limiting the sample to black families, the IGE is estimated at 0.77. Therefore, within-race relative mobility appears to have been low.

It is also possible to predict the "true" rank-rank slope despite non-classical measurement error from the rank transformation. Using a generalized errors-in-variables model adapted from Haider and Solon (2006), Nybom and Stuhler (2017) propose a method to correct the bias in the rank-rank slope based on the association between two father observations.[34] However, since rank-rank measures are also influenced by transitory fluctuations in the son's outcome, unlike the IGE, measurement error for the son must also be addressed. Since I do not have multiple son

[33] However, the group effects are still important for the population estimate. If one decomposes the population IGE into between-race and across-race components using the decomposition from Hertz (2008), then the between-race component contributes about half of the population elasticity and the within-race component makes up the other half. See Appendix C for more detail.

[34] Following Nybom and Stuhler (2017), let λ_f be the attenuation factor from mismeasuring the father's rank and λ_s be the attenuation factor from mismeasuring the son's rank. That is, let $\tilde{y} = a + \lambda\tilde{y}^* + \widetilde{w}$, where $\tilde{y}^*$ is the true percentile rank and $\tilde{y}$ is the observed rank (see Haider and Solon (2006) for a similar model). Note that due to the percentile rank transformation, λ is less than or equal to one (Nybom and Stuhler, 2017). Based on this formulation, $\rho_{observed} = \lambda_s\lambda_f\rho_{true}$ where $\rho_{observed}$ is the rank-rank correlation between father and son based on one father observation and one son observation. Nybom and Stuhler (2017) show that a regression of the father's percentile rank on another father observation is equal to λ_f^2 if the error terms are uncorrelated. This regression between the 1920 and 1910 father's percentile ranks is 0.725, which suggests that $\lambda_f = 0.852$. Similarly, $\lambda_f = 0.809$ for white fathers, $\lambda_f = 0.772$ for black fathers, and $\lambda_f = 0.821$ when using the 1950 score. I use these values to back out ρ_{true}, where I assume that $\lambda_f = \lambda_s$.

observations, I assume that error in the son's outcome is of the same magnitude as error in the father's outcome. Based on this assumption, the predicted rank-rank measure of mobility is 0.69, which is 79 percent higher than the baseline estimate with one father observation and white families (0.39). Using an IV strategy where the 1910 percentile rank is instrumented with the 1920 percentile rank produces a similar estimate of 0.71.[35]

Comparison to estimates based on the surname average

Since the data include surnames, it is possible to compare mobility estimates from linked data to estimates when the father's status is imputed with the surname average (similar to Clark (2014)). Clark argues that averaging by surname reduces measurement error and therefore captures such that the true correlation across generations. However, convergence of surname status may differ from the convergence of individual status. For example, surnames could proxy for influences from ethnicity or geography, which could then lead to a stronger correlation than found in individual data (Chetty et al., 2014; Ward, 2019a). However, after imputing the father's status with the surname average, the elasticity is estimated at 0.80 (see Table A2) – almost the same as the 0.82 2SLS estimate.[36] Yet, the similarity in estimates only occurs when taking the average of the average (that is, the average of the 1900-1920 father score). If one instead averages the 1910 father snapshot by surname, then the grouped elasticity is estimated at 0.67, which is less than the IV estimate. These results suggest that averaging a snapshot by surname will not fully eliminate error when imputing the father's economic status, but the method does come closer to the 2SLS estimate with linked data.

Robustness.

One limitation of the results is that I use imputed income instead of actual income. It is possible to check how the imputed-income IGE compares with the actual-income IGE based on results from Iowa – at least when using one father observation (Feigenbaum, 2018).[37] Based on the same geography as the Iowa sample, the IGE from my preferred score is higher than

[35] An overidentification test also suggests that the percentile rank instruments are valid.
[36] Note that this is a Two-Sample Two-Stage Least Squares (TS2SLS) estimator, where the father's status is replaced by the average status by surname. For applications, see Olivetti and Paserman (2015) for averaging by first name and Borjas (1994) for averaging by country of birth. I group by the NYSIIS version of the 1910 father's last name.
[37] Note that self-employed earnings are still unobserved in the 1940 Federal Census, leading to Feigenbaum (2018) to impute non-wage income.

Feigenbaum's actual-income IGE (0.39 v 0.21, see Table A3).[38] Therefore, my income score appears to overestimate the actual income IGE for Iowans, which reinforces the point that one should not directly compare the historical imputed income IGEs to modern-day actual income IGEs. While the income score method misses key variation within occupation, race and region, I can still demonstrate that measurement error attenuates the Iowa income score IGE. It is unclear how measurement error affects historical income mobility estimates, but it may be more severe due to transitory fluctuations within occupation.

A similar issue is that farmer income is difficult to impute historically, which is important because farmers were a large share of the labor force and farming was highly persistent across generations. Dropping sons of farmers causes the "true" IGE to fall by about 20 percent, from 0.84 to 0.67 (see Panel C of Table 5). Yet, I continue to find that accounting for measurement error and race are important for estimating the non-farmer IGE.

I show in Table A4 that mobility estimates based on other measures of status are also affected by the sample's racial composition and measurement error. For example, a similar pattern holds when using the 1901 Cost of Living Survey or occupational wealth data from 1850-1870.[39] Importantly, if I use the status measure from Song et al. (2019), then I find that the father-son slope increases from 0.32 to 0.70 after accounting for race and measurement error. While all other measures are influenced by measurement error and racial disparities, the magnitude of the father-son association depends on the measure of status. In general, 2SLS estimates range from 0.55 to 0.82 when status is adjusted to reflect gaps by race and region.

In Appendix D, I gauge the importance of measurement error and racial disparities for the Altham statistic (Altham and Ferrie, 2007; Ferrie, 2005). The key advantage of the Altham statistic is that one does not have to impute income (Long and Ferrie, 2013b). As opposed to my preferred IGE, the Altham statistic is not strongly influenced by the sample's racial composition. This occurs since racial disparities within occupation are not captured in the Altham statistic. On the other hand, measurement error can influence the Altham statistic – depending on how one codes of the father's "true" category, which is not immediately clear. To see why, consider the following

[38] Note that when we both use the same 1950 score, I estimate a 0.41 IGE while Feigenbaum estimates a 0.44 IGE, suggesting that the samples are not that different.)
[39] I use the occupational wealth scores from Olivetti and Paserman (2015).

problem: if a person is a farmer in one census and a white-collar worker in a second census, which is his “true” occupational category? Averaging is not an option since people need to be assigned to a discrete category and because occupations are not on a univariate scale. Despite the ambiguity for how to handle measurement error in the Altham statistic, I show in Appendix D that the association between the son’s occupation category and the father’s category is strengthened after using multiple father observations. Therefore, measures of occupational mobility (without imputing income) are also influenced by measurement error.

The importance of measurement error and racial disparities does not depend on my linking method. Instead of using my sample which is based on hand-linked data, one could use a fully automated method to link censuses, methods which are more transparent. In Table A5, I recreate the estimates when using linked data created with the conservative iterative linking methods described by Abramitzky, Boustan and Eriksson (2012). Note that the conservative method keeps a *unique* link within a 5-year range, which substantially reduces the issue of false positives (see Abramitzky et al., 2019). These alternative samples confirm that accounting for race and measurement error are key for estimating the IGE. Interestingly, the alternative samples produce similar IGEs as found with my linked data, with a 2SLS IGE estimate between 0.820-0.827 (as opposed to 0.818 in my data).

While I am conservative in my linking methodology, it is possible that false positives are driving the result that measurement error matters. It could be that a perfectly linked dataset would find that the father’s occupation was more strongly correlated across observations.[40] In the appendix, I gauge the importance of linking error with two different checks. First, I purposely make the data lower quality by replacing my links with false links for an increasing percentage of the sample.[41] Going from replacing zero of my links to the 1920 census with false links (i.e., my main sample) to an extreme of replacing 25 percent of links with false ones increases the IV estimate from 0.82 to 0.84 (see Figure A3). As a second check, I make the better purposely better

[40] While false positives are often thought to attenuate IGEs due to a false link between father and son (Bailey et al., 2019), an upward bias may occur due to a false link between the father and a second observation. To see why, consider the 2SLS estimate where I instrument the 1910 father observation with the 1920 father observation. Linking the 1910 father to a wrong 1920 observation attenuates the first stage/denominator and potentially increases the IGE. At the same time, such error also attenuates the reduced form/numerator and lowers the IGE, leaving the overall bias unclear.
[41] I randomly match a father in my data to a male from the 1920 census with the same age, state of birth and race combination. This decision reflects the linking methodology, although technically age is allowed to match on a 5-year band for my main sample.

by limiting it to higher-quality 1910-1920 links based on the linking scores. The higher-quality data still leads to 2SLS estimates at or above 0.80 (See Figure A4). Therefore, linking error does not appear to be driving the results.

VI. Reevaluating the trend in intergenerational mobility across the 20th century

The evidence so far shows that the IGE was higher in the early 20th century than previously estimated with linked father-son data. This revision suggests that the trend in relative mobility over the 20th century should also be revised: instead of a decrease or flat trend in relative mobility over the past 100 years, it may be that relative mobility has increased. Since it is possible to use income scores across the 20th century, I re-estimate the trend in father-son mobility based on the IGE of imputed income.

To estimate the trend in relative mobility, I compare the early 20th century data to data from the Panel Study of Income Dynamics (PSID). To mimic the census data, I use white and black fathers who are observed ten years apart in 1968 and 1978 and the son's outcome from 30 years later in 1997. The father's and son's occupations are both observed at the 3-digit level, which matches the detail of the earlier census data.[42] To keep the outcome consistent over time, I impute the father's income with the mean income by occupation, race and region according to microdata from the 1970 census; I do the same process for the son in 1997 with microdata from the 2000 census (Ruggles et al., 2019).[43] Since the PSID is much smaller than the historical linked data, I use wider age restrictions such that the sons are between 30-55 in 1997 and the fathers are between

[42] I use the 3-digit occupation codes for the father from the Retrospective Occupation-Industry File. The Retrospective Occupation-Industry File recodes the 1-digit occupations in the original dataset to the 3-digit level after going back through the original interviewer files. Kambourov and Manovskii (2008) argue that this retrospective coding of occupations has less measurement error than the original data.

[43] The original occupation data are reported in 1970 Census occupation codes. I use the 1970 IPUMS 1 percent sample to calculate the average earnings at the occupation, region and race level, which is used for the 1968 and 1978 PSID observations. For the son's income score in 1997, I first create a cross-walk between the 1970 and 1990 occupation codes using the most common 1990 code for each 1970 occupation in the 1970 IPUMS 1% random sample. Then I create the income score in 1997 as the average earnings at the 1990-occupation, region and race level in the 2000 census. For occupation codes that are in the PSID but not in the 1970 1 percent sample, then I replace the imputed income with the average earnings for the first digit of the 3-digit code. Those in the PSID who either did not report an occupation or were living outside of the United States are dropped from the sample. Note that the results also hold if I use the 2000 census to impute both the father and son's income in the PSID, which matches the early 20th century data since I use the 1940 score to impute the father and son's income (see Table A6).

25-50 in 1968. Yet the average age is similar across time where the average father in the PSID is 40 and the average son is 39. Ultimately, I am left with 697 father-son pairs.[44]

In contrast to the recent literature, these data suggest an *increase* in mobility over time (see Figure 4). The IGE (estimated via 2SLS) fell from 0.82 to 0.57 between 1910-1940 and 1968-1997 – a decrease in persistence of 30 percent. The rank-rank slope also decreased from 0.71 to 0.44 (see Table 6 and Figure A5). Similar declines are found when using measurement error formulas (IGE from 0.84 to 0.59; rank-rank from 0.69 to 0.46). These results suggest that equality of opportunity, when measured with imputed income, has improved over the course of the 20th century.

The difference in mobility over time is not just found in the pooled data but is also found within race: the white-family IGE decreased from 0.69 to 0.51, while the black-family IGE decreased from 0.77 to an imprecise 0.19. (Note that the trend in black relative mobility is unclear due to a small PSID sample.) The increase in mobility also holds if one uses different measures of status over time (see Table A6). Importantly, the decline in mobility is also found in measures that are purely occupational-based, suggesting that the increase in mobility over time is not entirely due to narrowing racial and regional gaps within occupation. Further, I also find a decline in the IGE if I use measures that more closely approximate income. For example, if I use wage income for wage workers in 1940 and also use total family income for the PSID, the estimates suggest that the IGE declined over the 20th century from 0.94 to 0.66 (see Table A6).

If one does not account for measurement error or racial disparities, then the data recreate some of the literature's result that mobility decreased over time (see Figure 4). That is, if one uses a white-only sample with one father observation, then the estimated IGE increases from 0.37 to 0.47 over time, implying a decline in relative mobility. However, this difference is not statistically significant due to the small PSID sample size. Nevertheless, this result suggests that racial disparities and measurement error bias historical estimates more than PSID estimates. Indeed, adding black families more strongly affects the historical IGE (44 percent increase from 0.37 to 0.53) than the modern-day one (14 percent increase from 0.47 to 0.53), which occurs because of

[44] I use the 1997 family cross-sectional weights for the analysis but increase black family weights to make their proportion of the sample representative of the population. This effectively raises the modern-day IGE since black families are given a greater weight.

the larger black-white gap in the past. Moreover, measurement error more strongly attenuates the historical IGE: correcting for measurement error increases the historical IGE from 0.53 to 0.82, while the PSID IGE only increases from 0.53 to 0.57. This result reflects the lower correlations across father observations in the historical data. For example, while the 1910 and 1920 father observations had a 0.67 correlation, the 1968 and 1978 father observations had a 0.89 correlation. Lower correlations in the historical data may occur because occupations were less stable in the past or because the early 20th century data are of poorer quality.

When using one father observation, my early 20th century and PSID results closely align with the trend in mobility from Song et al. (2019), but I can also demonstrate that accounting for measurement error revises the trend. When using their status measure, they find that the rank-rank slope was stagnant from 0.31 to 0.33 across the early and late 20th century.[45] When I use the same status measure, I find a similarly trendless line from 0.35 to 0.37 – before correcting for measurement error. (Note that I include black families in my estimate, while they do not.) When using the 2SLS method, I find a decline from 0.70 to 0.45 (see Table A6). This result once again suggests that measurement error is so severe in historical data that it biases the trend over time.

Since farming is a key reason for why historical mobility was low, the increase in mobility over time could just reflect the economy's shift away from agriculture. Indeed, the decrease of IGE over time is less stark after dropping the sons of farmers. Specifically, the 2SLS IGE falls from 0.63 to 0.57 instead of from 0.82 to 0.57 (see Panel C of Table 6). The difference in the early and late 20th century non-farmer IGEs is not statistically significant, although the difference is statistically significant for percentile ranks (from 0.67 to 0.44). This result suggests that the trend in relative mobility can be largely attributed to a structural shift away from agriculture, an argument that has been made elsewhere in the sociology literature (e.g., Guest et al., 1989; Blau and Duncan, 1967; Xie and Killewald, 2013; Song et al., 2019). Of course, the trend in the IGE is a function of numerous other changes over time (e.g., fertility, household formation, assortative mating, residential segregation, education premium, internal migration, institutions, etc.), so it cannot be simplified to one explanation. Explaining why mobility changed is outside the scope of this paper; instead, I focus on correctly measuring mobility in the first place. Ultimately, more

[45] I compare estimates from the 1900 birth cohort for the 1910-1940 data, and 1960 cohort for the 1968-1997 data. Since I primarily use the rank-rank slope for my results, I compare my results to theirs from Appendix Table S10 (Song et al., 2019). I also compare my result to their results from survey data.

research is needed on comparative mobility across countries and time periods to understand the forces behind changes in mobility.

VII. Conclusion

The main message of this paper is that intergenerational mobility was lower in the early 20th century than previously recognized in linked father-son data. To show why, I account for two measurement issues that had not been fully addressed: racial disparities in mobility and measurement error. First, I account for racial disparities by adding black families to the sample and using an income score that captures the historical black-white income gap. Second, I account for measurement error by using multiple father observations to more accurately capture his permanent economic status. These issues are not new to the literature (e.g., Solon, 1992; Duncan, 1968; Hertz, 2005), but due to various data limitations, they had not been fully addressed in historical linked studies that find high mobility. My preferred estimate is that the imputed income IGE was 0.82-0.84 in the early 20th century – 2.3 times higher than an estimate that does not account for race or measurement error (0.37). I also find that the within-race IGEs are also high, with the white elasticity at 0.69-0.72 and the black elasticity at 0.77-0.79.

