Like Great-Grandparent, Like Great-Grandchild? Multigenerational Mobility in American History

Zachary Ward* Kasey Buckles† Joseph Price‡

Baylor University University of Notre Dame BYU

June 2025

Abstract: Using data on 2.5 million great-grandchildren linked to their great-grandfathers in the US (1850–1940), we show that economic gaps persisted strongly across four generations despite major structural change. We find that one-third of the initial differences in economic status across white great-grandfathers remained in their great-grandchildren. When including both Black and white families, this persistence rises to about 50 percent, largely because the gap between Black and white families closed slowly over time. Grandparent and great-grandparent status matter beyond the father's status, indicating slower convergence to the mean than predicted by two-generational estimates. However, this excess persistence is largely driven by enduring racial inequality as ancestor effects are small within the white population.

Acknowledgments: We thank audiences at Auburn, Baylor, Duke, Florida State, Kansas State, St Andrews, Toronto, Vanderbilt, and the Washington-Area Economic History Seminar. Parts of this work were also presented at NYU and the University of Southern Denmark. We received helpful feedback at the 2025 MWEA, 2020 SEA, 2021 SEHO, 2021 SOLE, and 2023 SSHA meetings. We are especially grateful to Greg Clark, James Feigenbaum, Walker Hanlon, Jonas Helgertz, Bhash Mazumder, Santiago Perez, John Parman, Jan Stuhler, Kelly Vosters, and Ariell Zimran for their comments, some of whom provided feedback at very early stages of this project. Thanks to Josh Reynolds for help with compiling death dates for our sample. All errors are our own.

*Department of Economics, Hankamer School of Business, Baylor University, One Bear Place 98001, Waco, TX 76798, USA.

†Department of Economics, University of Notre Dame, South Bend, IN 46556, USA.

‡Department of Economics, Brigham Young University. 2129 WVB Provo, UT 84602

"Equality of opportunity" is often proxied by the extent to which parental income predicts that of their children. A strong prediction suggests that circumstances outside a child's control, like family background, play a major role in shaping lifetime outcomes. Researchers have studied how this link varies across countries (Corak, 2013), within countries (Chetty et al., 2014), and over time (Davis and Mazumder, 2024; Jácome, Kuziemko and Naidu, 2025; Ward, 2023), helping to identify when and where opportunity is more limited.

But focusing only on parents may miss deeper patterns across multiple generations. This is because having rich grandparents or great-grandparents, conditional on parental income, may also predict a child's economic status (Braun and Stuhler, 2018; Clark, 2014; Mare, 2011; Solon, 2018). Historical studies, in particular, consistently find evidence of these "grandparent effects" (Belloc et al., 2024; Dribe and Helgertz, 2016; Knigge, 2016; Long and Ferrie, 2018; Hällsten and Kolk, 2023; Modalsli, 2023; Olivetti, Paserman and Salisbury, 2018; Ward, 2020).[1] Some studies estimate remarkably large effects, with grandparent-to-grandchild status transmission more than double what would be expected from parent-child estimates alone.[2] This suggests that even the extensive literature on intergenerational mobility, including studies finding low relative mobility for cohorts born before World War II (Buckles et al., 2024*b*; Jácome, Kuziemko and Naidu, 2025; Ward, 2023), may understate just how limited opportunity was across generations.

In this paper, we present new estimates of multigenerational mobility between 1850 and 1940 using a novel four-generation dataset linking over 2.5 million adults to their parents, grandparents, and great-grandparents. Four-generation samples are rare, but by linking full-count US censuses (1850–1940) (Ruggles et al., 2024) at high match rates using recent advances in historical linking (Buckles et al., 2024*a*), we create, by far, the largest four-generational sample in the literature. Importantly, our data include both Black and white families, capturing mobility for the broader population rather than focusing exclusively on whites. Additionally, the dataset includes multiple observations of fathers, grandfathers, and great-grandfathers, allowing us to address measurement error, which is particularly important in mobility studies (Modalsli and Vosters, 2024; Solon, 1992; Mazumder, 2005).

[1] Historical studies strongly support grandparent effects, while modern studies yield mixed results (Anderson, Sheppard and Monden, 2018; Solon, 2018). Colagrossi, d'Hombres and Schnepf (2020) find consistent evidence of grandparent effects using data from 28 EU countries. An early example, Hodge (1966), finds modest grandparent effects based on a 1949 survey.

[2] Long and Ferrie (2018), using U.S. census data (1850–1910), report an occupational elasticity between grandfather and grandchild of 0.145—164% higher than implied by multiplying the father-son coefficient twice (0.234^2=0.055, see Table 5 in Long and Ferrie (2018)).

Using this new dataset, we present four main results. First, economic status is remarkably persistent across four generations. Our preferred estimate is that 33% of the differences in status across white great-grandparents remained in their great-grandchildren. This level of persistence directly contradicts the claim that advantages or disadvantages vanish within three generations (Becker and Tomes, 1986). It is all the more striking given the major changes in the American economy from the great-grandparents' era (1850–1860) to that of their great-grandchildren (1930–1940).

Second, four-generation transmission for white families was slightly stronger than predicted from two-generation estimates. When we predict a child's status using their parents, grandparents, and great-grandparents, we find a small but meaningful grandparent "effect" and zero "effect" from great-grandparents. (While we use the term *effect* to be consistent with the literature, all estimates are non-causal.) To address transitory measurement error, we use an instrumental variables approach, instrumenting each ancestor's status with a second observation (Modalsli and Vosters, 2024; Ward, 2023). To capture the total predictive power of ancestors, we sum the parent, grandparent, and great-grandparent coefficients, for what we call the *total ancestor effect*. For white families, the total ancestor effect is 9% larger than the father effect.

Our third result is that transitory measurement error, caused by short-term shocks or recording/transcription error, explains why previous research found larger grandparent effects. If we ignore these errors by relying on a single snapshot of the father, grandfather, and great-grandfather rather than using IV, we reproduce the literature's findings of very large ancestor effects. The total ancestor effect is 54% larger than the father effect, suggesting a large amount of excess persistence across multiple generations. However, once we correct for transitory measurement error using repeated observations and IV, the grandparent and great-grandparent effects shrink substantially. These results are consistent with measurement error driving large grandparent effects (Engzell, Mood and Jonsson, 2020; Ferrie, Massey and Rothbaum, 2021; Modalsli and Vosters, 2024). Intuitively, with a noisy measure of the father, distant ancestors act as proxies for the father's true status. If we improve the measurement of the father's status with repeated observations, the proxies from the grandfather and great-grandparent provide less predictive power.