The results raise the possibility that measurement error varies across source, time and space, which could then bias inference in comparative mobility research. For example, I show that the historical IGE was more strongly attenuated than the modern one, perhaps because the early 20th century data are of lower quality than the PSID or because there was substantially higher intragenerational mobility in the past. I also show that if one does not address measurement error in both datasets, then one would wrongly infer the white mobility was higher in the past. These results suggest that future comparative mobility studies should account for variation in measurement error with at least two father observations, perhaps by instrumenting one observation with a second one.[46]

Besides contributing to the intergenerational mobility literature, the results have broad implications for any historical research which uses occupation. Specifically, I find that about two-thirds of the observed variation in log occupational income is due to variation in permanent status,

[46] Note that there is evidence that mobility was even higher in the mid-19th century than the early 20th century (Song et al., 2019). A future study could estimate mobility using linked data where the father is observed twice, though how to include those who were emancipated is unclear.

while one-third is due to noise. Part of the noise may be due to errors in the data from coding or digitization, but it also may be due to high intragenerational mobility. Either way, any historical analysis that uses occupation as an independent variable to proxy for permanent status is also subject to attenuation bias.

When I estimate the trend in IGE over time using consistent methods, the results suggest the optimistic conclusion that equality of opportunity is greater today than in the past. An increase in mobility over time is consistent with others who estimate that institutional changes over the 20th century helped to improve outcomes for disadvantaged groups (e.g., Card and Krueger, 1992; Hoynes et al., 2016; Reber, 2010).[47] The results also suggest that the structural shift away from agriculture matters since there is no change in relative mobility after dropping the sons of farmers. Of course, there are many other factors besides structural and institutional change that affect relative mobility trends. But before we can understand what causes mobility to change over time, we must first accurately measure mobility by accounting for racial disparities and measurement error.

[47] Derenoncourt (2019) also shows evidence that institutional changes affected mobility rates in northern cities following the Great Migration. However, instead of institutional changes increasing equality of opportunity, they decreased it for African Americans.

References

Aaronson, Daniel, and Bhashkar Mazumder. "Intergenerational economic mobility in the United States, 1940 to 2000." *Journal of Human Resources* 43.1 (2008): 139-172.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson. "Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration." *American Economic Review* 102.5 (2012): 1832-56.

Abramitzky, Ran, Leah Platt Boustan, Katherine Eriksson, James J. Feigenbaum, and Santiago Pérez. *Automated linking of historical data*. No. w25825. National Bureau of Economic Research, 2019.

Ager, Philipp, Leah Platt Boustan, and Katherine Eriksson. *The intergenerational effects of a large wealth shock: White southerners after the Civil War*. No. w25700. National Bureau of Economic Research, 2019.

Altham, Patricia ME, and Joseph P. Ferrie. "Comparing contingency tables tools for analyzing data from two groups cross-classified by two characteristics." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 40.1 (2007): 3-16.

Altonji, Joseph G., and Thomas A. Dunn. "Relationships Among the Family Incomes and Labor Market Outcomes of Relatives." *NBER Working Paper* w3724 (1991).

Bailey, Martha, et al. *How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth*. No. w24019. National Bureau of Economic Research, 2019.

Bingley, Paul, and Alessandro Martinello. "Measurement error in income and schooling and the bias of linear estimators." *Journal of Labor Economics* 35.4 (2017): 1117-1148.

Borjas, George J. "Long-run convergence of ethnic skill differentials: The children and grandchildren of the great migration." *ILR Review* 47.4 (1994): 553-573.

Card, David, and Alan B. Krueger. "School quality and black-white relative earnings: A direct assessment." *The Quarterly Journal of Economics* 107.1 (1992): 151-200.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. "Is the United States still a land of opportunity? Recent trends in intergenerational mobility." *American Economic Review* 104, no. 5 (2014): 141-47.

Chetty, Raj, et al. *Race and economic opportunity in the United States: An intergenerational perspective*. No. w24441. National Bureau of Economic Research, 2018.

Chetty, Raj, Hendren, N., Kline, P., & Saez, E. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics* 129.4 (2014): 1553-1623.

Chetty, Raj, David Grusky, Maximilian Hell, Nathaniel Hendren, Robert Manduca, and Jimmy Narang. "The fading American dream: Trends in absolute income mobility since 1940." *Science* 356, no. 6336 (2017): 398-406.

Cilliers, Jeanne, and Johan Fourie. "Occupational Mobility during South Africa's Industrial Take-Off." *South African Journal of Economics* 86.1 (2018): 3-22.

Clark, Gregory, and Neil Cummins. "Intergenerational wealth mobility in England, 1858–2012: surnames and social mobility." *The Economic Journal* 125.582 (2014): 61-85.

Clark, Gregory. *The son also rises: Surnames and the history of social mobility*. Vol. 49. Princeton University Press, 2014.

Collins, William J., and Marianne H. Wanamaker. *Up from Slavery? African American Intergenerational Economic Mobility Since 1880*. No. w23395. National Bureau of Economic Research, 2017.

Connor, Dylan Shane. "Class Background, Reception Context, and Intergenerational Mobility: A Record Linkage and Surname Analysis of the Children of Irish Immigrants." *International Migration Review* (2018): 0197918318806891.

Corak, Miles. "Inequality from generation to generation: The United States in comparison." Chapter 6 in *The Economics of Inequality, Poverty and Discrimination in the 21st century*, edited by Robert S. Rycroft. Santa Barbara, CA: ABC-CLIO (2013).

Craig, Jacqueline, Katherine Eriksson, and Gregory T. Niemesh. "Marriage and the Intergenerational Mobility of Women: Evidence from Marriage Certificates 1850-1910."

Davis, Jonathan, and Bhashkar Mazumder. "The decline in intergenerational mobility after 1980." (2017).

Derenoncourt, Ellora. “Can you move to opportunity? Evidence from the Great Migration” (2019).

Duncan, Otis Dudley. "Inheritance of poverty or inheritance of race?" *On understanding poverty* (1968): 85-110.

Dupraz, Yannick, and Andreas Ferrara. "Fatherless: The Long-Term Effects of Losing a Father in the US Civil War." (2018).

Feigenbaum, James J. "Automated census record linking: A machine learning approach." (2016).

Feigenbaum, James J. "Intergenerational mobility during the great depression." (2015).

Feigenbaum, James J. "Multiple measures of historical intergenerational mobility: Iowa 1915 to 1940." *The Economic Journal* 128.612 (2018): F446-F481.

Ferrie, Joseph P. "History lessons: The end of American exceptionalism? Mobility in the United States since 1850." *Journal of Economic Perspectives* 19.3 (2005): 199-215.

Ferrie, Joseph, Catherine Massey, and Jonathan Rothbaum. *Do grandparents and great-grandparents matter? Multigenerational mobility in the US, 1910-2013*. No. w22635. National Bureau of Economic Research, 2016.

Goldin, Claudia Dale, and Lawrence F. Katz. *The race between education and technology*. Harvard University Press, 2008.

Goldin, Claudia, and Lawrence F. Katz. "Education and income in the early twentieth century: Evidence from the prairies." *The Journal of Economic History* 60.3 (2000): 782-818.

Goldin, Claudia, and Robert A. Margo. "The great compression: The wage structure in the United States at mid-century." *The Quarterly Journal of Economics* 107.1 (1992): 1-34.

Grusky, D.B., 1986. American Social Mobility in the 19^{th} and 20^{th} centuries. CDE Working Paper 86-28.

Guest, Avery M., Nancy S. Landale, and James C. McCann. "Intergenerational occupational mobility in the late 19th century United States." *Social Forces* 68.2 (1989): 351-378.

Haider, Steven J. "Earnings instability and earnings inequality of males in the United States: 1967–1991." *Journal of labor Economics* 19.4 (2001): 799-836.

Haider, Steven, and Gary Solon. "Life-cycle variation in the association between current and lifetime earnings." *American Economic Review* 96.4 (2006): 1308-1320.

Hertz, Tom. "A group-specific measure of intergenerational persistence." *Economics Letters* 100.3 (2008): 415-417.

Hertz, Tom. "Rags, Riches and Race: The Intergenerational Economic Mobility of Black and White Families in the United States." In Samuel Bowles, Herbert Gintis, and Melissa Osborne (editors). Unequal Chances: Family Background and Economic Success. Princeton: Princeton University Press and Russell Sage. (2005).

Hout, Michael, and Avery M. Guest. "Intergenerational occupational mobility in Great Britain and the United States since 1850: Comment." *American Economic Review* 103.5 (2013): 2021-40.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. "Long-run impacts of childhood access to the safety net." *American Economic Review* 106.4 (2016): 903-34.

Inwood, Kris, Chris Minns, and Fraser Summerfield. "Occupational income scores and immigrant assimilation. Evidence from the Canadian census." *Explorations in Economic History* (2019).

Kambourov, Gueorgui, and Iourii Manovskii. "Rising occupational and industry mobility in the United States: 1968–97." *International Economic Review* 49.1 (2008): 41-79.

Karbownik, Krzysztof, and Anthony Wray. *Educational, Labor-market and Intergenerational Consequences of Poor Childhood Health*. No. w26368. National Bureau of Economic Research, 2019.

Kosack, Edward, and Zachary Ward. "El sueño americano? The generational progress of Mexican Americans prior to World War II" (2018).

Lee, Chul-In, and Gary Solon. "Trends in intergenerational income mobility." *The Review of Economics and Statistics* 91.4 (2009): 766-772.

Long, Jason, and Joseph Ferrie. "Grandfathers matter (ed): occupational mobility across three generations in the US and Britain, 1850–1911." *The Economic Journal* 128.612 (2018): F422-F445.

Long, Jason, and Joseph Ferrie. "Intergenerational occupational mobility in Great Britain and the United States since 1850." *American Economic Review* 103.4 (2013): 1109-37.

Long, Jason, and Joseph Ferrie. "Intergenerational occupational mobility in Great Britain and the United States since 1850: reply." *American Economic Review* 103.5 (2013): 2041-49.

Long, Jason, and Joseph Ferrie. "Grandfathers matter (ed): occupational mobility across three generations in the US and Britain, 1850–1911." *The Economic Journal* 128.612 (2018): F422-F445.

Margo, Robert A. "Obama, Katrina, and the persistence of racial inequality." *The Journal of Economic History* 76.2 (2016): 301-341.

Mazumder, Bhashkar, and Miguel Acosta. "Using occupation to measure intergenerational mobility." *The ANNALS of the American Academy of Political and Social Science* 657.1 (2015): 174-193.

Mazumder, Bhashkar. "Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data." *Review of Economics and Statistics* 87.2 (2005): 235-255.

Mitchener, Kris James, and Ian W. McLean. "US regional growth and convergence, 1880–1980." *The Journal of Economic History* 59.4 (1999): 1016-1042.

Modalsli, Jørgen and Kelly Vosters, 2019. "Spillover bias in multigenerational income regressions," Discussion Papers 897, Statistics Norway, Research Department.

Modalsli, Jørgen. "Intergenerational mobility in Norway, 1865–2011." *The Scandinavian Journal of Economics* 119.1 (2017): 34-71.

Nybom, Martin, and Jan Stuhler. "Biases in standard measures of intergenerational income dependence." *Journal of Human Resources* 52.3 (2017): 800-825.

Olivetti, Claudia, and M. Daniele Paserman. "In the name of the son (and the daughter): Intergenerational mobility in the united states, 1850-1940." *American Economic Review* 105.8 (2015): 2695-2724.

Parman, John. "American mobility and the expansion of public education." *The Journal of Economic History* 71.1 (2011): 105-132.

Pérez, Santiago. "Intergenerational Occupational Mobility across Three Continents." *The Journal of Economic History* (2019): 1-34.

Pérez, Santiago. "The (South) American Dream: Mobility and Economic Outcomes of First-and Second-Generation Immigrants in Nineteenth-Century Argentina." *The Journal of Economic History* 77.4 (2017): 971-1006.

Piketty, Thomas, and Emmanuel Saez. "Income inequality in the United States, 1913–1998." *The Quarterly journal of economics* 118.1 (2003): 1-41.

Reber, Sarah J. "School desegregation and educational attainment for blacks." *Journal of Human resources* 45.4 (2010): 893-914.

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. IPUMS USA: Version 9.0 [dataset]. Minneapolis, MN: IPUMS, 2019. https://doi.org/10.18128/D010.V9.0

Saavedra, Martin Hugo, and Tate Twinam. "A Machine Learning Approach to Improving Occupational Income Scores." (2018).

Saez, Emmanuel, and Gabriel Zucman. "Wealth inequality in the United States since 1913: Evidence from capitalized income tax data." *The Quarterly Journal of Economics* 131.2 (2016): 519-578.

Sobek, Matthew. "Work, status, and income: Men in the American occupational structure since the late nineteenth century." *Social Science History* 20.2 (1996): 169-207.

Solon, Gary. "Intergenerational Income Mobility in the United States." *American Economic Review* 82.3 (1992): 393-408.

Song, Xi, Catherine G. Massey, Karen A. Rolf, Joseph P. Ferrie, Jonathan L. Rothbaum and Yu Xie. "Long-Term Decline in Intergenerational Mobility in the United States since the 1850s." *Proceedings of the National Academy of Sciences,* Nov 2019; DOI: 10.1073/pnas.1905094116.

Tan, Huiren. “A Different Land of Opportunity: The Geography of Intergenerational Mobility in the Early 20th-century US.” Working paper. (2018).

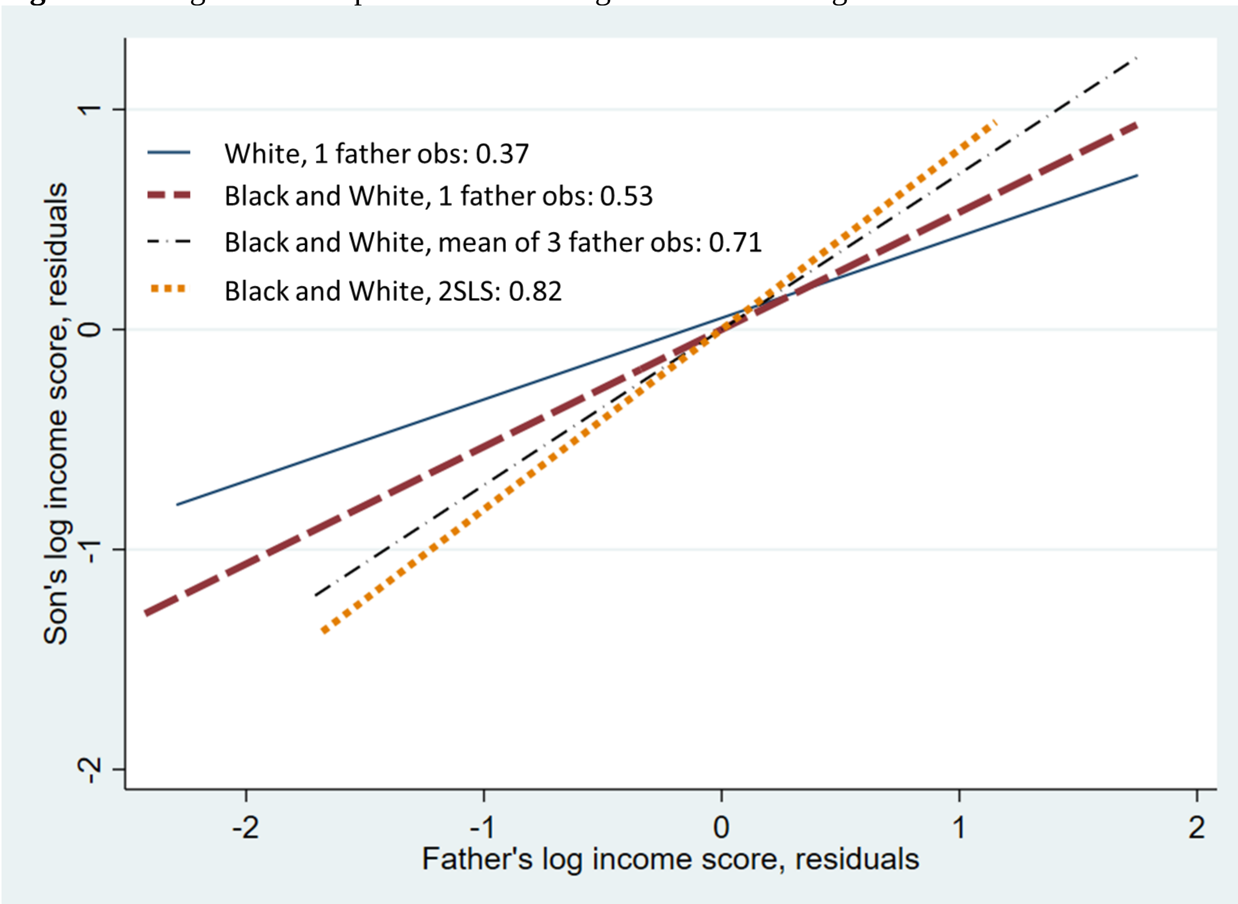
Ward, Zachary. "The Not-So-Hot Melting Pot: The Persistence of Outcomes for Descendants of the Age of Mass Migration." *American Economic Journal: Applied Economics*, forthcoming. 2019a.

Ward, Zachary. *Internal Migration, Education and Upward Rank Mobility: Evidence from American History*. No. 04. Centre for Economic History, Research School of Economics, Australian National University, 2019b.

Xie, Yu, and Alexandra Killewald. "Intergenerational occupational mobility in Great Britain and the United States since 1850: Comment." *American Economic Review* 103.5 (2013): 2003-20.

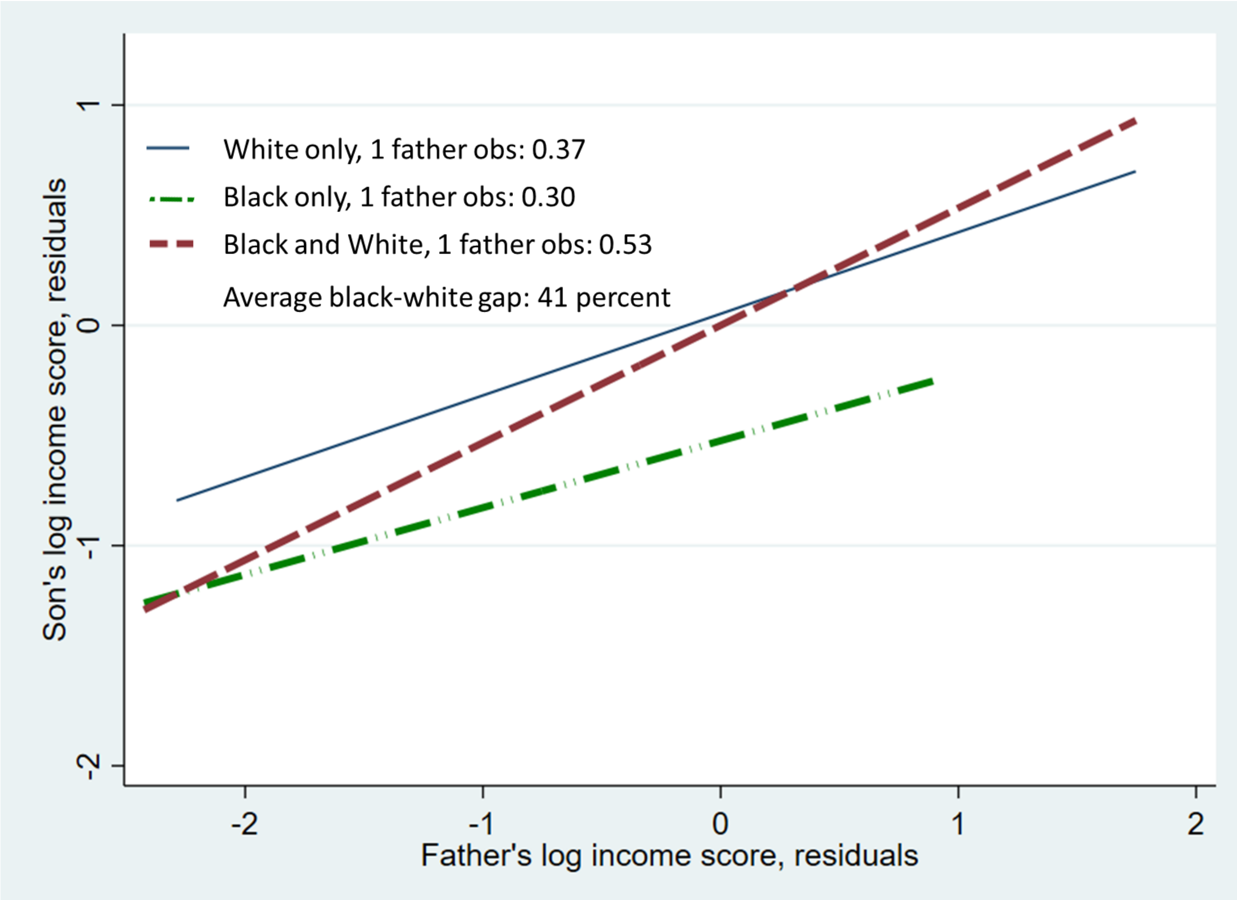
Zimmerman, David J. "Regression toward mediocrity in economic stature." *The American Economic Review* (1992): 409-429.

Figure 1. Intergenerational persistence is stronger after accounting for race and measurement error



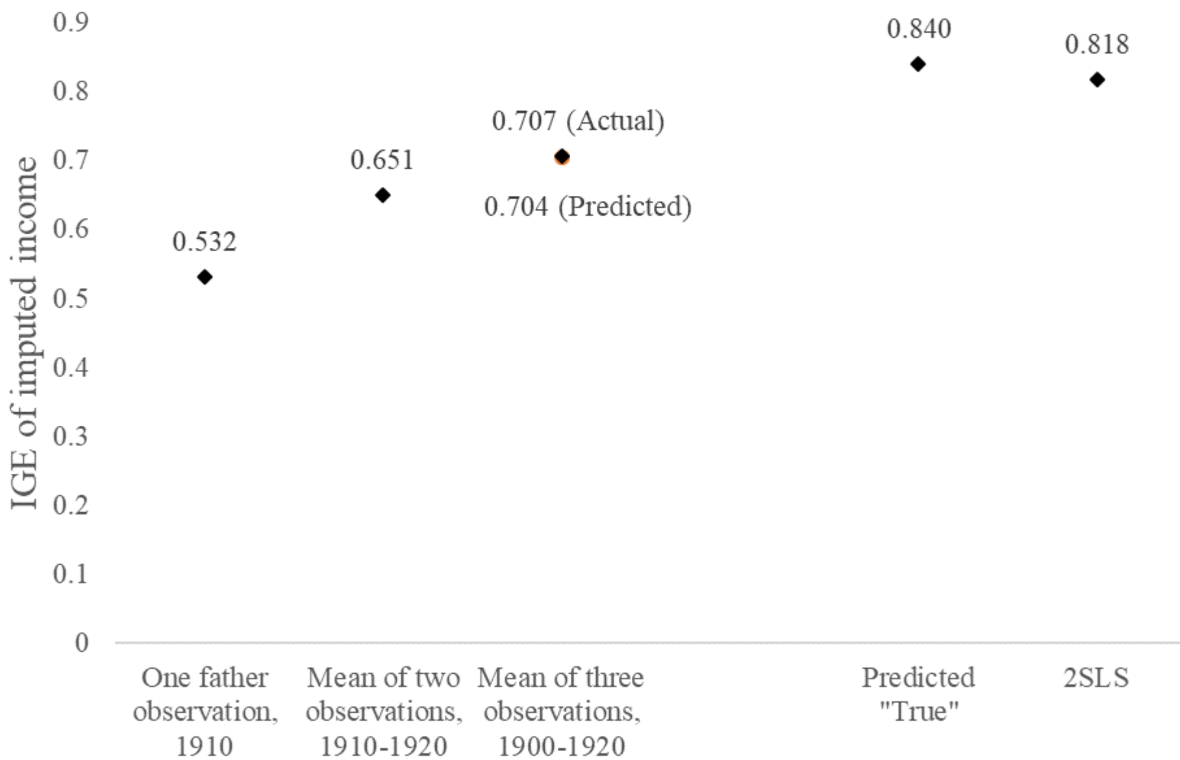
Notes: Data are a linked sample of fathers and sons from 1900, 1910, 1920 and 1940 United States Censuses. All results are intergenerational elasticity estimates after removing lifecycle effects with a quartic in age for the father and son. The first line ("White, 1 father obs.") is a common IGE estimate in the historical literature, where I regress the son's 1940 log income score on the father's 1910 log income score when limiting the sample to white families and one father observation. The second line ("Black and White, 1 father obs.") is the same regression, but pools black families with white families. The third line ("Black and White, mean of 3 father obs.") averages the log income score of the father from the 1900, 1910 and 1920 censuses, which reduces measurement error. The fourth line ("Black and White, 2SLS") instruments the 1910 father's income score with the 1920 income score. The point of this method is to purge the 1910 income score of measurement error under the assumption that error is uncorrelated across census observations ten years apart.

Figure 2. Equality of opportunity is less for the overall population than for the white population



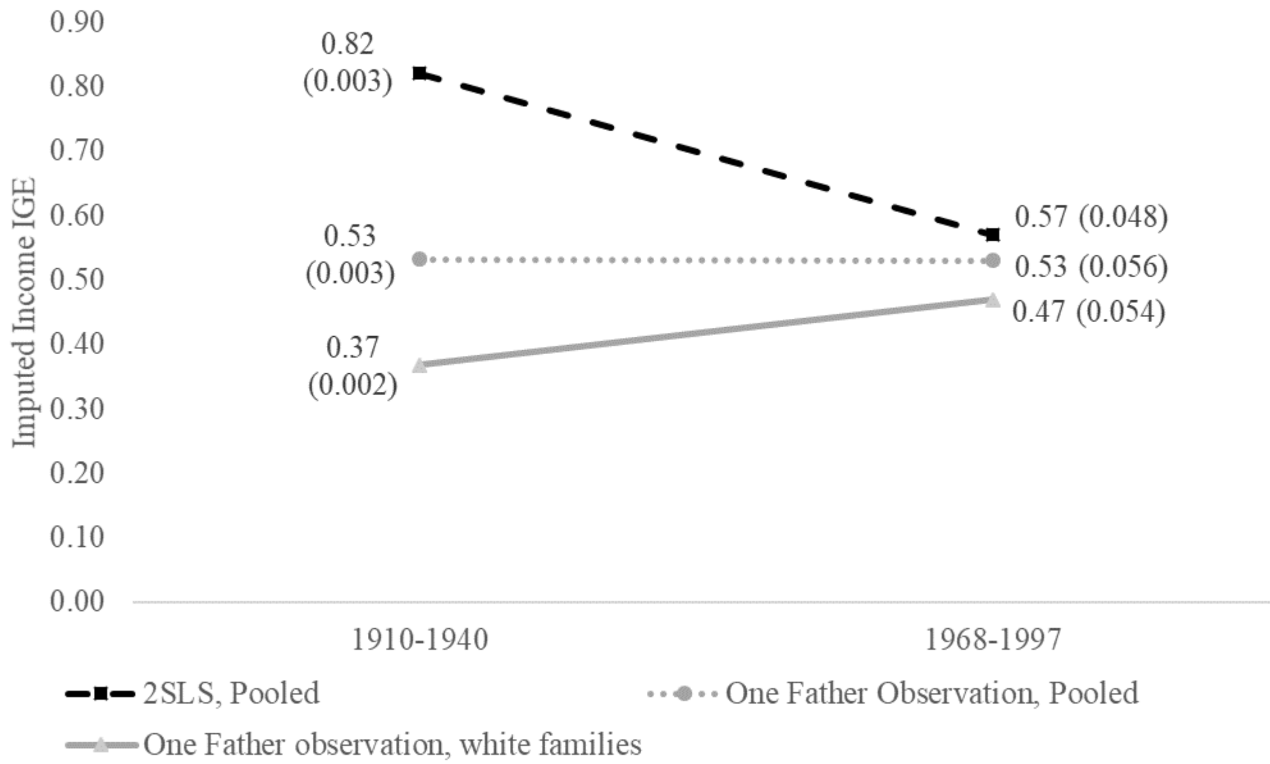
Notes: Data are a linked sample of fathers and sons from 1900, 1910, 1920 and 1940 United States Censuses. The results show a log-log estimate of mobility (i.e. intergenerational elasticity). See Table 3 for underlying coefficients for different scores. The estimates do not account for measurement error in the father's observation.

Figure 3. The one-father IGE is attenuated by measurement error



Notes: Data are a linked sample of fathers and sons from 1900, 1910, 1920 and 1940 United States Censuses. The results show a log-log estimate of mobility (i.e. intergenerational elasticity). The predicted elasticity (0.704) is based off of the classical measurement error formula. The actual elasticity (0.707) uses the actual average of three father observations. The predicted "true" estimate is based off the classical measurement error formula, after completely eliminating noise.

Figure 4. The trend in the income score IGE over 20th century



Notes: Data are from the early 20th century linked sample and the PSID. The underlying coefficients are reported in Panel A of Table 6. Standard errors are in parentheses.

Table 1. Descriptive Statistics of linked dataset

	Fathers			Sons		
	White	Black	Pooled	White	Black	Pooled
Age	38.17	38.05	38.15	36.54	36.58	36.55
	(4.81)	(4.86)	(4.82)	(4.28)	(4.44)	(4.30)
Log Income Score	9.80	8.98	9.71	9.89	9.06	9.81
	(0.43)	(0.28)	(0.49)	(0.48)	(0.37)	(0.53)
Percentile Rank of Inc. Score	53.39	8.36	48.63	54.10	11.72	49.90
	(27.42)	(10.03)	(29.58)	(27.20)	(11.62)	(28.99)
Log Occupational Score	9.88	9.59	9.85	10.04	9.71	10.01
	(0.45)	(0.31)	(0.45)	(0.45)	(0.42)	(0.46)
Percentile Rank of Occ. Score	41.11	19.57	38.83	48.89	25.80	46.60
	(36.11)	(25.37)	(35.75)	(29.94)	(21.28)	(30.00)
Observations	320,169	6,190	326,359	387,922	6,942	394,864

Notes: Data are a linked sample of fathers and sons from 1900, 1910, 1920 and 1940 United States Censuses, is the pooled sample of white and black families (Ruggles et al., 2019). Descriptive statistics are all weighted for representativeness; see Appendix Tables B5 and B6 for comparisons to the general population. Income and occupational scores are adjusted to be in 2016 dollars. The (single) log income score of the father is from 1910, and the log income of sons is from 1940. Income scores are imputed income based on occupation, race and region; occupational scores are imputed income based solely on occupation. Note that the average rank is not 50 because the equal rank is given to those with the same income score.

Table 2. The father's outcome is not strongly correlated across the 1900, 1910 and 1920 censuses

Log 1940 Income Score				Log 1950 Occupational Score			
	1900	1910	1920		1900	1910	1920
1900	1			1900	1		
1910	0.650	1		1910	0.544	1	
1920	0.646	0.672	1	1920	0.526	0.593	1
White Collar				Farmer			
	1900	1910	1920		1900	1910	1920
1900	1			1900	1		
1910	0.499	1		1910	0.542	1	
1920	0.492	0.544	1	1920	0.496	0.669	1
Unskilled				Semi-Skilled			
	1900	1910	1920		1900	1910	1920
1900	1			1900	1		
1910	0.230	1		1910	0.463	1	
1920	0.207	0.317	1	1920	0.419	0.510	1
Log 1940 Income Score, White Families				Log 1940 Income Score, Black Families			
	1900	1910	1920		1900	1910	1920
1900	1			1900	1		
1910	0.544	1		1910	0.373	1	
1920	0.536	0.568	1	1920	0.343	0.390	1

Notes: Data are a linked sample of fathers and sons from 1900, 1910, 1920 and 1940 United States Censuses (Ruggles et al., 2019). Correlation matrices are all weighted for representativeness. Life-cycle effects are removed after controlling for a quartic in the father's age. Income scores are imputed income based on occupation, race and region; occupational scores are imputed income based solely on occupation. There are 326,359 total fathers, 320,169 white fathers and 6,190 black fathers.