We show that the small grandparent effect that remains for the white population results largely from measurement error that repeated observations do not fully correct, rather than from a direct causal effect. This error occurs because even a complete income or occupational history ("permanent income") does not fully capture a person's productive capability (Braun and Stuhler, 2018; Clark, 2014). Instead of *transitory* error, we call this *permanent* error. To gauge its importance, we assume a latent factor model and use an approach suggested by Braun and Stuhler (2018) and Lindahl et al. (2015), where we instrument the father's status with the great-grandfather's status. This

method suggests that the transmission of the latent factor is only slightly stronger than the repeated measurement IV estimate, which corrects for transitory error. Based on the assumptions of the latent factor model, permanent error is small compared to transitory error (about one-eighth the size).

Our fourth main finding is that including Black families in the sample raises both the total persistence of status across four generations and the amount of excess persistence that goes beyond what two-generation models predict. While 33% of status differences across *white* great-grandfathers remained in their great-grandchildren, 51% remained when analyzing differences across *Black and white* great-grandfathers. Using an AR(3) model, the total ancestor effect is 20% larger than the father's effect alone in the Black and white sample. In contrast, the total ancestor effect was only 9% larger in the white-only sample. The higher persistence in the full population is due to strong between-group persistence (i.e., a barely converging Black-white gap), which causes convergence to the mean across multiple generations to be slower for the overall population (Solon, 2018). As a result, two-generational models underestimate long-run persistence for the Black and white population, but they come closer to the mark for white families alone.

A key limitation of our approach is that historical data lack direct income measures, so we rely mostly on occupation-based status. This makes our estimates less comparable to modern income mobility studies (e.g., Chetty et al. (2014); Corak (2013)).[3] However, in smaller samples where wealth proxies and wage data are available, we still find strong four-generation persistence: a rank-rank slope of 0.23-0.35 between property wealth in the first generation and home values in the fourth, and an rank-rank slope of 0.20 between property wealth in the first generation and wage income in the fourth. Unfortunately, these wealth and income measures only exist for either the great-grandparents or great-grandchildren, not the intermediate generations. We therefore cannot directly test for excess persistence in these measures.[4]

This paper contributes to a growing literature on multigenerational mobility (see reviews in Solon (2018) and Stuhler (2024)). As in previous work by others, we find that relative mobility is lower when extended kin are included, though the amount of excess persistence we detect is small compared to other studies (Adermon, Lindahl and Palme, 2021; Belloc et al., 2024; Celhay and Gallegos, 2025; Collado, Ortuño-Ortín and Stuhler, 2023; Clark, 2023; Ferrie, Massey and Rothbaum, 2021). One key contribution is separating transitory error, central in intergenerational studies, from permanent error measuring latent advantages, central in multigenerational studies. This

[3]Our primary measure adjusts occupational status for differences by race and region (Ward, 2023). Without these adjustments, four-generation persistence for white families remains high at 34% (Song et al., 2020).

[4]Given the occupational-based measure, it is possible that we underestimate excess persistence due to generations being especially productive within occupation across generations (Kalsi and Ward, 2025).

is only possible thanks to our rich dataset that has both four generations and repeated observations; many other studies rely on recalls of grandparents' status, and thus do not have repeated observations. Our findings show that transitory error plays a major role. This result may be amplified by the messy nature of historical records in our data, but the result is consistent with evidence from modern-day Norway (Modalsli and Vosters, 2024).

This paper extends recent research on pre-World War II intergenerational mobility and assortative mating by Buckles et al. (2024*b*) and Ward (2023) to a multigenerational context. Their studies highlighted the roles of measurement error, racial disparities, and strong assortative mating in two-generational data. Using multigenerational data, we confirm these findings from a different perspective. We show that relying on single snapshots of fathers' status leads to implausibly high estimates of grandparent and great-grandparent effects, reflecting significant measurement error in historical data.[5] High persistence across multiple generations, coupled with evidence that mothers mattered for child outcomes (Althoff, Brookes Gray and Reichardt, 2025; Espín-Sánchez, Ferrie and Vickers, 2025), implies that assortative mating was high (Braun and Stuhler, 2018; Stuhler, 2024), corroborating earlier findings from two-generational data (Buckles et al., 2024*b*). Further evidence for strong assortative mating comes from our finding that a great-grandfather's status predicts their great-grandchild's *spouse's* education (slope = 0.23) nearly as well as the *own* great-grandchild's education (slope = 0.25), indicating that people tended to marry others from similar family backgrounds. Still, prior work has shown high turnover in the top tail of the economic distribution, suggesting that elite status was not locked in even during this period of slow average mean reversion (Kalsi and Ward, 2025). Overall, the evidence indicates that pre–World War II America was far from a land of equal opportunity for the overall population.

Our paper makes a unique contribution to the multigenerational and extended kin literature by highlighting how group effects, particularly by race, significantly impact multigenerational mobility estimates (Solon, 2018).[6] One key result is that including Black families in our analysis increases both the estimated strength of status transmission across multiple generations and the degree to which this multigenerational persistence exceeds what would be predicted from two-generation estimates. This finding shows how sample composition can substantially affect mobility estimates, linking our work to broader research on racial inequality and mobility (Abramitzky et al., 2021; Chetty et al., 2020; Collins and Wanamaker, 2022; Hertz, 2008; Jácome, Kuziemko and Naidu, 2025; Pfeffer and Killewald, 2018; Ward, 2023). This result is especially relevant today, as economic gaps

[5] We also use alternative estimation methods, including rare-surname estimators (Clark, 2014) and an IV approach where we instrument the father with the great-grandfather's status (Braun and Stuhler, 2018; Lindahl et al., 2015), and these also indicate low relative mobility, with persistence (AR(1)) for whites estimated around 0.63–0.70.

[6] The transmission of group effects relates to the general idea that the transmission of multiple traits, each with a different rate of persistence, can generate excess persistence across multiple generations (Stuhler, 2012, 2024).

between races have converged much more slowly than gaps across families *within* racial groups over the past 80 years (Bayer and Charles, 2018; Chetty et al., 2020; Derenoncourt et al., 2024). This slower between-group convergence suggests that modern national-level two-generational estimates are likely to miss the mark on the convergence of economic gaps across multiple generations in the United States.

1. Theory and Estimation

1.1 The statistical approach

Intergenerational relative mobility is typically measured by regressing a child's economic status (G2)—such as income, education, wealth, or occupation—on their parent's status (G1). Using demeaned variables (to omit the constant), the standard regression is:

$$y_{i,g} = \beta_{-1} y_{i,g-1} + \varepsilon_{i,g} \tag{1}$$

The coefficient β_{-1} captures mobility between those one generation apart: values close to zero indicate high mobility (parental status matters little), while values near one indicate low mobility. With multigenerational data, researchers can run similar regressions for later generations—for example, from G2 to G3 (grandchild), or G3 to G4 (great-grandchild).[7]

Alternatively, one can directly estimate transmission across multiple generations at once, such as the direct regression between the first (G1) and fourth (G4) generation:

$$y_{i,G4} = \beta_{-3} y_{i,G1} + \varepsilon_{i,G4} \tag{2}$$

Researchers then test whether the direct four-generation effect (β_{-3}) exceeds the product of adjacent-generation coefficients (β_1^3), indicating "excess persistence" ($\beta_{-3} > \beta_{-1}^3$). For instance, if each two-generation coefficient equals 0.5, the predicted four-generation effect is $0.5^3 = 0.125$. If the actual four-generation coefficient is 0.20, then transmission across four generations is $\frac{0.20-0.125}{0.125} = 60\%$ stronger than implied by two-generation estimates.

[7] These slopes may differ if social mobility varies over time, but we assume stable slopes here for a more straightforward discussion. In the empirical section, we will allow the slopes to vary across generations.

Instead of comparing coefficients from separate regressions, one can estimate an AR(3) regression including the parent's, grandparent's, and great-grandparent's status simultaneously:[8]

$$y_{i,G4} = \beta_p y_{i,G3} + \beta_{gp} y_{i,G2} + \beta_{ggp} y_{i,G1} + u_{i,G4} \tag{3}$$

Researchers can then directly test whether earlier generations "matter" by testing if $(\beta_{gp} > 0)$ and $(\beta_{ggp} > 0)$. The overall predictive power of ancestors is captured by the sum of coefficients: $\widehat{\beta}_p + \widehat{\beta}_{gp} + \widehat{\beta}_{ggp}$, which we call the *total ancestor effect*.

Positive coefficients for β_{gp} or β_{ggp} could reflect a direct, causal effect of grandparents or great-grandparents on their descendants. However, there would be at least two other alternative explanations. First, any omitted variable that is both correlated with both the ancestor's and the offspring's outcome (like race or place-based effects) could create a positive coefficient. Second, ancestor effects could reflect measurement error.

1.2 Measurement error in a latent factor model

Measurement error can create a spurious grandparent or great-grandparent effect, even when these ancestors have no direct causal influence on their descendants. The literature highlights two main types of measurement error. First, short-term snapshots of a parent's income or occupation often misrepresent their permanent income due to data error or transitory fluctuations—an issue widely discussed in the intergenerational mobility literature (Solon, 1992; Mazumder, 2005; Ward, 2023). Second, as shown in multigenerational studies, even precise measures of permanent income fail to capture latent advantages that are unobservable in the data but still benefit children and grandchildren (Stuhler, 2012; Braun and Stuhler, 2018; Clark, 2014; Clark and Cummins, 2015).

Permanent income can differ from latent ability because even lifetime outcomes are shaped by random chance. Random events over a person's life, like a recession, a health shock, or landing a great job, can change their income path. A person may also be born with limited opportunities to obtain jobs at the top of the economic ladder because they belong to a discriminated group or live in a place with few chances to advance. Another reason permanent income can differ from the latent factor is preferences. Some people (including, say, academics) choose lower-paying work in

[8]The AR(3) regression includes all three ancestor generations simultaneously, so the coefficients for great-grandparent and grandparent status (β_{ggp} and β_{gp}) represent effects conditional on the father's status. These differ from the coefficients in separate bivariate regressions (e.g., β_{-3}).

nonprofits or public service because it aligns with their values.[9] Others might take lower-paying jobs to live near family.

To incorporate these measurement errors into a simple model in the spirit of Braun and Stuhler (2018) and Clark (2014), suppose you have a single observation of individual i's income from generation g and time period t, denoted y_{igt}. Let this observation equal their permanent income y^*_{ig} plus a transitory error term ε_{igt}, such that $y_{igt} = y^*_{ig} + \varepsilon_{igt}$ (Solon, 1992).[10] Additionally, permanent income y^*_{ig} is an imperfect measure of a latent factor x_{ig}, which differs due to a "permanent" error term u_{ig}. Let these errors be independent and uncorrelated with the true latent ability:

$$y_{igt} = x_{ig} + u_{ig} + \varepsilon_{igt} \tag{4}$$

$$x_{ig} = \lambda x_{i,g-1} + v_{ig} \tag{5}$$

The latent factor is transmitted across a *single* generation by λ, with added random noise v_{ig}. We emphasize that this latent "transmission" is still non-causal, like most estimates in the intergenerational mobility literature, as it captures mechanisms due to a direct effect from the father, but mechanisms from the environment (like place-based effects discussed before) that are correlated with the father's latent factor. But because it is across one generation, this model assumes that no latent transmission occurs across multiple generations. We will later provide supporting evidence for this claim by splitting the sample by grandparents who died before or after their grandchild's birth. But current evidence suggests that this is a reasonable assumption–Ager, Boustan and Eriksson (2021) and Bleakley and Ferrie (2016) show that random shocks to grandparental wealth had little to no impact on grandchildren's outcomes in American history.

Based on this model, the regression slope between a child's income and their parent's income measured j years earlier is:

$$\frac{\text{Cov}(y_{igt}, y_{i,g-1,t-j})}{\text{Var}(y_{i,g-1,t-j})} = \theta_{g-1}\lambda \tag{6}$$

where the reliability ratio θ_{g-1} that attenuates the slope is:

$$\theta_{g-1} = \frac{\sigma^2_{x,g-1}}{\sigma^2_{x,g-1} + \sigma^2_{u,g-1} + \sigma^2_{\varepsilon,g-1}} \tag{7}$$

[9] See Clark (2014, Chapter 6) for a detailed discussion of luck and preferences.

[10] Appendix A of Braun and Stuhler (2018) demonstrates a similar approach for including transitory error, framing it in terms of "recall bias" since respondents report parental or grandparental education.

Here, λ represents the true transmission of the latent factor across generations. If a snapshot of income perfectly measured the latent factor (no permanent error, $\sigma_u^2 = 0$, and no transitory error, $\sigma_\varepsilon^2 = 0$), then the regression slope would equal the true transmission, λ. With measurement error ($\sigma_u^2 > 0$ or $\sigma_\varepsilon^2 > 0$), the slope is attenuated. If instead one regresses a great-grandchild's income on the great-grandparent's income, the slope becomes $\theta_{g-3}\lambda^3$ reflecting three generations of transmission of the latent factor, but with attenuation due to error in the great-grandparent's generation ($g-3$).