Table 3. Estimates of intergenerational persistence increase when including black families

	Log-log (IGE)			Rank-rank		
	White	Black	Pooled	White	Black	Pooled
Panel A. 1940 income score						
Father's outcome	0.371	0.304	0.533	0.387	0.363	0.503
	(0.002)	(0.018)	(0.003)	(0.002)	(0.018)	(0.002)
N	387,922	6,942	394,864	387,922	6,942	394,864
R^2	0.111	0.054	0.234	0.151	0.098	0.258
Panel B. 1950 occupational score						
Father's outcome	0.299	0.166	0.323	0.282	0.176	0.302
	(0.002)	(0.017)	(0.002)	(0.001)	(0.010)	(0.001)
N	387,922	6,942	394,864	387,922	6,942	394,864
R^2	0.089	0.015	0.098	0.115	0.044	0.129
Panel C. 1950 occupational score, adjusted for race, region and farm ownership						
Father's outcome	0.357	0.325	0.489	0.379	0.249	0.472
	(0.002)	(0.016)	(0.003)	(0.002)	(0.010)	(0.002)
N	387,922	6,942	394,864	387,922	6,942	394,864
R^2	0.128	0.060	0.238	0.147	0.089	0.232

Notes: Data are a linked sample of fathers of sons from the 1910 and 1940 United States Censuses (Ruggles et al., 2019). The columns show how mobility estimates vary when limiting the sample to white families or black families, and then pooling white and black families. The 1940 score is imputed earnings by occupation, race and region. The 1950 score is the IPUMS variable *occscore*. Panel C reports results for the 1950 score after adjusting for income differentials by race, region and farm ownership according to the 1940 census.

Table 4. Persistence is higher when averaging three observations of father's occupation

	Log-log (IGE)			Rank-rank		
	1 obs	Mean of 2 obs	Mean of 3 obs	1 obs	Mean of 2 obs	Mean of 3 obs
Panel A. 1940 income score, black and white families						
1900 Census	0.545			0.476		
	(0.002)	0.653		(0.002)	0.533	
1910 Census	0.533	(0.003)	0.707	0.503	(0.002)	0.556
	(0.003)	0.652	(0.003)	(0.002)	0.539	(0.002)
1920 Census	0.558	(0.003)		0.512	(0.002)	
	(0.003)			(0.002)		
Panel B. 1940 income score, only white families						
1900 Census	0.379			0.359		
	(0.002)	0.486		(0.002)	0.419	
1910 Census	0.371	(0.002)	0.547	0.387	(0.002)	0.447
	(0.002)	0.490	(0.002)	(0.002)	0.428	(0.002)
1920 Census	0.398	(0.002)		0.399	(0.002)	
	(0.002)			(0.002)		
Panel C. 1940 income score, only black families						
1900 Census	0.297			0.244		
	(0.015)	0.438		(0.014)	0.419	
1910 Census	0.304	(0.018)	0.522	0.363	(0.020)	0.501
	(0.018)	0.439	(0.019)	(0.018)	0.457	(0.023)
1920 Census	0.306	(0.018)		0.327	(0.022)	
	(0.017)			(0.018)		
Panel D. 1950 occupational score, black and white families						
1900 Census	0.300			0.281		
	(0.002)	0.404		(0.001)	0.332	
1910 Census	0.323	(0.002)	0.458	0.302	(0.002)	0.366
	(0.002)	0.417	(0.002)	(0.001)	0.341	(0.002)
1920 Census	0.342	(0.002)		0.315	(0.001)	
	(0.002)			(0.001)		

Notes: Data are a linked sample of fathers and sons from 1900, 1910, 1920 and 1940 United States Censuses. The columns show how mobility estimates vary when taking the average of log father's income score from different censuses. The 2-observation column either averages 1900-1910 or 1910-1920; the 3 observation column averages 1900-1910-1920. IGE stands for intergenerational elasticity estimate. The 1940 score is imputed earnings by occupation, race and region. The 1950 score is the IPUMS variable *occscore*.

Table 5. Estimates of the "true" rate of persistence after eliminating error

	Number of father obs.			Correcting for measurement error			
	One	Two	Three	Implied "Three"	Implied "True"	2SLS	N
Panel A. IGE							
1940 Income Score	0.533	0.652	0.707	0.704	0.839	0.818	394,864
	(0.003)	(0.003)	(0.003)			(0.003)	
1940 Income Score, only white	0.371	0.490	0.547	0.549	0.721	0.685	387,922
	(0.002)	(0.002)	(0.002)			(0.003)	
1940 Income Score, only black	0.304	0.439	0.522	0.515	0.790	0.771	6,942
	(0.018)	(0.018)	(0.019)			(0.040)	
1950 Occupational Income	0.323	0.417	0.458	0.462	0.588	0.577	394,864
	(0.002)	(0.002)	(0.002)			(0.003)	
Panel B. Rank-rank							
1940 Income Score	0.503	0.539	0.556		0.694	0.706	394,864
	(0.002)	(0.002)	(0.002)			(0.002)	
1940 Income Score, only white	0.387	0.428	0.447		0.591	0.608	387,922
	(0.002)	(0.002)	(0.002)			(0.003)	
1940 Income Score, only black	0.363	0.457	0.501		0.610	0.610	6,942
	(0.018)	(0.022)	(0.023)			(0.033)	
1950 Occupational Income	0.302	0.341	0.366		0.448	0.457	394,864
	(0.001)	(0.001)	(0.002)			(0.002)	
Panel C. IGE, drop sons of farmers							
1940 Income Score	0.374	0.480	0.531	0.530	0.670	0.629	167,505
	(0.004)	(0.004)	(0.004)			(0.005)	
1940 Income Score, only white	0.229	0.318	0.365	0.365	0.520	0.471	165,649
	(0.003)	(0.003)	(0.003)			(0.005)	
1940 Income Score, only black	0.351	0.433	0.487	0.470	0.565	0.604	1,856
	(0.024)	(0.024)	(0.025)			(0.044)	
1950 Occupational Income	0.181	0.258	0.295	0.301	0.449	0.400	167,505
	(0.004)	(0.004)	(0.005)			(0.007)	

Source: Data are from the 1900-1920, 1940 US Censuses (Ruggles et al., 2019)
Notes: The measurement error columns show the projections of mobility under the assumption of classical measurement error for the IGE measures, and nonclassical measurement error for the percentile rank measures. Note that measurement error in the son's percentile rank is not accounted for, which would increase the rank-rank slope even further. The Implied "Three" column predicted the IGE when averaging three father observation based on the results from the one and two father observations and when using the classical measurement error formulas. The Implied "Three" estimate cannot be made for the rank-rank estimates due to nonclassical measurement error. The Implied "True" column are the projections after eliminating the noise component. The 2SLS column instruments for the 1910 father observation with the 1920 father observation.

Table 6. Decreased persistence of economic status across the 20th century

	1 father observation		Implied "True"		2SLS	
	1910-1940	1968-1997	1910-1940	1968-1997	1910-1940	1968-1997
Panel A: IGE						
Black and white	0.533	0.533	0.839	0.585	0.818	0.570
	(0.003)	(0.050)			(0.003)	(0.055)
Only white	0.371	0.466	0.721	0.510	0.685	0.505
	(0.002)	(0.066)			(0.003)	(0.077)
Only black	0.304	0.194	0.790	0.209	0.771	0.185
	(0.018)	(0.126)			(0.040)	(0.180)
Panel B: Rank-rank						
Black and white	0.503	0.396	0.694	0.462	0.706	0.438
	(0.002)	(0.039)			(0.002)	(0.044)
Only white	0.387	0.309	0.591	0.369	0.608	0.351
	(0.002)	(0.045)			(0.003)	(0.053)
Only black	0.363	0.444	0.610	0.880	0.610	0.287
	(0.018)	(0.159)			(0.033)	(0.394)
Panel C: IGE, no sons of farmers						
Black and white	0.374	0.531	0.670	0.581	0.629	0.568
	(0.004)	(0.050)			(0.005)	(0.055)
Only white	0.229	0.462	0.520	0.504	0.471	0.500
	(0.003)	(0.066)			(0.005)	(0.077)
Only black	0.351	0.194	0.565	0.209	0.604	0.185
	(0.024)	(0.126)			(0.044)	(0.180)
Panel D: Rank-rank, no sons of farmers						
Black and white	0.396	0.397	0.652	0.460	0.665	0.437
	(0.004)	(0.040)			(0.005)	(0.045)
Only white	0.251	0.306	0.520	0.366	0.531	0.348
	(0.003)	(0.046)			(0.005)	(0.054)
Only black	0.306	0.461	0.552	0.915	0.544	0.280
	(0.025)	(0.159)			(0.052)	(0.396)

Source: Data are from the 1900-1920, 1940 US Censuses (Ruggles et al., 2019), and PSID.
Notes: The Implied “True” column are the projections after eliminating measurement error, which is assumed to be classical for the IGE. Note that classical measurement error in the son’s percentile rank is not accounted for. For the 1968-1997 data, there are 697 total sons, 567 of which are white and 130 of which are black. For the early 20th century data, there are 394,864 sons, 387,922 of which are white and 6,942 of which are black. Income scores are used throughout, which is estimated to be the mean income by occupation, race and region. Dropping sons of fathers leads to 167,505 total fathers in the early 20th century, 165,649 of which are white and 1,856 of which are black.

Online Appendix

Table A1. Results are similar whether using a triple-linked sample (1900-1910-1920-1940) or a double-linked sample (1910-1920-1940)

	White 1 Father obs	Black and white 1 Father obs.	Black and white Mean of 2 father obs.	Black and white 2SLS
Panel A. 1940 log Income Score				
Main sample of fathers linked 1900-1910-1920	0.371	0.533	0.652	0.818
	(0.002)	(0.003)	(0.003)	(0.003)
N	387,922	394,864	394,864	394,864
Alt. sample of fathers linked 1910-1920	0.369	0.526	0.647	0.817
	(0.001)	(0.002)	(0.002)	(0.002)
N	804,939	825,251	825,251	825,251
Panel B. 1950 log occupational Income Score				
Main sample of fathers linked 1900-1910-1920	0.299	0.323	0.417	0.576
	(0.002)	(0.002)	(0.002)	(0.003)
N	387,922	394,864	394,864	394,864
Alt. sample of fathers linked 1910-1920	0.296	0.320	0.413	0.577
	(0.001)	(0.001)	(0.001)	(0.002)
N	804,939	825,251	825,251	825,251

Data: The main linked sample is from the 1900-1910-1920-1940 censuses, while the alternative linked sample is from 1910-1920-1940.

Notes: Both samples are weighted based on characteristics of the 1940 population.

Table A2. Estimating when imputing father's status with the surname average

	Individual	Surname Mean	2SLS benchmark
Panel A. Three father observations (1900-1920)			
1940 Income Score	0.707***	0.802***	0.818***
	(0.002)	(0.005)	(0.002)
1950 Occupational Income	0.458***	0.537***	0.577***
	(0.002)	(0.005)	(0.003)
Panel B. One father observation (1910)			
1940 Income Score	0.533***	0.667***	0.818***
	(0.002)	(0.004)	(0.002)
1950 Occupational Income	0.323***	0.408***	0.577***
	(0.002)	(0.004)	(0.003)

Data: The main linked sample is from the 1900-1910-1920-1940 censuses.

Notes: The "Individual" column contains the main estimates from the paper. The "Surname mean" estimate the association between the son's log income score and the average income score by surname. Surnames are grouped after using the NYSIIS algorithm to clean differences. The "2SLS benchmark" are the 2SLS estimates when instrumenting the 1910 father observation with the 1920 observation. The table shows that the group average method with surnames leads to similar estimates as the 2SLS estimate when averaging the average income score from three father observations.

Table A3. The imputed-income IGE overstates the actual-income IGE for sample of Iowans when using one father observation

	I	II	III	IV	V
No. of father observations	1	1	1	3	3
Father data	1915 Iowa	1910 US	1910 US	1900/20 US	1900/20 US
Son data	1940 US	1940 US	1940 US	1940 US	1940 US
Race in sample	White	White	Black and White	White	Black and white
Panel A: Same Iowa counties and cities in Feigenbaum (2018)					
Log actual income	0.208				
	(0.032)				
Log Occupational Income, 1950	0.441	0.411	0.407	0.558	0.556
	(0.021)	(0.022)	(0.022)	(0.024)	(0.024)
Log Income Score, 1940		0.391	0.398	0.590	0.599
		(0.027)	(0.027)	(0.029)	(0.029)
Observations		2,334	2,337	2,334	2,337
Panel B: Drop Sons of Farmers					
Log actual income	0.301				
	(0.037)				
Log Occupational Income, 1950	0.229	0.135	0.133	0.307	0.311
	(0.027)	(0.031)	(0.031)	(0.038)	(0.038)
Log Income Score, 1940		0.203	0.228	0.364	0.400
		(0.028)	(0.031)	(0.034)	(0.038)
Observations		1,254	1,257	1,254	1,257

Notes: Column I estimates are from Feigenbaum (2018). Columns II-V use my main sample when limited to the same counties in Iowa as in Goldin and Katz (2000).

Table A4. Alternative ways of imputing income or measuring status

Race	White	White	White	Black	Black	Black	Pooled	Pooled	Pooled
No. of father observations	1	3	2SLS	1	3	2SLS	1	3	2SLS
Income Score, 1940	0.371	0.547	0.685	0.304	0.522	0.771	0.533	0.707	0.818
	(0.002)	(0.002)	(0.003)	(0.018)	(0.019)	(0.040)	(0.003)	(0.003)	(0.003)
Occupational Score, 1950	0.299	0.425	0.539	0.166	0.304	0.543	0.323	0.458	0.577
	(0.002)	(0.002)	(0.003)	(0.017)	(0.022)	(0.056)	(0.002)	(0.002)	(0.003)
Occ. Inc. Score, 1950, adjusted (race/region)	0.357	0.488	0.591	0.325	0.482	0.692	0.489	0.621	0.708
	(0.002)	(0.002)	(0.003)	(0.016)	(0.017)	(0.033)	(0.003)	(0.003)	(0.003)
Occ. Inc. Score, 1890	0.334	0.449	0.536	0.273	0.388	0.543	0.354	0.472	0.560
	(0.001)	(0.002)	(0.002)	(0.013)	(0.015)	(0.028)	(0.002)	(0.002)	(0.002)
Occ. Inc. Score, 1890, adjusted (race/region)	0.373	0.483	0.557	0.387	0.509	0.660	0.461	0.574	0.641
	(0.001)	(0.002)	(0.002)	(0.012)	(0.013)	(0.022)	(0.002)	(0.002)	(0.003)
Occ. Inc. Score, 1890-1950	0.251	0.337	0.402	0.192	0.273	0.373	0.271	0.361	0.429
	(0.001)	(0.001)	(0.002)	(0.014)	(0.016)	(0.028)	(0.001)	(0.002)	(0.002)
Occ. Inc. Score, 1890-1950, adjusted (race/region)	0.295	0.383	0.439	0.334	0.437	0.561	0.396	0.493	0.547
	(0.001)	(0.001)	(0.002)	(0.013)	(0.015)	(0.023)	(0.002)	(0.002)	(0.002)
Occ. Inc. Score, 1901	0.130	0.226	0.338	0.088	0.151	0.174	0.189	0.314	0.451
	(0.002)	(0.003)	(0.005)	(0.014)	(0.019)	(0.049)	(0.002)	(0.003)	(0.005)
Occ. Inc. Score, 1901, adjusted	0.170	0.289	0.454	0.111	0.188	0.232	0.384	0.562	0.753
	(0.002)	(0.002)	(0.005)	(0.013)	(0.017)	(0.038)	(0.003)	(0.004)	(0.006)
Duncan Socioeconomic Index	0.326	0.485	0.662	0.093	0.148	0.190	0.370	0.544	0.736
	(0.002)	(0.002)	(0.004)	(0.019)	(0.027)	(0.068)	(0.002)	(0.002)	(0.004)
Occ. Education Score, 1950	0.168	0.251	0.548	0.082	0.092	0.172	0.176	0.260	0.562
	(0.014)	(0.016)	(0.058)	(0.028)	(0.044)	(0.079)	(0.014)	(0.016)	(0.056)
Occ. Wealth Score, 1850-1870	0.236	0.393	0.561	0.090	0.150	0.180	0.265	0.436	0.615
	(0.002)	(0.003)	(0.005)	(0.017)	(0.025)	(0.071)	(0.002)	(0.003)	(0.005)
Hauser-Warren Socioeconomic Index	0.279	0.415	0.567	0.111	0.183	0.235	0.300	0.445	0.604
	(0.003)	(0.003)	(0.005)	(0.018)	(0.026)	(0.063)	(0.003)	(0.003)	(0.005)
Occ. Human Capital rank from Song et al. (2019)	0.324	0.473	0.660	0.110	0.197	0.360	0.349	0.506	0.702
	(0.002)	(0.002)	(0.004)	(0.016)	(0.022)	(0.056)	(0.002)	(0.002)	(0.004)
Observations	387,922	387,922	387,922	6,942	6,942	6,942	394,864	394,864	394,864

Notes: Data are from the 1900-1910-1920-1940 linked sample. Each cell is from a different regression. The 2SLS column instruments the 1910 observation with the 1920 observation. When using income or occupational scores, they are logged. Note that while for the main results I always rank within the dataset, following Chetty et al. (2014), Song et al. (2019) rank the occupations for the population and then apply them to individuals in their data. Since Song et al. do not rank within the linked data, I also do not when using their score. Therefore, when averaging father observations, I take the simple average and do not re-rank within the data. This leads to a greater increase in the association than found for ranks in the main text.