Substituting these slopes into the comparison between the direct three-generational slope ($\widehat{\beta}_{-3}$) and the product of the two-generational slopes ($\widehat{\beta}^3_{-1}$) shows that there is evidence of excess persistence when:

$$\begin{aligned} \widehat{\beta}_{-3} &> \widehat{\beta}^3_{-1} \\ \theta_{G1}\lambda^3 &> \theta_{G1}\theta_{G2}\theta_{G3}\lambda^3 \\ 1 &> \theta_{G2}\theta_{G3} \end{aligned} \tag{8}$$

This result shows that excess persistence or spurious ancestor effects are due to measurement errors in the intermediate generations—specifically, the parent's or grandparent's ($\theta_{G2} < 1$ or $\theta_{G3} < 1$). Measurement error in the earliest generation (great-grandparent θ_{G1}) does not bias toward evidence of excess persistence. Instead, it attenuates estimates of the true rate of persistence across four generations.

1.3 Repeated-measure instrumental variables estimates

One approach to addressing measurement error is a repeated-measurement instrumental variables strategy, using a second observation of a parent's income ($y_{i,g-1,s}$) to instrument for the first ($y_{i,g-1,t}$). Since our data come from censuses conducted a decade apart, it is unlikely that transitory components persist over such a long period (Modalsli and Vosters, 2024). However, this approach corrects only for transitory errors, not for permanent deviations from underlying status.

To see why, recall that the IV estimate is the ratio of the reduced-form slope and the first-stage slope. The reduced form is the regression of the offspring's outcome on the second observation of the ancestor, which yields the same attenuated transmission as the regression on the first observation of the ancestor $\theta_{g-1}\lambda$. The first stage is the regression of the second ancestor's observation on the first, which covary as the latent factor and permanent error are the same across observations, but

transitory error is not. Putting this together yields an IV slope of:[11]

$$\beta_{-1}^{IV} = \frac{\text{Cov}(y_{ig}, y_{i,g-1,s})}{\text{Cov}(y_{i,g-1,t}, y_{i,g-1,s})} = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2}\lambda \tag{9}$$

This equation shows that the IV estimate will not capture the transmission of the latent factor if lifetime income imperfectly measures the latent factor. Note that we plug the IV estimates into the excess persistence formula (as in Equation 8), we will find excess persistence even if none exists, if variation in permanent error σ_u^2 is large.

1.4 Group effects

So far, we have assumed a homogeneous population, but the United States has important inequalities across various groups. This matters because if group effects are substantial and transmit differently across generations than within-group transmission, then a grandparent effect will emerge. For example, persistent racial inequalities can create a correlation between Black great-grandparents and their great-grandchildren—not through direct interaction but due to limited opportunities affecting both generations. Empirically, this takes the form of omitted variable bias from a fixed effect (Solon, 2014, 2018).

To understand this issue within the latent variable framework, suppose that permanent deviations in income u_{ig} from the latent factor x_{ig} are no longer random but are a group fixed effect plus noise. This fixed effect shifts the intercept in earnings up or down $(u_{ig} = \eta_{g,group} + \omega_{ig})$ Assume this fixed effect diminishes over generations at a rate $\gamma < 1$, meaning $\eta_{group,g} = \gamma\eta_{group,g-1}$, possibly due to declining discrimination over time. Under these assumptions, regressing a child's outcome on the outcome of an ancestor m generations ago is a weighted average of latent factor transmission and fixed-effect transmission.

$$\frac{\text{Cov}(y_{ig}, y_{i,g-m})}{\text{Var}(y_{i,g-m})} = \lambda^m \left(\frac{\sigma_x^2}{\sigma_y^2}\right) + \gamma^m \left(\frac{\sigma_{\eta_{\text{group}}}^2}{\sigma_y^2}\right) \tag{10}$$

Here, the weights reflect the share of total income variation in generation $g - m$ due to latent factor, $\frac{\sigma_x^2}{\sigma_y^2}$, and due to between-group differences, $\frac{\sigma_{\eta_{\text{group}}}^2}{\sigma_y^2}$. Total variation σ_y^2 is now the sum of

[11] $\beta_{-1}^{IV} = \frac{\frac{\text{Cov}(y_{ig}, y_{i,g-1,s})}{\text{Var}(y_{i,g-1,s})}}{\frac{\text{Cov}(y_{i,g-1,t}, y_{i,g-1,s})}{\text{Var}(y_{i,g-1,s})}} = \frac{\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2+\sigma_e^2}\lambda}{\frac{\sigma_x^2+\sigma_u^2}{\sigma_x^2+\sigma_u^2+\sigma_e^2}} = \frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2}\lambda.$

variation in the latent factor, the group fixed effect, and the idiosyncratic errors from both the permanent and transitory components.

In the absence of measurement error, this is the same within-group/between-group decomposition of the slope discussed in intergenerational mobility research (Bailey, Cole and Massey, 2020; Hertz, 2008; Jácome, Kuziemko and Naidu, 2025; Ward, 2023).[12] However, because measurement error inflates the total variance σ_y^2, the OLS estimate with group effects is still attenuated relative to the "true" transmission with group effects. Using IV instead of OLS eliminates bias from transitory error, but permanent error still causes the estimates to be attenuated.[13]

But if we ignore measurement error to illustrate the importance of group effects alone, and plug the slopes into the test for excess persistence,[14] then the formula reduces to

$$\text{plim}\left[\widehat{\beta}_{-3} - \widehat{\beta}^3_{-1}\right] = \left[\lambda^3 \frac{\sigma_x^2}{\sigma_y^2} + \gamma^3 \frac{\sigma^2_{\eta_{\text{group}}}}{\sigma_y^2}\right] - \left[\lambda \frac{\sigma_x^2}{\sigma_y^2} + \gamma \frac{\sigma^2_{\eta_{\text{group}}}}{\sigma_y^2}\right]^3. \tag{11}$$

If group effects persisted at the same rate of transmission within group ($\lambda = \gamma$) and there is no measurement error ($\frac{\sigma_x^2}{\sigma_y^2} + \frac{\sigma^2_{\eta_{\text{group}}}}{\sigma_y^2} = 1$), then there is no excess persistence, as expected. But if ($\lambda \neq \gamma < 1$), then by Jensen's inequality there will be positive excess persistence.[15] Therefore, if any group effect persists, as suggested by the extensive literature on intergenerational mobility gaps by race, ethnicity, and region, this would lead to excess persistence, even in the absence of measurement error.

What if the group effect is not AR(1)? In American history, several group inequalities, like those by race or place, have not followed a smooth path toward equality (Bayer and Charles, 2018; Derenoncourt et al., 2024; Gray et al., 2024; Margo, 2016; Wanamaker, 2017). After the Civil War,

[12] This assumes that within-group transmission is identical across groups.

[13] In this setting, the two repeated measures share the permanent components $x_{i,g-1}$, $\eta_{\text{group},g-1}$, and the idiosyncratic component $\omega_{i,g-1}$, while transitory errors are uncorrelated. The IV estimator is then given by

$$\frac{\text{Cov}(y_{ig}, y_{i,g-1,s})}{\text{Cov}(y_{i,g-1,t}, y_{i,g-1,s})} = \lambda \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\eta^2 + \sigma_\omega^2} + \gamma \frac{\sigma_\eta^2}{\sigma_x^2 + \sigma_\eta^2 + \sigma_\omega^2},$$

which is larger than the OLS estimate because the transitory error is once again removed.

[14] Also assume that the share of variance within- and between-group remains stable across generations. If the between-group share gets smaller over time, as would be expected, then iterated term that is cubed would be weighted more toward the within-group transmission λ, which would also lead to evidence of excess persistence if $\lambda < \gamma$.

[15] If $\lambda \neq \gamma < 1$, then by the convexity of the cubic and Jensen's inequality, $\lambda^3 + \gamma^3 > (\lambda + \gamma)^3$, leading to positive excess persistence.

initial Black progress stalled with the end of Reconstruction (1877) and resumed only during the First Great Migration (1916-1940). Regional gaps also have an uneven pattern: North-South income and wealth gaps *widened* after the destruction of the Civil War, then stagnated, and began to narrow only after the 1930s (Lindert and Williamson, 2016; Wright, 1986). If group effects do not follow an AR(1) process but instead are AR(2) or AR(3), then grandparent and great-grandparent effects could be *positive or negative* even if the within-group transmission is AR(1). It depends on the context, as there is not a theoretical reason why all group effects behave in a specific way. In the end, the rate of multigenerational mobility must be estimated from the data.

2. Data

2.1 Building a multigenerational sample

We use complete-count U.S. censuses from 1850 to 1940, provided by IPUMS (Ruggles et al., 2024). Individuals are linked across censuses using the Census Tree (Price et al., 2023), described by Buckles et al. (2024*a*). The Census Tree has two key advantages over other linking methods. First, it uses genealogical records from FamilySearch.org to link women from childhood to adulthood, allowing us to trace both maternal and paternal lines. Second, it achieves a higher linking rate than automated methods like the Census Linking Project. This matters because our analysis requires multiple links to observe four generations with repeated measures.

We explain the data creation process in more detail in Appendix A, but the process is shown for an example family in Figure 1. To create the data, we start with adults from the 1930 or 1940 U.S. censuses, whom we call the fourth generation (G4). We trace each G4 individual back to their childhood home to find their parents (G3). Then we follow those parents to their own childhood homes, identifying their parents (G2). Finally, we repeat this step once more, tracing G2 individuals back to their childhood homes to find their parents (G1). We can do this back to the 1850 Census, as censuses before this one did not enumerate the entire household by name or record occupations.

To address transitory measurement error, we use a second observation of the same individual. For each parent, grandparent, and great-grandparent, we link them not only to their child but also to another point in their own adult life. The "main" (or endogenous) ancestor observation in our regression is the one taken when the child is living in the same household (the bolded cases in Figure 1). The repeated-measures IV method uses a separate adult observation of the same ancestor as the instrument (the italicized cases in Figure 1). Once again, we emphasize that the purpose of this IV method is to address measurement error, not to identify a treatment effect.

We limit the sample as follows. First, we restrict the fourth generation to individuals aged 25 to 60, with an average age of 34. Modern studies often focus on mid-life outcomes to reduce life-cycle bias (Nybom and Stuhler, 2017). However, this bias is likely less important for historical occupation data, as occupational trajectories differ from earnings trajectories (Ward, 2023), though we will test for life-cycle bias using a method from Modalsli and Vosters (2024). We keep only lineages where the child co-resides with family at age 14 or younger, since older children remaining at home may disproportionately represent farmer families. Additionally, we keep only co-residing parents aged 20 to 60. When we use a second observation as an instrument for the first (typically from an adjacent census), it must also fall within the 20 to 60 age range.

One challenge in building multi generational links is that descendants of enslaved individuals cannot be traced back to the 1850 and 1860 censuses. To address this, we trace these families back three generations, with the earliest generation appearing in the 1870 and 1880 censuses. For Black grandfathers who lived in the South in these censuses, we impute their father's "occupation" as "enslaved," and assign them the lowest status measure (effectively 0 on a 0-100 scale, explained in the next section). We then append these imputed records with the directly linked four-generation samples.

However, linking Black families involves more errors than linking white families, potentially making their estimates less accurate or representative. But including Black individuals is important for capturing overall persistence. Mathematically, the impact of including Black families on relative mobility estimates is driven almost entirely by between-group differences rather than within-group variation (see Ward (2023) for more detailed discussion). This is because within-Black variation of status was small, as almost the entire group was concentrated in the bottom percentiles, and therefore the within-Black mobility estimates do not have a lot of variation to contribute to the overall slope. As a result, the impact of linking errors on our overall estimates, which impact the within-Black transmission estimates, is likely to be minor for overall mobility estimates.

After putting together our sample, we observe 2.5 million great-grandchildren to their great-grandfathers. This is the largest four-generation sample used in AR(3) regressions to date, far larger than any in the existing literature (see Figure 2). The large sample size is key for precisely estimating ancestor effects.

2.2 Custom weights

To address potential bias in our sample, which includes only individuals who can be linked back four generations, we create custom weights to make our sample more representative. Following

Bailey, Cole and Massey (2020), we predict the probability of being included in our four-generation sample based on the great-grandchild's observable characteristics such as age, education/literacy, race, marital status, urban status, gender, and region. We then create weights using these predicted probabilities, with the most substantial adjustment being a five-fold upweighting of Black individuals to correct for their underrepresentation in the linked sample. This weighting successfully reduces observable differences between our linked sample and the general population in both 1930 and 1940 censuses. Further details on our weighting methodology are provided in Appendix B.

2.3 Imputing relative economic status

We estimate mobility using a 0–100 scale of economic status inferred from occupation, race, and region of residence. Unlike traditional historical methods which base status only on occupation, our approach captures significant inequalities within occupations, which is important for accurately measuring mobility (Saavedra and Twinam, 2020). Our primary measure is the "adjusted Song score," developed by Ward (2023) and built upon the method from Song et al. (2020). This measure ranks individuals by percentile according to the mean literacy rate within their decadal birth cohort, occupation, race, and region from the 1850-1930 Censuses. For the 1940 Census, we use mean years of education by these same characteristics.