Table A5. Main results robust to alternatively linked sample

	White 1 Father obs	Black and white 1 Father obs.	Black and white Mean of 3 father obs.	Black and white 2SLS
Panel A. 1940 log Income Score				
Main sample	0.371	0.533	0.707	0.818
	(0.002)	(0.003)	(0.003)	(0.003)
ABE (NYSIIS), Unique ± 2 years of birth	0.376	0.539	0.713	0.827
	(0.003)	(0.004)	(0.004)	(0.005)
ABE (Exact), Unique ± 2 years of birth	0.377	0.540	0.711	0.820
	(0.003)	(0.004)	(0.004)	(0.005)
Panel B. 1950 log occupational Income Score				
Main sample	0.299	0.323	0.458	0.577
	(0.002)	(0.002)	(0.002)	(0.003)
ABE (NYSIIS), Unique ± 2 years of birth	0.307	0.330	0.466	0.588
	(0.002)	(0.003)	(0.003)	(0.004)
ABE (Exact), Unique ± 2 years of birth	0.309	0.332	0.464	0.582
	(0.002)	(0.003)	(0.003)	(0.005)

Notes: Data are a linked sample of fathers of sons from the 1900-1910-1920-1940 United States Censuses (Ruggles et al., 2019). The table recreates the results in Figure 1, but with a linked sample based on the iterative linking method described by Abramitzky et al., (2012) where only those who are unique in plus/minus 2 year of birth range are kept. There are 394,864 total observations in the main sample, with 387,922 white father-son pairs and 6,942 black father-son pairs. There are 209,251 total observations in the ABE NYSIIS sample, with 204,869 white father-son pairs and 3,953 black father-son pairs. There are 207,803 total observations in the ABE Exact name sample, with 203,550 white father-son pairs and 4,253 black father-son pairs. I block on race for the ABE match, in addition to name (exact or NYSIIS), and birthplace.

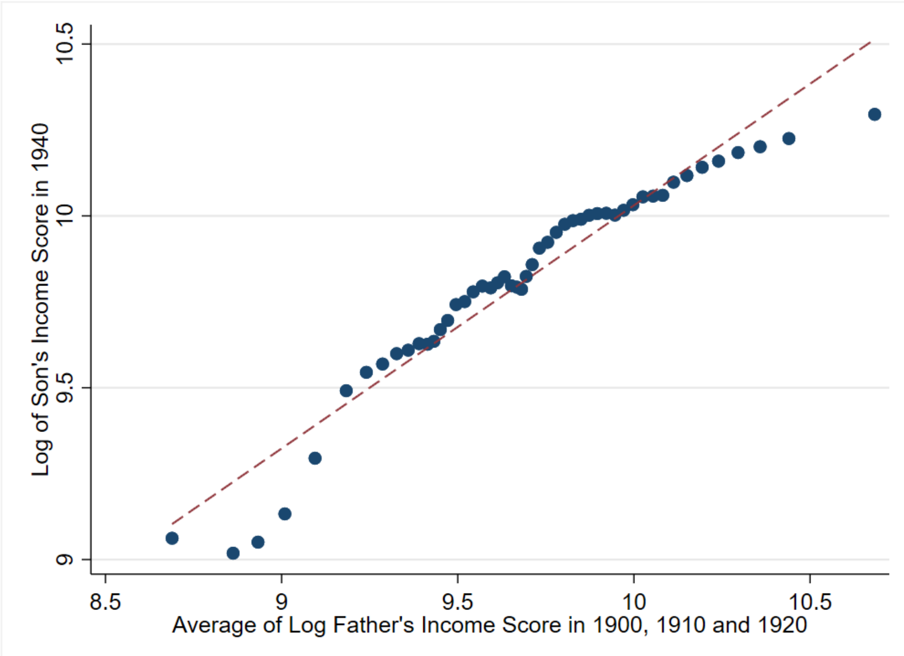
Table A6. Trend in mobility over time, alternative measures of status

	1910-1940			1968-1997		
No. of father observations	1	2	2SLS	1	2	2SLS
Income Score, 1940	0.533	0.652	0.818	0.533	0.558	0.570
	(0.003)	(0.003)	(0.003)	(0.050)	(0.052)	(0.055)
Income Score, 1940 for early, 2000 for PSID	0.533	0.652	0.818	0.438	0.463	0.500
	(0.003)	(0.003)	(0.003)	(0.044)	(0.045)	(0.051)
Occupational Score, 1950	0.323	0.417	0.577	0.305	0.321	0.352
	(0.002)	(0.002)	(0.003)	(0.048)	(0.051)	(0.057)
Occ. Inc. Score, 1890	0.354	0.432	0.560	0.335	0.351	0.383
	(0.002)	(0.002)	(0.002)	(0.047)	(0.048)	(0.053)
Duncan Socioeconomic Index	0.370	0.489	0.736	0.333	0.346	0.371
	(0.002)	(0.002)	(0.004)	(0.041)	(0.044)	(0.048)
Hauser-Warren Socioeconomic Index	0.300	0.397	0.600	0.423	0.449	0.483
	(0.003)	(0.003)	(0.005)	(0.045)	(0.048)	(0.052)
Log income*	0.601	0.739	0.931	0.412	0.535	0.662
	(0.004)	(0.004)	(0.006)	(0.072)	(0.073)	(0.103)
Occ. human capital rank (Song et al., 2019)	0.349	0.462	0.702	0.373	0.407	0.445
	(0.002)	(0.002)	(0.004)	(0.038)	(0.040)	(0.046)
Observations	394,864	394,864	394,864	697	697	697

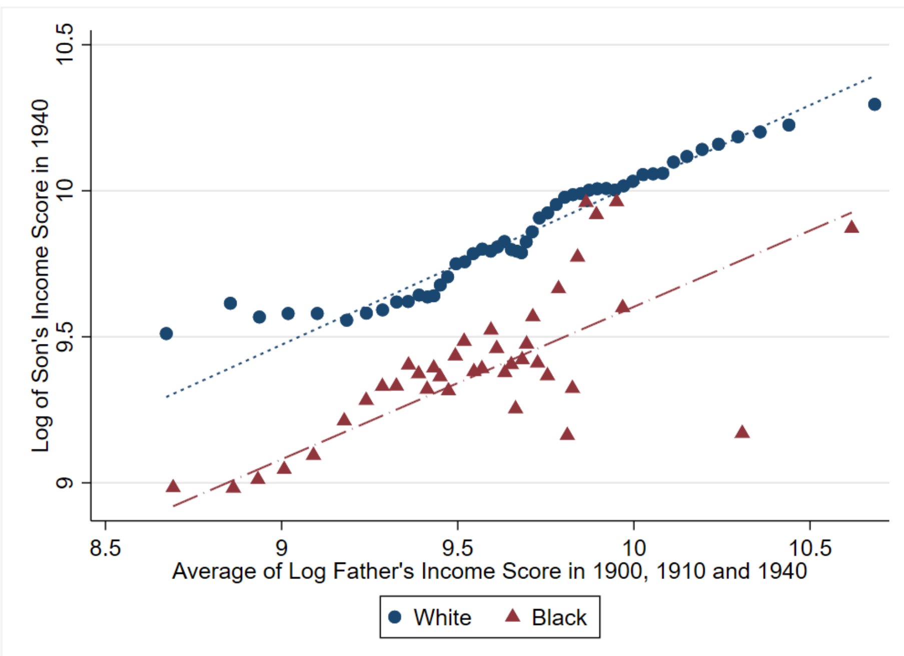
Notes: Data are from the PSID and from a linked sample of the 1900, 1910, 1920 and 1940 Censuses. There are up to 15 missing observations of fathers linked to sons for the PSID that do not have an assigned score due to missing information in the Duncan Socioeconomic Index, Harren-Warren Socioeconomic Index, and occupation human capital rank. This is primarily because there are no matching 3-digit occupation codes.
*Log income is not truly log income, but it more closely captures income than the income score. For this measure, I use wage income in 1940 for wage workers instead of the 1940 income score. I continue to use the 1940 income score for self-employed workers. Since wage income is unavailable in 1940, I use the income score for fathers. Note that 18,189 sons are dropped from this regression due to missing or zero wage income. For the PSID I use total family income.

Figure A1. Log-log associations appear to be linear

Panel A. Pooled



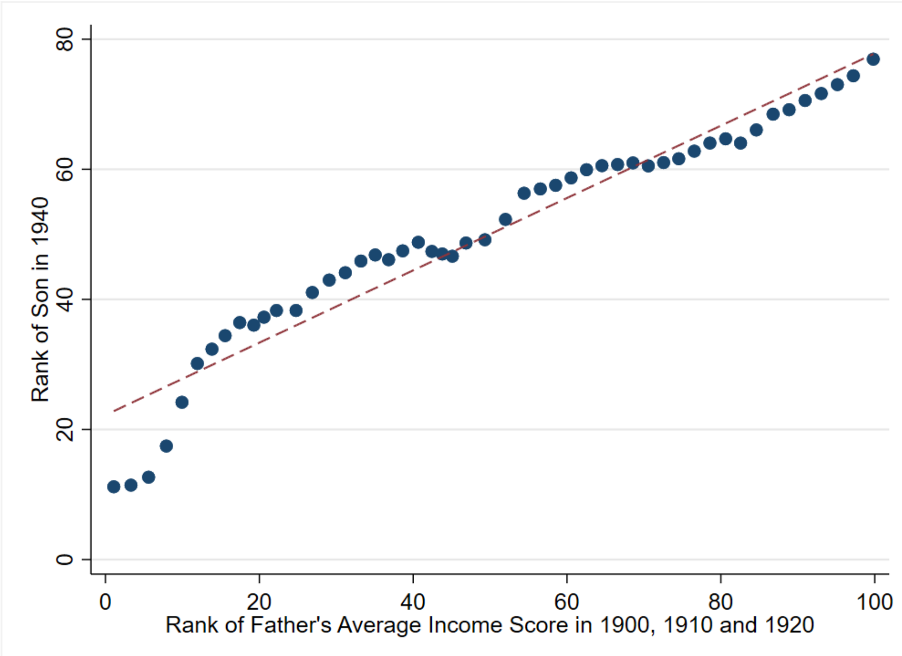
Panel B. Within race



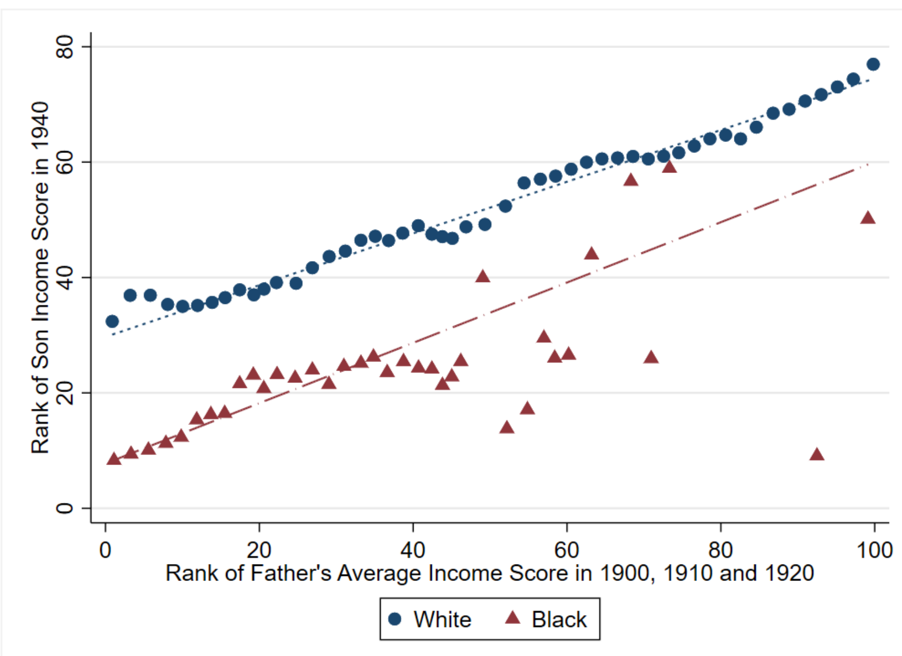
Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: The figure shows the IGE bin scatter plot when using the average of three father observations.