Measuring the relative status of occupations based on human capital levels is common in sociology, whereas economists typically use income or wealth. However, the advantage of using human capital over income-based measures is that human capital information is consistently available across censuses from 1850 to 1940, while income and wealth is not. This consistency allows for tracking changes in relative status of certain occupations, such as the declining relative status of farmers. Income-based methods, by contrast, rely on detailed occupational income data that first became available in 1940, making earlier projections difficult and potentially unreliable. Indeed, estimates of intergenerational mobility depend heavily on how occupational status is imputed. To show the robustness of our estimates to various occupational measures, we will also use the Song score (Song et al., 2020), as well as income scores predicted from observable characteristics (occupation, race, sex, and location) using data from the 1940 Census, the first to systematically record wages for wage earners, as used in prior research (Collins and Wanamaker, 2022; Saavedra and Twinam, 2020; Abramitzky et al., 2021). We are also able to estimate persistence when using more detailed measures, such as directly comparing the fourth-generation's education in 1940 to their ancestor's status, or comparing home values in 1930 and 1940 to ancestor's wealth between 1850 and 1870. Across these various methods, we consistently find that great-grandparents' outcomes are strongly predictive of the great-grandchild's.

2.4 First descriptive statistics

We provide descriptive statistics of our sample in Table 1. We have 2.5 million G4 observations, linked to 1.6 million different fathers, 1.0 million grandfathers, and 0.7 million great-grandfathers. The first and fourth generations are born an average of 85 years apart, with the first generation born in 1819 on average, and the fourth generation in 1904. Our sample is 52% male, though we find essentially no difference in the transmission of status by gender, similar to Buckles et al. (2024*b*). This suggests high assortative mating as the wife's economics status during our period was largely determined by the occupation of her husband. Second, the Black share is 11%, which matches the population, but this is entirely because our reweighting procedure assigns Black individuals a much higher weight than whites.

There was a noticeable shift in the occupation distribution as the economy developed. Most great-grandfathers were farmers (62%), with few in professional occupations (7%) or clerical/sales roles (1%). Over time, the share of farmers fell significantly, reaching 14% by the fourth generation (recall we use the husband's occupation if the fourth-generation observation is female). Instead, there was growth in low-skilled manual jobs (factory operatives) and service positions, along with an increase in professional and clerical/sales jobs. This shift reflects broader economic changes driven by technological advances, potentially weakening the transfer of occupation-specific human capital across generations (for example, as farming skills valued initially became less relevant later).

3. Results for white families

3.1 Measurement error amplifies excess persistence but attenuates four-generational transmission

We first present estimates of multigenerational mobility for white families. We then add Black families to the sample, which allows us to account for the role of racial group effects.[16]

First, to mimic the results from the literature, we estimate the G4-G1 association when using a single snapshot of the great-grandfather, despite this method suffering from measurement error from transitory shocks. Using the adjusted Song score and a single-year snapshot, we find a slope of 0.23, implying that roughly one-quarter of status differences across white great-grandfathers persisted to their great-grandsons (represented by the triangle markers in the first panel of Figure 3).

[16] See binscatter plots between great-grandfather and great-grandchild by race in Figure A1.

If status depended only on the prior generation, then this 0.234 G4-G1 association should be the product of the G4-G3, G3-G2, and G2-G1 associations. We estimate each of these slopes for adjacent generations, allowing them to vary by generation.

When multiplying them together, the predicted great-grandfather/great-grandchild association is 0.10 (see the circle markers in Figure 3). Because the predicted association of 0.10 is smaller than the actual grandfather-grandson of 0.23, the difference suggests that there is significant "excess persistence" across multiple generations. That is, convergence to the mean is slower across multiple generation than predicted from 2-generation data. Later, we will estimate the AR(3) regression where we directly include the father, grandfather, and great-grandfather in the same regression, and will find positive grandparent and great-grandparent effects.

The difference between the *predicted* G4-G1 association (0.10) and *actual* one (0.23) is very large, suggesting that 2-generational estimates badly miss the mark on long-run mobility; the results imply that status transmission between a great-grandfather and great-grandchild is 134% stronger than predicted from two-generational data. This finding–that there was a large amount of excess persistence–is the essential conclusion of much of the historical research on multigenerational mobility.

These results support the claim that great-grandparents and grandparents "mattered" in intergenerational transmission. However, the point estimates may be biased by significant measurement error (Modalsli and Vosters, 2024). To address this, rather than immediately applying a repeated-measurement IV approach, we first average the two socioeconomic status measures of fathers with each other, the two grandfather observations with each other, and the two great-grandfather observations with each other. Based on the statistical discussion, this averaging increases the reliability ratio from $\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2+\sigma_\varepsilon^2}$ to $\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2+\frac{\sigma_\varepsilon^2}{2}}$, thereby reducing transitory measurement error. While this adjustment improves precision, it does not fully resolve measurement error or address permanent mismeasurement of a latent factor. If transitory error were small, averaging should leave coefficient estimates unchanged. Yet Figure 3 reveals that averaging the great-grandfather's status increases the G4–G1 transmission estimate from 0.234 to 0.274. This result aligns with prior findings in historical data (Ward, 2020, 2023), confirming that transitory measurement error attenuates estimates.

While it is expected that the G4-G1 transmission coefficient increases from 0.23 to 0.27 after averaging two observations, the more important finding is that the gap between the actual and predicted G4-G1 association narrows significantly. After averaging, the G4-G1 association is only 64 percent stronger than the predicted G4-G1 slope (0.167), which is computed by multiplying the G4-G3, G3-G2, and G2-G1 associations when two parental outcomes are averaged. This is a

substantial reduction from the 134 percent estimate obtained without averaging. The gap between the predicted G4-G1 transmission and actual G4-G1 transmission shrinks because measurement error attenuates each G4-G3, G3-G2, and G2-G1 estimate, and this attenuation compounds when their product is taken. Otherwise, if the true model was that grandparents "mattered" by having a direct causal impact on the grandchild, then this gap should not decrease. But reducing measurement error decreases the overall attenuation bias, as shown in Equation 8.

When using the repeated-IV measure to fully eliminate transitory error, where we instrument the first great-grandfather observation with the second, the G4-G1 association increases further, reaching 0.33.[17] This is 41 percent higher than the estimate from a single snapshot (0.23). This result essentially reproduces the findings of Ward (2023), which identified measurement error in historical census data when using two-generation data. However, the novel contribution here is new evidence of *long-run*, four-generational persistence in economic status. If one-third of initial gaps across white great-grandfathers persisted to great-grandchildren, this contradicts the assertion of Becker and Tomes (1986) that nearly all economic differences disappear within three generations. It also implies that assortative mating was strong throughout American history; otherwise, gaps between families should dwindle more rapidly (Stuhler, 2024). We return to this point on assortative mating later.

The second key result is that accounting for transitory measurement error nearly eliminates evidence of excess persistence. We re-estimate the product of each two-generation associations using IV in each generation. After this adjustment, the actual great-grandfather to great-grandson association (0.331) becomes just 4% stronger than the predicted coefficient (0.317), a marginal difference that undermines earlier claims of multigenerational influence. Importantly, this implies that great-grandfathers and grandfathers did not “matter” as much as previously argued, beyond what we would expect given single-generation transmission (with the caveat that these estimates remain noncausal). This lack of excess persistence under IV estimation contrasts starkly with the original OLS approach. When using single-generation snapshots of fathers, grandfathers, and great-grandfathers, OLS produced a G4–G1 transmission rate 126% stronger than predicted, dramatically overstating multigenerational effects.

We can show that the transmission of status between great-grandparent and great-grandchild holds when using alternative proxies for status (see Table 2). Instead of using an occupation measure that adjusts for race and region, we can use an occupation-only measure from (Song et al., 2020). This yields the same one-third estimate. We can also use educational attainment for all individuals

[17] The results are very similar if we assume that classical measurement error holds and we infer the amount of error from the amount of variation in one snapshot relative to an average of two. See Figure A2a.

(including women) in 1940, eliminating the need to infer status from the spouse's occupation. Regressing 1940 educational attainment (percentile ranked) on the adjusted Song score yields a G4-G1 slope of 0.248, implying an AR(1) estimate of 0.63, which is less than but close to our preferred estimate (0.69). Similarly, we can measure relative status based on neighborhood quality, as proxied by the status of nearby neighbors (see Buckles et al. (2024*b*)), which produces an AR(1) estimate of 0.72.[18]

We can also estimate how wage income from 1940 is associated with the great-grandfather's adjusted Song score, which gives a slope of 0.203 (AR[1] = 0.59). This sample is a selected one, however, as wages are only observed for wage workers, dropping farmers and business owners in the fourth generation. If we drop this same group for our primary regression using the Adjusted Song score, this yields a lower AR[1] slope of 0.65 than our main estimate of 0.69, suggesting that limiting the sample to wage workers reduces persistence estimates.

Finally, we use wealth measures (real property values from 1850–1870 and home values from 1930–1940). These associations are imperfect, since many individuals did not own property or homes, making their latent productive capability unobservable. Among those with positive wealth in both the first and fourth generations, the rank-rank slope is 0.347 (AR[1] = 0.703). When we include individuals with zero wealth, the slope falls to 0.234 (AR[1] = 0.616).[19] This drop occurs because descendants of those with zero property in the 19th century have higher home values than those whose great-grandfathers held small amounts of property. This pattern likely reflects negative selection among those with minimal holdings. For example, descendants of small-scale farmers had lower home values than descendants of urban workers with no real estate. A similar non-linear pattern in wealth mobility over two generations has been documented in Sweden (Black et al., 2020). Finally, the rank-rank slope between 1940 wage income and 1850–1870 wealth is 0.201, yielding an AR[1] estimate of 0.59.

3.2 An AR(3) model of long-run mobility

A different way of showing excess persistence is to directly estimate the grandfather and great-grandfather effects in an AR(3) model (Equation 3), the results of which are graphically shown in Figure 4.

[18] In this measure, one's status is the average of the adjusted Song scores of the 10 closest household heads listed either above or below an individual on census manuscripts. Non-relatives (like domestic servants) are assigned their own adjusted Song score.

[19] These are estimated using 2SLS.

The first column (OLS) shows large effects from grandfather and great-grandfather status when using single snapshots, which ignores transitory error. The estimated effects are 0.338 for fathers, 0.136 for grandfathers, and 0.046 for great-grandfathers, all statistically significant. This naive approach that ignores transitory measurement error yields a "surprising" positive great-grandparent effect: given the average 80-year gap in birth years across four generations, most great-grandfathers had died before the birth of the great-grandchild. Summing the effects together, these coefficients imply a total ancestral effect of 0.52. This combined estimate is 54% larger than the father effect alone, suggesting again that two-generation models understate the amount of long-run persistence.

However, as we now know from the prior section, these grandfather and great-grandfather effects are inflated due to transitory measurement error in historical data. As Modalsli and Vosters (2024) explain, measurement error in the father's status causes "spillover bias" on the grandfather effect, inflating the coefficient. Intuitively, the grandfather and great-grandfather effects are not capturing large direct effects on the grandchild; instead, they are simply providing additional information about the father's latent factor. If this is correct, reducing measurement error by averaging or using IV should increase the father effect while decreasing the effects of the grandfather and great-grandfather. But if these ancestor effects are real and not just artifacts of spillover bias, they should increase when using an average of two measurements, as this reduces measurement error in their status estimates.

After averaging two observations, the results are consistent with the transitory measurement-error model: the father effect increases in magnitude (from 0.34 to 0.42), while the grandfather effect decreases slightly (from 0.14 to 0.13) and the great-grandfather effect drops by almost half (from 0.05 to 0.03). The total ancestral effect (father + grandfather + great-grandfather) rises from 0.52 to 0.58. Notably, the gap between the father effect and the total ancestral effect narrows after averaging: initially, the total effect was 54% larger than the father effect alone, but after averaging, it is only 36% larger. This indicates that averaging observations reduces the relative influence of older generations (grandfather and great-grandfather) compared to the father effect. Once again, this result reflects the significant measurement error in historical data.

Finally, when applying IV to fully eliminate transitory measurement error, the father effect increases sharply (from 0.34 to 0.60), the grandfather effect is halved (from 0.14 to 0.06), and the great-grandfather effect shrinks further and turns negative (from 0.03 to a statistically significant -0.005). The total ancestral effect (sum of all three) reaches 0.65, which is only 9% larger than the father effect alone. This total is also 25% larger than the corresponding estimate from the first column (using single snapshots), demonstrating that even with four generations of data, the long-run ancestral effect remains underestimated unless multiple observations are used. As such,

accurately capturing the long-run transmission of status requires both extended lineage data and repeated measurements.

Another test for the role of measurement error is to improve only the father's measurement, either by averaging observations or with the repeated-measurement instrument, while leaving the grandfather and great-grandfather measured with a single, noisy snapshot. Figure A3 shows what happens: once the father's measure improves, his coefficient rises, while the coefficients on the grandfather and great-grandfather fall.

3.3 Robustness of AR(3) estimates

A standard concern in mobility studies is life-cycle bias: if generations are observed at different ages, or not near mid-life, estimates of persistence can be biased (Haider and Solon, 2006; Nybom and Stuhler, 2017). To check this, Modalsli and Vosters (2024) propose reversing the endogenous and exogenous variables in the IV model. Since these are measured 10–20 years apart, any life-cycle bias would show up as a shift in estimates. But as Figure A4 shows, the results remain steady. The total ancestor effect stays around 0.66, suggesting that life-cycle bias is not a major concern.

Another potential issue is that our status measure, which imputes status within occupation by region, may overstate persistence if it puts too much weight on enduring regional inequalities. To address this, we use the occupation-only score from Song et al. (2020), which ignores regional variation within occupations. When we estimate the AR(3) model with this score (Figure A5), we find broadly similar patterns. The total ancestor effect is larger than the father effect when using a single snapshot (0.522 compared to 0.332), and it increases when using IV. Contrary to the idea that including region inflates persistence, the total ancestor effect is actually larger (0.77) with the occupation-only score. In contrast to our preferred measure, the occupation-only measure still has large grandparent and great-grandparent effects when using IV, suggesting greater measurement error in occupation-only scores compared to our more detailed status measure.