Figure A2. Rank-rank associations appear to be linear

Panel A. Pooled

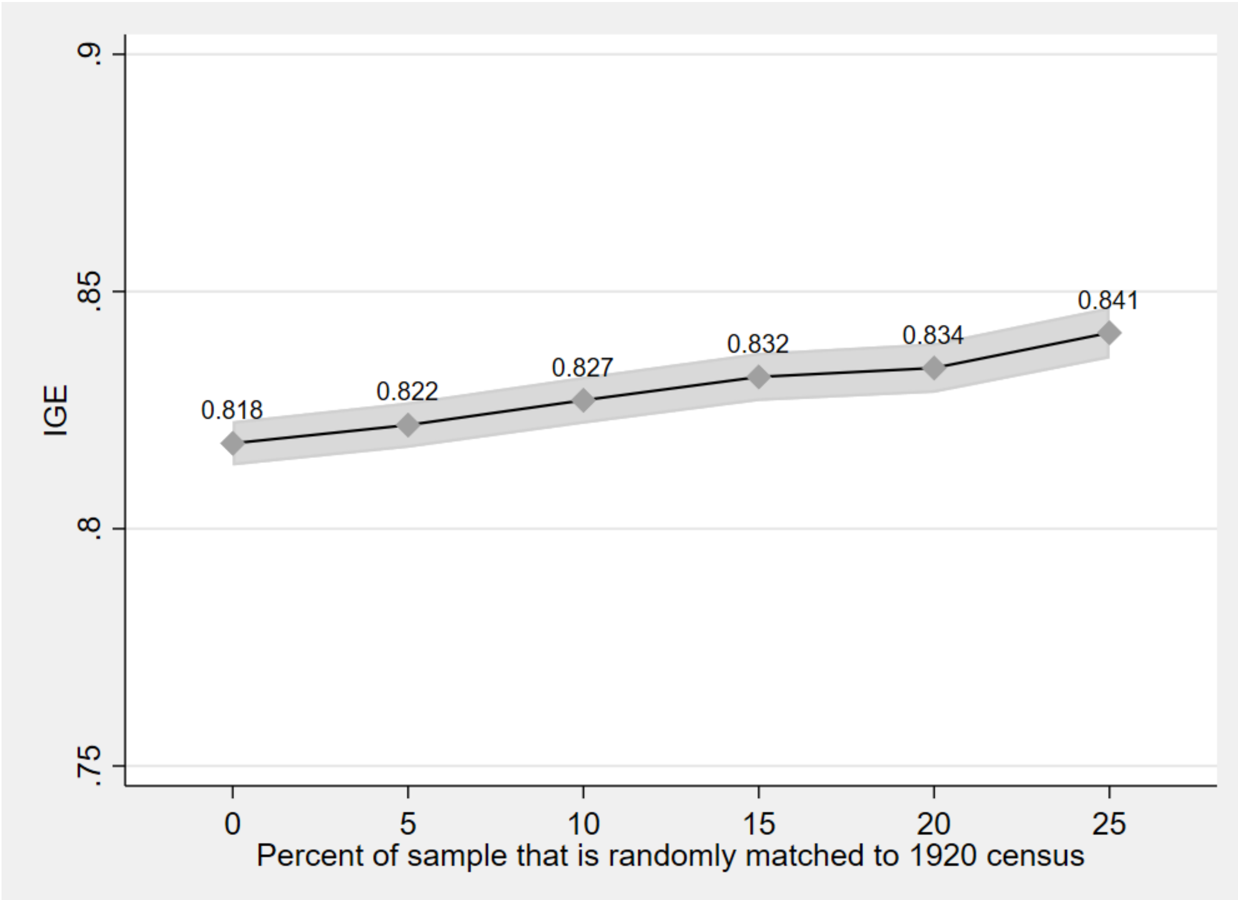


Panel B. Within race



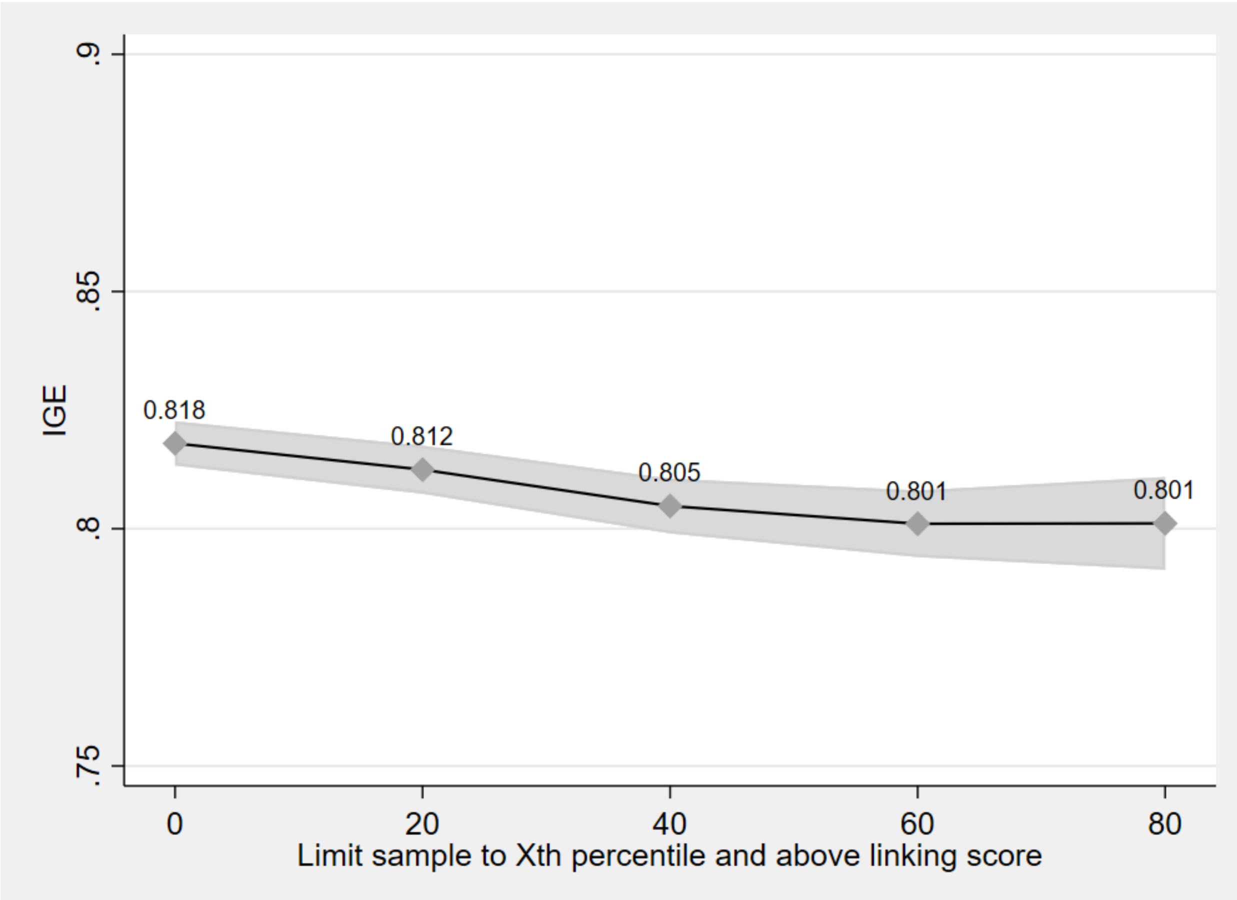
Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: The figure shows the rank-rank bin scatter plot when using the average of three father observations.

Figure A3. Falsely linking more fathers between 1910 and 1920 does not strongly affect the IGE



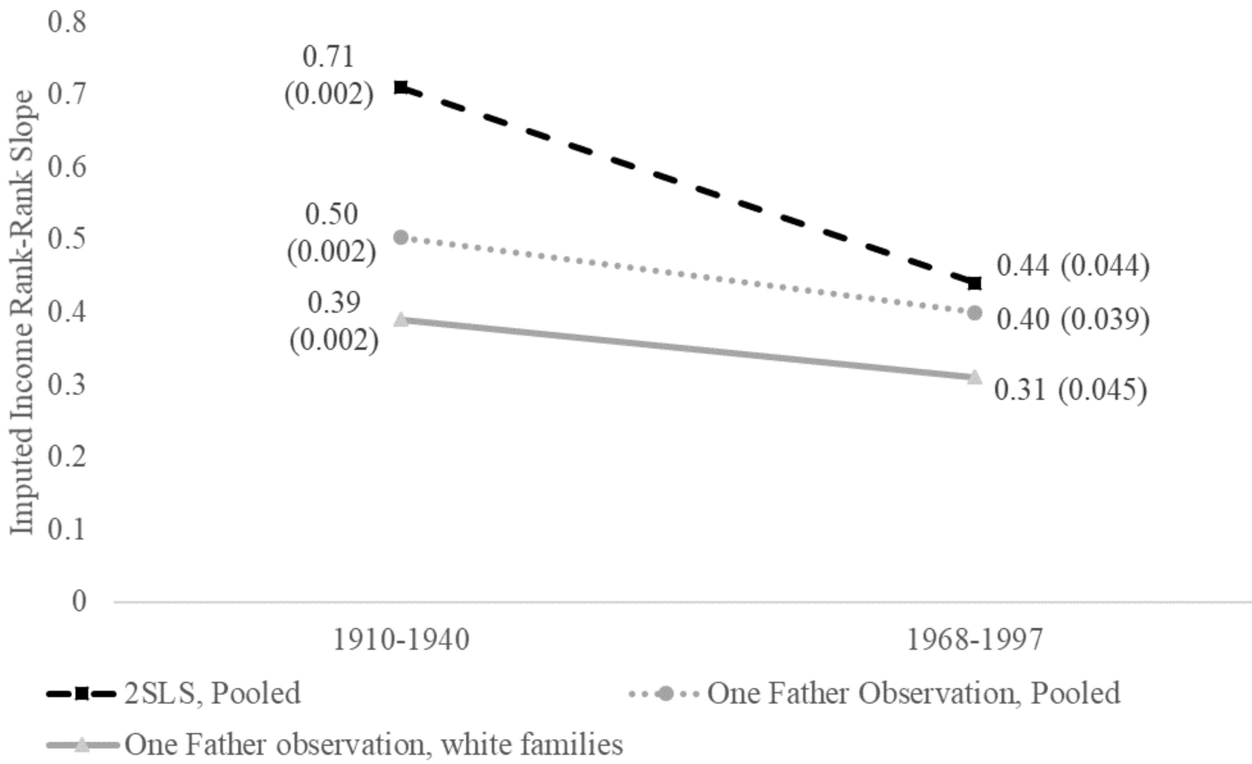
Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: The figure shows how replacing actual links with false links influences the 2SLS estimate of persistence between father and son. Recall that the 2SLS estimate instruments the 1910 father observation with the 1920 father observation. The 0.818 estimate for zero percent of the sample is the same estimate from the main paper. The 0.841 estimate for 25 percent comes from randomly replacing 25 percent of the main sample of 1910 fathers with random links to the 1920 census after requiring age, race and birth place to match exactly.

Figure A4. Restricting the data to higher-quality links does not strongly affect the IGE



Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: The figure shows how restricting the main sample to higher-quality links influences the 2SLS estimate of persistence between father and son. Recall that the 2SLS estimate instruments the 1910 father observation with the 1920 father observation. The 0.818 estimate for the zeroth percentile and above is the same estimate from the main paper. The 0.800 estimate for the 80th percentile and above comes after restricting the sample to the top 20th percent of linking scores for the white population and for the black population, as predicted from the probit model. Note that the 2SLS estimate may change not only due to reduced measurement error, but also because the sample composition has changed.

Figure A5. The trend in the income score rank-rank slope over 20th century

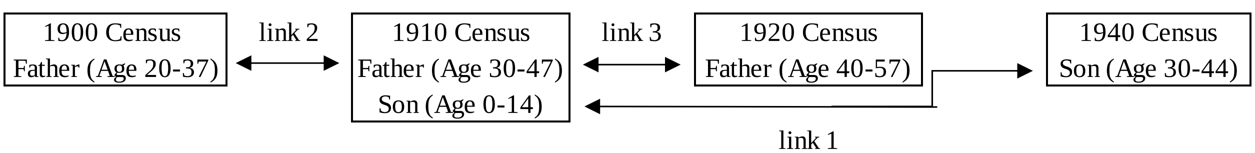


Note: This figure recreates Figure 4 from the main paper but reports the rank-rank slope. See Table 6 in main text for underlying coefficients.

Appendix B. Details on linking data

I combine three different linked datasets in this paper: 1910-1940 (sons from childhood to adulthood); 1910-1920 (fathers to another observation); and 1910-1900 (fathers to another observation) – see Figure B1. The first link (1910-1940) is the common way to build an intergenerational dataset in historical studies: take sons and fathers from the same household in 1910 and then link the sons forward to the 1940 census to get his adult occupation. I take this linked data from Kosack and Ward (2018), who built it to estimate mobility gaps across Anglo, African and Mexican Americans between 1910 and 1940. The second and third links take the fathers in 1910 and find them in both the 1900 and 1920 censuses. I will describe the 1900-1910 and 1910-1920 links in detail but will not do so for the 1910-1940 link since this link is fully explained in Kosack and Ward's (2018) Appendix B. (Note that I also use the 1910-1920 link in Ward (2019a).) All links are made with the same method (that is, based on Feigenbaum (2016)), so describing the 1910-1920 and 1900-1910 links is sufficient to understand the 1910-1940 link.

Figure B1. Linking Process to build dataset



Building the set of potential matches.

I build new datasets of US-born whites and US-born blacks by linking the 1900-1910 and 1910-1920 censues. I use the same broad strategy as in Feigenbaum (2016) where I build a set of potential links, handlink a subset of them, and then train a probit to pick the best link.

I first extract the entire set of US-born white and black males who are over 10 and under 40 years of age in both 1900 and 1910. After dropping those with the exact same combinations of first name string, last name string, race, state of birth and year of birth, I then search for all possible combinations in the census ten years later that meet the following criteria

1) First letter of first name match
2) First letter of last name match
3) Jaro-Winkler distance of first name is less than 0.20

4) Jaro-Winkler distance of last name is less than 0.20
5) Year of birth is less than three years in difference
6) State of birth and race match exactly

The first two criteria differ from Feigenbaum (2016), who does not block on first letters of last or first name; I keep these criteria to reduce computing costs and keep the matching process managable when matching complete to complete-count censuses. The race match requirement also misses some matches because race identification may change between censuses; therefore, the US-born black results only apply to fathers and sons listed as black in all censuses. Finally, I do not block on mother or father's state of birth because there appears to be some error in how these variables are recorded, perhaps because another person of the household was answering the enumerator for the entire household. However, mother and father's state of birth is useful for choosing the best matches so I will incorporate it into the probit model.

Based on these linking criteria, not everyone in the starting census has a potential match in the second census. For the white population, about 70 to 80 percent of the starting sample has a possible match ten years later; for the black poulation, only about 60 percent of the starting sample has a possible match. The different rates for the black and white population may reflect differential mortality between the two groups, or that true matches in the black population are less likely to meet the above criteria. Either way, the results suggest that the maximum linking rate is not near 100 percent even if I could find a true link among the set of potential matches. However, I first need to determine which of the potential matches is the true link.

Choosing the best link.

After creating the set of potential matches, I draw a sample of 2,000 black and 2,000 white individuals and all of their potential matches in the later census. I do this each for the 1900-1910 match and the 1910-1920 match, where the 2,000 are drawn from the starting census. These four datasets will form the basis of the training data, but first I need to handlink the people in the dataset.

From the dataset of potential matches, I handpick which is the best match. If there are two close potential links that look similar to the original link, then I do not pick a match since I am not confident which one is the true link. The matching rates for the training data are given in Table B1. After going through this handlinking process, I am able to find a true link for 53 to 63 percent of the white population with at least one match, and 42 to 52 percent of the black population with at least one potential match. Part of the reason I fail to find a link for all of the training data are because none of the potential links are close in names or year of birth; part of the reason is because there are multiple good matches. The 1900-1910 match has lower linking rates for both the black and white population, which may reflect that earlier census data was of lower quality.

Table B1. Details for the handlinked dataset

	1900-1910		1910-1920	
	White	Black	White	Black
Random sample in base year	2,000	2,000	2,000	2,000
Potential links ten years later	17,133	8,320	15,993	7,763
Successfully linked	1,073	911	1,263	1,030
Handlinking Rate for training data (given 1 potential match)	53.7	42.2	63.2	51.5

With the training dataset of potential links and actual links in hand, I model the true link as a function of observable differences between matches. I include the Jaro-Winkler distance in the first and last name; absolute difference in year of birth; number of potential links and its square; mother's place of birth and father's place of birth. I also include information on whether there are unique and exact matches for either the first or last name in terms of NYSIIS codes or exact string match; this is based on the handlinking process where having the same last name that was unique (that is, no other potential links has the same last name) was a strong predictor of a link. The probit models for each of the 1900-1910 and 1910-1920 matches, separately by black and white, are shown in Table B3.

The probit models give a predicted match score for each potential link in the training dataset. From this information, I set two tuning parameters to determine who will be included in my linked dataset. The first parameter is the cut off for predicted probability, where a potential link needs to have a predicted probability above this level to be included in the linked dataset. The second parameter is the ratio of the 1st best probability to the 2nd best probability; this ensures that I do not keep a match that has a close alternative. I set these parameters to maximize the efficiency of the algorithm in terms of true positive rate (TPR, or the percentage of true links that I keep), as long as the positive predictive value (PPV) is at least 0.9. The positive predictive value is the ratio of true positives to total matches; viewed from the opposite direction, it sets the false positive rate to 10 percent. This false positive rate is slightly lower than Feigenbaum's (2016) training data in Iowa and thus is on the conservative end; however, one could easily change this parameter to be more or less restrictive. A consequence of the decision to limit false positives is that it reduces the matching rate for the full sample. See Table B3 for the tuning parameters and the resulting PPV and TPR.

Table B2. Predicting the handlinked match using a probit model

	1900-1910 White	1900-1910 Black	1910-1920 White	1910-1920 Black
Jaro-Winkler Distance, First name	-6.989***	-5.912***	-6.566***	-6.318***
	(0.650)	(0.582)	(0.589)	(0.555)
Jaro-Winkler Distance, Last name	-14.79***	-13.14***	-13.57***	-12.98***
	(0.955)	(0.995)	(0.876)	(0.948)
Year of Birth Difference = 1	-0.439***	-0.156	-0.158	-0.308*
	(0.130)	(0.169)	(0.126)	(0.159)
Year of Birth Difference = 2	-1.029***	-0.377**	-0.760***	-0.585***
	(0.156)	(0.169)	(0.154)	(0.160)
Year of Birth Difference = 3	-1.396***	-0.714***	-1.161***	-0.826***
	(0.172)	(0.180)	(0.168)	(0.168)
No. of potential links	-0.0673***	-0.101***	-0.0623***	-0.164***
	(0.0197)	(0.0225)	(0.0190)	(0.0229)
No. of potential links squared	0.00103	0.00211**	0.00143**	0.00468***
	(0.000678)	(0.000936)	(0.000655)	(0.000940)
Unique and Exact NYSIIS First name match	0.365**	0.113	0.424***	0.162
	(0.172)	(0.128)	(0.161)	(0.119)
Unique and Exact NYSIIS Last name match	0.0892	0.892**	-0.0570	0.948**
	(0.304)	(0.380)	(0.266)	(0.379)
Unique and Exact NYSIIS First AND Last name match	0.634***	0.762***	0.948***	0.897***
	(0.140)	(0.120)	(0.132)	(0.119)
Unique Exact Last name String match	0.777***	0.205	1.288***	0.0764
	(0.200)	(0.224)	(0.219)	(0.250)
Middle initial match, if have one	1.095***	0.596*	1.191***	1.188***
	(0.107)	(0.337)	(0.110)	(0.332)
NYSIIS last name match AND Year of Birth Diff=0	1.067***	0.727**	1.031***	0.268
	(0.223)	(0.321)	(0.196)	(0.308)
NYSIIS last name match AND Year of Birth Diff=1	1.009***	0.659**	0.867***	0.523*
	(0.226)	(0.300)	(0.182)	(0.297)
NYSIIS last name match AND Year of Birth Diff=2	0.875***	0.291	0.732***	0.225
	(0.248)	(0.300)	(0.207)	(0.299)
2 Potential links with NYSIIS last name match	-0.707***	-0.585**	-0.387**	-0.739***
	(0.203)	(0.250)	(0.182)	(0.247)
>2 potential links with NYSIIS last name match	-1.061***	-0.651**	-0.525***	-0.398
	(0.212)	(0.291)	(0.167)	(0.281)
2 Potential links with last name string match	-0.636***	-0.585***	-1.425***	-0.476**
	(0.169)	(0.198)	(0.194)	(0.215)
>2 Potential links with last name string match	-1.246***	-1.314***	-1.354***	-1.276***
	(0.130)	(0.147)	(0.125)	(0.142)
One potential link	0.780***	0.638***	0.790***	0.650***
	(0.183)	(0.130)	(0.171)	(0.121)
Difference in length of last name strings	-0.343***	-0.498***	-0.336***	-0.552***
	(0.0570)	(0.0711)	(0.0499)	(0.0683)
Mother place of birth match	0.258***	0.258**	0.537***	0.249**
	(0.0780)	(0.107)	(0.0799)	(0.107)
Father place of birth match	0.426***	0.278***	0.520***	0.171*
	(0.0779)	(0.101)	(0.0781)	(0.101)
Constant	1.128***	0.741***	0.458**	1.299***
	(0.189)	(0.226)	(0.185)	(0.227)
Observations	15,993	8,320	15,993	7,763

Notes: Data are from the handlinked sample between 1900-1910 or 1910-1920. The coefficients are from a probit model that predicts the correct link.

Table B3. Tuning parameters for determining who to keep in the linked sample

Census Years	Race	Cutoff for predicted probability	Score Ratio of 1st best link to 2nd best	PPV	TPR
1900-1910	White	0.383	2.1	0.900	0.814
	Black	0.518	4.4	0.901	0.673
1910-1920	White	0.383	2.1	0.900	0.814
	Black	0.518	4.4	0.901	0.673

Notes: PPV stands for positive predictive value and gives the ratio of true positives to all links. TPR stands for true positive rate and gives the proportion of true links that would appear in the final linked dataset.

I then predict the linking scores for the full to full count match with the probit model; afterwards, I keep only those who meet the parameters set in Table B3. See Table B4 for the linking rates when applying this process to the full-count data. I link of 29 to 33 percent of the white population, and 13 to 15 percent of the black population. These linking rates are lower than Feigenbaum's link from the 1915 Iowa Census to the 1940 Federal Census of near 60 percent. This may be due to a number of reasons: because Iowa is a smaller state and thus has fewer other potential matches, because the data quality is higher from Iowa, because modelling the hand linking process is easier for Iowans versus the rest of the country, or because there are lower mortality rates for Iowans relative to the rest of the country. While the linking rate is somewhat low, I still have millions of individuals linked across censuses.

Table B4. Applying the probit model to the full 1910-1920 link, details

	1900-1910 Census		1910-1920 Census	
	White	Black	White	Black
Starting group in base year	15,353,841	2,377,438	18,524,622	2,745,128
Starting group in base year with a potential link in ten years later	11,184,130	1,307,528	15,448,111	1,581,988
Potential links ten years later	122,813,248	5,508,884	152,390,867	6,262,320
Linked	4,403,006	298,212	6,113,276	400,439
Overall Linking Rate	28.7	12.5	33.0	14.6
Linking Rate given Potential Match	39.4	22.8	39.6	25.3

Getting into the sample used in the main analysis

To be included into the final linked sample used in this paper, a father must be in both the 1900-1910 and 1910-1920 link, and the son first observed in 1910 must be found in 1940. Essentially this means a father-son observation must survive being triple linked. Beyond these criteria, I want fathers to be at the midpoint of their lifecycle throughout the sample. Therefore, I limit the sample such that fathers are between 30 and 47 years old in 1910; this implies that they are between 20 and 37 in 1900 and between 40 and 57 in 1920.[48] Finally, I keep individuals where there is an occupational response for the father all three censuses (1900, 1910 and 1920) and for the son in 1940.

The resulting sample is of 394,864 sons linked to 320,168 fathers. This number is only 4.6 percent of the 1910 sons that I could have possibly linked to the 1940 census. Given that the general linking rate of two censuses is around 25 percent, it would be expected that about (0.25)(0.25)(0.25) = 1.6 percent of individuals would be linked three times. The actual linking rate is higher than 1.6 percent since being successfully linked is not independent across censuses.

Weighting

Only a select group (4.6 percent) of the original population shows up in the triple-linked sample. Therefore, this group may be unrepresentative of the original population and provide misleading information on the convergence of economic gaps. I address this problem by reweighting the data to be representative of the population. To weight the data, I match the 1940 sons' outcomes to the 1940 census in terms of age, high school degree, region and race.[49]

The representativeness of the sample is shown in Table B5. There is selection into the linked sample, where sons with white-collar jobs and farmers are more likely to be in the sample than unskilled or semi-skilled fathers. Further, those in the Midwest and West are more likely to be in the sample than those in the South or in the Northeast. Therefore, estimating the mobilities using the unweighted data will erroneously reflect Midwestern rural states like Iowa, rather than

[48] The original linking process limits those to age 10 to 40 in 1910, but I further link those up to 47 years old using the same predicted match scores and rules from Tables B2 and B3.

[49] I have also used the inverse proportional weighting process, as suggested by Bailey et al. (2019). This involves pooling the linked sample with the linkable sample, and estimating which observables are associated with being successfully linked. Let q be the share of linked records and p be the predicted probability. The weight is $[(1-p)/p] \times [q/(1-q)]$.

the full population. The weighted representative characteristics are also shown in Table B5, which still shows small differences with the 1940 population in terms of age, years of education and being a farmer or skilled worker.

Table B5. Representativeness of the linked sample based on 1940 characteristics

	I	II	III	IV
	1940 Census	Linked Sample		Difference
		Unweighted	Weighted	(III-I)
Black	0.10	0.02	0.10	0.00
	(0.30)	(0.13)	(0.30)	(0.00)
Age	36.52	36.55	36.55	0.03***
	(4.30)	(4.17)	(4.30)	(0.00)
Log Income Score, 1940	9.82	9.90	9.81	-0.01***
	(0.52)	(0.48)	(0.53)	(0.00)
High School Degree	0.26	0.34	0.26	0.00
	(0.44)	(0.47)	(0.44)	(0.00)
Years of Education	8.60	9.90	9.25	0.65***
	(3.76)	(3.25)	(3.32)	(0.00)
White Collar	0.30	0.34	0.30	-0.00***
	(0.46)	(0.47)	(0.46)	(0.00)
Farmer	0.13	0.15	0.16	0.03***
	(0.33)	(0.36)	(0.36)	(0.00)
Unskilled	0.22	0.18	0.22	-0.00***
	(0.42)	(0.39)	(0.42)	(0.00)
Skilled	0.35	0.32	0.32	-0.03***
	(0.48)	(0.47)	(0.47)	(0.00)
Northeast	0.25	0.21	0.25	0.00
	(0.43)	(0.41)	(0.43)	(0.00)
Midwest	0.32	0.41	0.32	0.00
	(0.46)	(0.49)	(0.46)	(0.00)
South	0.32	0.24	0.32	0.00
	(0.47)	(0.43)	(0.47)	(0.00)
West	0.11	0.14	0.11	0.00
	(0.31)	(0.35)	(0.31)	(0.00)
Observations	11,814,981	394,864	394,864	

Notes: Data are from the 1940 full-count census (column I) and from the linked sample (columns II and III). The difference in Column IV tests how the weighted linked sample's characteristics are different from the 1940 population's characteristics. Only those with reported education, occupations, and black/white individuals are included in the sample.

Appendix C. Decomposing the IGE into the within-group and between-group components

It is clear that adding black families to the historical IGE matters, partially because the black-white gap was large and converged slowly (Margo, 2016). It is possible to decompose the population-level IGE into the sum of between-group effects and within-group effects. Following Hertz (2008), the within-race contribution to the pooled elasticity can be written as:

$$\text{Within effect} = \frac{n_b}{n_{all}}\beta_b\left(\frac{\sigma_b^2}{\sigma_{all}^2}\right) + \frac{n_w}{n_{all}}\beta_w\left(\frac{\sigma_w^2}{\sigma_{all}^2}\right) \tag{C1}$$

where n_b/n_{all} is the black share of the population, β_b is the within-black elasticity, σ_b^2 is the population variance of black fathers' outcomes, and σ_{all}^2 is the population variance of all fathers' outcomes.[50] Subscript *w* indicates the white population.

The between-group contribution to the pooled elasticity can be written as

$$\text{Between effect} = \frac{\delta}{\sigma_{all}^2}\left(\frac{n_b}{n_{all}}[x_b - x_{all}]^2 + \frac{n_w}{n_{all}}[x_w - x_{all}]^2\right) \tag{C2}$$

where δ is the between-group IGE, and $x_b - x_{all}$ is the black-population gap in economic outcomes for the father's generation.

The decomposition method demonstrates that about half of the predicted "true" elasticity (0.84) comes from the within-group effect (0.44) and the other half comes from the between-group effect (0.40, see Panel A of Table C1). While one may expect that adding more father observations only increases the within-group effect, the between-group effect also increases. The reason is that the total amount of variation falls (σ_{all}^2) when eliminating the error component of income scores, which then inflates both the between- and within-group effects.

In contrast to the 1940 income score, the between-group effect for the 1950 occupational score is only 12 percent (or 0.07) of the pooled IGE, which is the main reason why the IGE is lower for the 1950 occupational score than for the 1940 income score. The lower between-group component for the 1950 occupational score can be entirely attributed to its smaller measured black-white gap for fathers (26 log points v. 79 log points). This finding reinforces that imputed income scores that better capture racial disparities lead to higher estimates of persistence.

[50] The population variance is used for convenience in notation rather than the sample variance. See Hertz (2008, footnote 1). In practice, this does not matter for the large sample.

It is possible to use the decomposition to address a few limitations in the data. For example, the 0.84 IGE may be overestimated because I impute income in 1910 with data from 1940 and therefore miss the convergence of black and white actual incomes from 1910 to 1940. Indeed, my data estimate that the black-white economic gap *widened* since the between-group elasticity is 1.03, which contrasts with estimates that the black-white income gap *converged* with an elasticity of 0.93 (Margo, 2016).[51] At the same time, if the black-white gap truly converged at 0.93 between 1910 and 1940, then I should also have measured a 7 percent wider black-white gap in 1910. Plugging these estimates into the decomposition suggests that the "true" 1910-1940 elasticity should be 0.81 instead of 0.84 – not that large of a difference from the one estimated in the data.[52]

Decomposing elasticities over time.

It is possible to decompose the elasticities over time to understand how between- and within-group effects have changed. Table C2 shows that that IGE fell over time primarily because the between-group component dropped in magnitude, which occurs due to a smaller black-white gap in the PSID than in the earlier census data (47 percent instead of 55 percent). Therefore, the trend in the IGE is a function of how far black families need to jump to overcome the initial disparities.[53] Otherwise, if mobility was truly constant over time (e.g., Clark, 2014), then to offset the decline in the between-group component, the within-group component would need to have increased. For example, assume that mobility was constant over the early 20th and late 20th century such that I should have estimated a 0.84 "true" elasticity in the PSID. For this result to hold, alongside a smaller black-white gap in 1968, then the within-race component would have had to increase by about 75 percent (from 0.34 to 0.60). Therefore, for population-level mobility to have been constant over time, then within-group effects would have needed to increase to offset the converging racial and ethnic averages over the past 150 years (Margo, 2016; Ward, 2019a).

---

[51] Margo (2016) estimates the black-white income gap was 0.32 in 1900 and 0.38 in 1940. If one linearly interpolates these estimates, then the convergence between 1910-1940 was from 0.335 to 0.38. Based on these numbers then between-group elasticity should be (1-0.38) / (1-0.335) = 0.932.

[52] To update the total variation in the data given a wider black-white gap in 1910, I adjust all black fathers' income scores downward and re-estimate the total amount of variation in the data.

[53] Collins and Wanamaker (2017) argue that that black-white mobility gap has been mostly constant between 1880 and 2000 – when mobility is measured with percentile ranks. However, since the actual black-white gap was larger in the late 19th and early 20th centuries, then a given black-white percentile rank gap is more difficult to overcome in the past than for today.

Table C1. Decomposing the population IGE into between-race and within-race components

	Pooled	Between effect	Within Effect	Components of between effect: Black-white gap	Between-elasticity	Total variation	Components of within effect: Black elas.	White elas.	Black variation over total	White variation over total
Panel A. IGE, 1940 Income Score										
Predicted "True"	0.839	0.400	0.439	-0.793	1.041	0.149	0.797	0.722	0.200	0.650
Three father observations	0.707	0.340	0.367	-0.793	1.041	0.172	0.522	0.547	0.269	0.718
One father observation	0.533	0.257	0.276	-0.802	1.029	0.230	0.304	0.371	0.354	0.794
Panel B. IGE, 1950 Occupational Income Score										
Predicted "True"	0.588	0.068	0.519	-0.267	1.224	0.118	0.516	0.554	0.269	1.014
Three father observations	0.458	0.056	0.401	-0.267	1.224	0.138	0.304	0.425	0.378	1.018
One father observation	0.323	0.040	0.283	-0.274	1.190	0.198	0.166	0.299	0.472	1.020
Panel C. Predicted IGEs										
Black-white converges like Margo (2016)	0.811	0.407	0.404	-0.879	0.932	0.161	0.797	0.723	0.184	0.599

Notes: Data are a linked sample of fathers and sons from 1900, 1910, 1920 and 1940 United States Censuses. Decomposition according to Hertz (2008). The first counterfactual assumes that the black-white gap converged between 1910 and according to Margo (2016) and thus changes the between-group elasticity. The predicted 1870-1900 elasticity uses data for the black-white gap and between-group elasticity from Margo (2016). Further, it decreases the within-race elasticity by 19 percent from the 1910-1940 period based on results from Olivetti and Paserman (2015)

Table C2. Decomposing the "true" IGE over time

				Components of between effect			Components of within effect			
	Pooled	Between effect	Within Effect	Black-white gap	Between-elasticity	Total variation	Black elasticity	White elasticity	Black variation in scores over total	White variation in scores over total
Panel A. IGE										
1910-1940	0.839	0.400	0.439	-0.793	1.041	0.149	0.797	0.722	0.200	0.650
1968-1997	0.585	0.241	0.344	-0.635	0.820	0.115	0.209	0.509	0.667	0.717
Panel B. IGE, no sons of farmers										
1910-1940	0.670	0.338	0.332	-0.889	0.938	0.127	0.565	0.520	0.609	0.634
1968-1997	0.581	0.241	0.340	-0.636	0.821	0.115	0.208	0.504	0.666	0.715

Source: Data are from the 1900-1920, 1940 US Censuses (Ruggles et al., 2019), and PSID.
Notes: The pooled elasticity is decomposed into the sum of the between effect and within effect, according to the formulas in Hertz (2008). The black-white gap is estimated for fathers. The variation in income scores are estimated for fathers.

Appendix D. The influence of measurement error and racial disparities on the Altham Statistic

Perhaps the most important methodological choice is that I use imputed income to measure economic status. This method allows me to place people on a univariate scale such that I can calculate an IGE or rank-rank slope. Sorting people on a univariate scale is directly criticized by Long and Ferrie (2013b) since it is unclear how well income imputations capture actual income further back in time. This criticism is more important when one goes as far back as Ferrie (2005) and Long and Ferrie (2013a) do – that is, back to 1850. However, I am less concerned about this criticism for my early 20th century data since occupational income was strongly correlated between 1890 and 1950 (Sobek, 1996).

Nevertheless, it is important to understand how alternative measures of mobility that do not rely on imputed income change when accounting for race and measurement error. One way to measure mobility is based on the row and column associations in an occupational transition matrix (i.e., the Altham statistic) (Altham and Ferrie, 2007).[54] Given two $r \times s$ transition matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$, with p_{ij} and q_{ij} as elements, the Altham statistic is:

$$d[\boldsymbol{P},\boldsymbol{Q}] = \left[\sum_{i=1}^{r}\sum_{j=1}^{s}\sum_{l=1}^{r}\sum_{m=1}^{s}\left|\log\left(\frac{p_{ij}p_{lm}q_{im}q_{lj}}{p_{im}p_{lj}q_{ij}q_{lm}}\right)\right|^{2}\right]^{1/2}$$

Often researchers report the Altham statistic of a matrix $\boldsymbol{P}$ from an independent matrix $\boldsymbol{J}$ where each element is one, indicating perfect mobility. The standard method places fathers and sons into four occupational categories (white collar, farmer, semi-skilled and unskilled).

In contrast to the IGE, the Altham statistic is barely influenced when going from a white-only sample to a black and white sample (see Table D1). While the preferred IGE increases by about 44 percent when adding black families, the Altham statistic increases by only 5 percent from 15.5 to 16.3. (Note that these estimates do not account for measurement error with multiple father observations.) The lack of movement when adding black families is likely because the Altham statistic does not capture black-white disparities within occupation category. For example, recall

[54] For recent examples of studies that use the Altham statistic, see Cilliers and Fourie (2018), Ferrie (2005), Long and Ferrie (2013a), Long and Ferrie (2018), Modalsli (2017), Pérez (2017), and Pérez (2019). I use the code from Long and Ferrie (2013a) to calculate the Altham statistics.

that the IGE based on the 1950 occupational income score, which assigns the same income for everyone in an occupation, increased by only 8 percent when adding black families. Based on this result, it is unsurprising that the Altham statistic does not change much when adding black families. This result reiterates the point that to fully account for racial disparities, one must both include black families and use a status measure that captures the historical black-white income gap.

Accounting for measurement error in the Altham statistic is not straightforward. The problem arises because averaging the father's occupational category across censuses does not place him in one discrete category. For example, if a father is a farmer in one census and a white-collar worker in another census, which category is his "true" one? I take two approaches to address this problem. First, since I have three father observations, I can place the father in his most-observed category and then measure the Altham statistic.[55] However, this method is not helpful for those who only have access to two father observations. Therefore, I take a second approach where I restrict the sample to sons with fathers who are "truly" in a specific occupation group; that is, if he is observed in the same group in all three observations. Others could use this approach if they only had access to two father observations. At the same time, this method drops part of the sample where fathers switch categories.

Based on the first approach where fathers are placed in their most-observed occupation group, the Altham statistic barely increases (from 16.3 to 17.5). With this method, the three father-observation Altham statistic with black and white families (17.5) is 13 percent higher than the Altham statistic for white families with only one father observation (15.5). To put this increase in context, Long and Ferrie (2013a, Table 2) estimate that the Altham statistic increased by 42 percent from 14.6 for 1880-1900 data to 20.8 for 1950-1973 data. Therefore, this approach suggests that the Altham statistic is not strongly influenced by measurement error.

The second approach to address measurement error, where I restrict the sample to fathers observed in the same group three times, does strongly influence the Altham statistic. Based on this method, the Altham statistic increases by 47 percent from the baseline estimate of 15.5 to updated estimate of 22.8. Therefore, there is substantially less intergenerational occupational mobility for the subsample where the father's "true" occupation group can be more precisely pinpointed. At

[55] Even with three censuses, 7.6 percent of fathers are observed in three different categories. I place these fathers in the observed category from 1910.

the same time, this method suffers from dropping 54 percent of the sample, making it unclear whether it is preferable to the first method.

Rather than using the Altham statistic, a simple regression of the son's occupation group on the father's more clearly shows that measurement error matters for occupation categories. For this regression, the father's outcome is an average of zero-one variable across three censuses. In this model, going from a one to the average of three-father observations increases the white-collar coefficient by 50 percent, from 0.34 to 0.50 (see Table D2). Instrumenting the 1910 father observation with the 1920 one further increases the estimate to 0.66. However, note that this 0.66 estimate may overstate the relationship between father and son due to non-classical measurement error in categorical variables (Bingley and Martinello, 2017; Dupraz and Ferrara, 2018). At the same time, there is non-classical error in the son's occupational category, which may attenuate the IV estimate, leaving the overall persistence of occupational category from father to son unclear. Nevertheless, it appears that measurement error matters – even for broad occupation groups. However, it may not show up in the Altham statistic depending on how one assigns a father to his "true" category.[56]

[56] Another way one could account for measurement error is to average the three transition matrices from 1900-1940, 1910-1940, and 1920-1940, and then calculate the Altham statistic for this averaged matrix. However, this method does not narrow down on the father's true occupation category since a given observation in one matrix is not linked to the other matrix.

Table D1. Altham statistics are not influenced by racial composition but can be by measurement error.

	Independence	White, occupation in 1910	Black, occupation in 1910	Pooled, occupation in 1910	Pooled, most common occupation 1900-1920
Independence					
White, occupation in 1910	15.6				
Black, occupation in 1910	15.9	2.4			
Pooled, occupation in 1910	16.4	1.8	3.5		
Pooled, most common occupation 1900-1920	17.9	2.9	4.4	1.4	
Pooled, observed in same occupation 1900-1920	22.9	8.1	8.8	6.9	5.7

Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: Each cell reports the Altham statistic between the row and the column matrix. See Tables D3-D5 for the underlying occupation transition matrices. The "occupation in 1910" rows are Altham statistic as calculated in single-linked data. The "most common occupation 1900-1920" sets the father's occupation group in the one he is in at least two of the three censuses between 1900 and 1920). If the father is observed in three different categories in 1900, 1910 and 1920, I use the 1910 category. The "observed in same occupation 1900-1920" limits the sample to father who are observed three times as either a farmer, white-collar worker, unskilled worker or semi-skilled worker (n=179,663).

Table D2. Persistence of occupation categories is influenced by multiple father observations

	1 obs.	Mean of 3 obs.	2SLS	Most commonly observed	Father in same occ. across three censuses
Panel A. White-Collar					
Father's outcome	0.338	0.504	0.657	0.369	0.484
	(0.002)	(0.003)	(0.004)	(0.002)	(0.003)
Observations	394,864	394,864	394,864	394,864	183,685
Panel B. Farmer					
Father's outcome	0.239	0.323	0.369	0.249	0.319
	(0.001)	(0.002)	(0.002)	(0.001)	(0.002)
Observations	394,864	394,864	394,864	394,864	183,685
Panel C. Semi-skilled					
Father's outcome	0.112	0.228	0.376	0.139	0.214
	(0.003)	(0.004)	(0.010)	(0.003)	(0.007)
Observations	394,864	394,864	394,864	394,864	183,685
Panel D. Unskilled					
Father's outcome	0.153	0.236	0.337	0.168	0.228
	(0.002)	(0.003)	(0.004)	(0.002)	(0.003)
Observations	394,864	394,864	394,864	394,864	183,685

Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: The last two columns assign the father his occupation category in the same way as in the Altham statistic in Table D1. More specifically, the most commonly observed category set the father's outcome equal to one if he is observed in the occupation most often (i.e., in at least two of the three censuses between 1900 and 1920). If the father is observed in three different categories in 1900, 1910 and 1920, I use the 1910 category. The observed in same occupation column limits the sample to father who are observed three times as either a farmer, white-collar worker, unskilled worker or semi-skilled worker.

Table D3. Occupational Transition Matrix for White Families

	Father's Occupation				
Son's occ. in row	White Collar	Semi-Skilled	Unskilled	Farmer	Row sum
White, occupation in 1910					
White Collar	34,206	29,510	12,419	31,902	108,037
	(58.7)	(37.0)	(28.8)	(21.0)	
Semi-skilled	15,990	35,810	17,344	44,068	113,212
	(27.4)	(44.9)	(40.2)	(28.9)	
Unskilled	5,961	12,069	10,411	33,025	61,466
	(10.2)	(15.1)	(24.1)	(21.7)	
Farmer	2,132	2,316	2,963	43,249	50,660
	(3.7)	(2.9)	(6.9)	(28.4)	
Column sum	58,289	79,705	43,137	152,244	333,375
White, most commonly observed occupation					
White Collar	35,275	30,040	11,267	31,455	108,037
	(61.2)	(36.5)	(27.2)	(20.7)	
Semi-skilled	14,944	37,685	16,927	43,657	113,213
	(25.9)	(45.8)	(40.9)	(28.7)	
Unskilled	5,538	12,368	10,586	32,973	61,465
	(9.6)	(15.0)	(25.6)	(21.7)	
Farmer	1,864	2,126	2,626	44,044	50,660
	(3.2)	(2.6)	(6.3)	(29.0)	
Column sum	57,621	82,219	41,406	152,129	333,375
White, same occupation group in 1900, 1910 and 1920					
White Collar	17,942	13,034	2,384	15,987	49,347
	(70.7)	(37.4)	(25.6)	(18.9)	
Semi-skilled	5,074	16,600	3,836	21,585	47,095
	(20.0)	(47.6)	(41.2)	(25.5)	
Unskilled	1,839	4,721	2,720	17,818	27,098
	(7.2)	(13.5)	(29.2)	(21.0)	
Farmer	535	520	367	29,338	30,760
	(2.1)	(1.5)	(3.9)	(34.6)	
Column sum	25,390	34,875	9,307	84,728	154,300

Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: The top panel is when using the 1910-1940 linked sample. The middle panel uses the triple linked sample, where the father's occupation is the one in which he is observed in most often. That is, if the father is observed in three different categories in 1900, 1910 and 1920, I assign the father to his category in the 1910 census. The bottom panel limits the sample to those with fathers who are observed in the same occupation group in the 1900, 1910 and 1920 censuses. The weighted number of observations are reported in the cells, rounded to the nearest whole number. Since weighted observations are used, the number does not match the raw number of observations in the paper.

Table D4. Occupational Transition Matrix for Black Families

Son's occ. in row	White Collar	Semi-Skilled	Unskilled	Farmer	Row sum
	Father's Occupation				
Black, occupation in 1910					
White Collar	170	246	764	869	2,049
	(20.8)	(9.4)	(7.5)	(3.8)	
Semi-skilled	225	726	2,202	3,550	6,703
	(27.5)	(27.8)	(21.5)	(15.4)	
Unskilled	376	1,555	6,604	12,421	20,956
	(46.0)	(59.4)	(64.6)	(54.0)	
Farmer	47	89	651	6,158	6,945
	(5.7)	(3.4)	(6.4)	(26.8)	
Column sum	818	2,616	10,221	22,998	36,653
Black, most commonly observed occupation					
White Collar	143	254	839	813	2,049
	(24.0)	(11.7)	(7.3)	(3.6)	
Semi-skilled	171	645	2,427	3,460	6,703
	(28.7)	(29.7)	(21.2)	(15.4)	
Unskilled	251	1,197	7,506	12,001	20,955
	(42.1)	(55.2)	(65.7)	(53.4)	
Farmer	31	73	657	6,184	6,945
	(5.2)	(3.4)	(5.7)	(27.5)	
Column sum	596	2,169	11,429	22,458	36,652
Black, same occupation group in 1900, 1910 and 1920					
White Collar	63	59	372	401	895
	(42.6)	(11.0)	(8.4)	(3.4)	
Semi-skilled	44	195	990	1,716	2,945
	(29.7)	(36.2)	(22.5)	(14.4)	
Unskilled	41	267	2,927	6,117	9,352
	(27.7)	(49.6)	(66.4)	(51.3)	
Farmer	0	17	119	3,684	3,820
	0.0	(3.2)	(2.7)	(30.9)	
Column sum	148	538	4,408	11,918	17,012

Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: The top panel is when using the 1910-1940 linked sample. The middle panel uses the triple linked sample, where the father's occupation is the one in which he is observed in most often. That is, if the father is observed in three different categories in 1900, 1910 and 1920, I assign the father to his category in the 1910 census. The bottom panel limits the sample to those with fathers who are observed in the same occupation group in the 1900, 1910 and 1920 censuses. The weighted number of observations are reported in the cells, rounded to the nearest whole number. Since weighted observations are used, the number does not match the raw number of observations in the paper.

Table D5. Occupational Transition Matrix for Black and White Families

	Father's Occupation				
Son's occ. in row	White Collar	Semi-Skilled	Unskilled	Farmer	Row sum
Pooled, occupation in 1910					
White Collar	33,448	29,329	13,105	33,355	109,237
	(58.0)	(36.1)	(24.6)	(18.8)	
Semi-skilled	15,890	36,158	19,481	48,341	119,870
	(27.5)	(44.5)	(36.6)	(27.2)	
Unskilled	6,209	13,448	16,972	45,984	82,613
	(10.8)	(16.5)	(31.9)	(25.9)	
Farmer	2,166	2,394	3,622	50,125	58,307
	(3.8)	(2.9)	(6.8)	(28.2)	
Column sum	57,713	81,329	53,180	177,805	370,027
Pooled, most commonly observed occupation					
White Collar	34,298	29,918	12,102	32,919	109,237
	(60.7)	(35.9)	(22.9)	(18.6)	
Semi-skilled	14,701	37,945	19,320	47,905	119,871
	(26.0)	(45.5)	(36.6)	(27.0)	
Unskilled	5,613	13,362	18,096	45,543	82,614
	(9.9)	(16.0)	(34.3)	(25.7)	
Farmer	1,864	2,179	3,279	50,984	58,306
	(3.3)	(2.6)	(6.2)	(28.7)	
Column sum	56,476	83,404	52,797	177,351	370,028
Pooled, same occupation group in 1900, 1910 and 1920					
White Collar	17,432	13,095	2,798	16,831	50,156
	(70.5)	(37.0)	(20.3)	(17.1)	
Semi-skilled	4,958	16,782	4,846	23,839	50,425
	(20.1)	(47.5)	(35.2)	(24.2)	
Unskilled	1,806	4,929	5,646	24,384	36,765
	(7.3)	(13.9)	(41.0)	(24.7)	
Farmer	521	541	482	33,643	35,187
	(2.1)	(1.5)	(3.5)	(34.1)	
Column sum	24,717	35,347	13,772	98,697	172,533

Source: Data are from the 1900, 1910, 1920 and 1940 US censuses (Ruggles et al., 2019).
Notes: The top panel is when using the 1910-1940 linked sample. The middle panel uses the triple linked sample, where the father's occupation is the one in which he is observed in most often. That is, if the father is observed in three different categories in 1900, 1910 and 1920, I assign the father to his category in the 1910 census. The bottom panel limits the sample to those with fathers who are observed in the same occupation group in the 1900, 1910 and 1920 censuses. The weighted number of observations are reported in the cells, rounded to the nearest whole number. Since weighted observations are used, the number does not match the raw number of observations in the paper.