We also use log predicted wage income from the 1940 Census, similar to Collins and Wanamaker (2022), and find a total ancestor effect of 0.73 and relatively small (but positive) grandparent and great-grandparent effects (See Figure A6). While this score has the advantage of being based on income, rather than human capital, it requires us to back-cast 1940 relative earnings to 1850.

Instead of relying on a univariate measure of occupational status, we test whether broad occupational categories of ancestors help predict a grandchild's occupation. In other words, does it matter not just that your father was white-collar, but that your grandfather and great-grandfather

were too? Figure A7 explores this using five categories: professional/managerial, clerical/sales, farmer, semi-skilled, and unskilled. We find two main results. First, a basic OLS model shows strong effects from both grandfathers and great-grandfathers. For professional workers, including ancestors raises the total ancestor effect by 76% compared to using only the father's occupation; for unskilled workers, the increase is 60%. The gains are smaller for semi-skilled workers and farmers, but still positive. Second, when we average across outcomes, the effects of both father and ancestors grow larger.[20] Overall, these results suggest that long-run mobility is lower than implied by two-generation estimates, and that single snapshots miss the underlying, unobserved, status of fathers.

A large share of our sample is descended from farmers, so we examine how results change when we exclude them (Figure A8). Dropping farmer ancestors reduces overall persistence: the total ancestor effect for the white population falls from 0.65 to 0.56. At the same time, excess persistence becomes more pronounced, with the total ancestor effect now 32% larger than the father effect, compared to 9% in the full sample. If the latent factor model is correct, then this suggests that there is more permanent measurement error in the non-farming population than for the full population, leading to stronger grandparent and great-grandparent effects. We show evidence consistent with this point later when we decompose measurement error into permanent and transitory components.

3.4 Assortative Mating

One limitation of our data is that we do not directly observe mothers' impact on their children. However, the strong persistence of social status across four generations suggests that if mothers mattered for child outcomes, assortative mating had to be strong. Otherwise, marriages between parents of different statuses would have caused rapid regression to the mean (see Appendix C for an explanation using a latent variable model, following Stuhler (2024)). If one argues that assortative mating was weak, then this is the same as arguing that only fathers mattered, perhaps because they trained their children in job-specific skills or had occupation-based networks. But this view conflicts with evidence that mothers also played an important role in shaping child's outcomes in American history (Althoff, Brookes Gray and Reichardt, 2025; Espín-Sánchez, Ferrie and Vickers, 2025).

Our data further show that multigenerational transmission (G4-G1) is equally strong whether traced through paternal or maternal lines (father's father's father, father's mother's father, etc.; Table 3). If mothers dominated transmission, we would expect stronger links through the mother's

[20] We do not apply the repeated IV method here because binary 0/1 variables violate the classical measurement error assumption, as any error is likely negatively correlated with the true value. Using IV with non-classical measurement error in the independent variable causes upward bias. But if the dependent variable is also measured with error, the direction of bias is ultimately unclear.

mother's father, but we do not see that. Rather, the slopes across four great-grandfathers are all similar in size (ranging from 0.33 to 0.35), suggesting that mothers and fathers matter equally to child outcomes, and assortative mating was high.

We can use our data to show that assortative mating appears to have been strong based on the strength of "transmission" between one's ancestors to one's *spouse*, who is from a different family background. Table 4 shows that the strength of association between first and four generations is very similar when switching between one's own and one's spouse's outcomes. First, recall that for women we use the husband's status as their outcome because most women did not work during this time period; that is, we use the great-grandson-in-law. Table 4 shows that the slope coefficient for the actual great-grandson (0.351) is nearly the same as for the great-grandson-in-law (0.342). This result is surprising as the great-grandson-in-law did not come from the same family investments during childhood. Rather, it suggests that spouses had very similar backgrounds.

We observe a similar pattern when using years of education in 1940, ranked by percentiles to make a straightforward comparison across measures. The slope between G4 education rank and G1 status is 0.251, which is close to the 0.234 slope between an individual's great-grandparent and their spouse's education. Education is useful in this context because it is measured for both men and women, so the strength of the association is not driven by proxying women's outcomes using the husband's occupation.

If the rate of assortative mating was actually 0.58, as suggested by the direct regression of an individual's education rank on their spouse's, most models of multigenerational mobility would predict a much weaker association with in-laws, about 40% less than the direct association with their own great-grandparent. However, the association in our data is not much weaker. This result suggests that the direct assortative mating regression understates the true extent of matching on latent characteristics, possibly due to measurement error in the 1940 education variable or match on other unobservables.[21]

3.5 Direct contact from grandparents

The additional persistence from the small grandparent effect could result from a direct causal impact of the grandfather on the grandchild. To test this, we divide the sample into two groups:

[21] Using the framework in Espín-Sánchez, Ferrie and Vickers (2025) with multigenerational data, one can estimate assortative mating. When we apply this method using Proposition 4 from that paper, we obtain implausibly high assortative mating estimates—greater than one. This could reflect a violation of this model's assumptions, such as the assumption that parent-child transmission parameters are stable across generations, which does not appear to hold in our data.

(1) grandchildren whose grandparents were alive at their birth and (2) those whose grandparents had already passed away. We determine grandparental death dates using data from FamilySearch, similar to the data used by Black et al. (2023). We merge these death dates into our dataset and split the sample accordingly.[22] This split results in 9% of white sample having their grandparent die before they were born. Others have used the same approach (Belloc et al., 2024; Braun and Stuhler, 2018; Helgertz and Dribe, 2022; Neidhöfer and Stockhausen, 2019).

The results show that the grandparent effect is not weaker for grandchildren whose grandparents had died before their birth (Figure 6). While the coefficients differ from our main analysis due to sample changes—limited to cases with observable grandparental death dates—the grandparent effect is actually larger for those whose grandparents had already passed away (0.09 vs. 0.06). This suggests that a large causal impact from the grandparent is unlikely to drive the result. At the same time, we do not have exogenous variation in grandparental death, and so we cannot make definitive causal claims (Goldstein, 2025). Nevertheless, this evidence is consistent with results using exogenous variation in grandparental wealth (Ager, Boustan and Eriksson, 2021; Bleakley and Ferrie, 2016)

3.6 Underlying transmission coefficient for white families based on latent variable model

If the small grandparent effect is not caused by a direct impact on the grandchild, it may be due to a group effect that spans generations, such as families living in the same high- or low- opportunity places. We will return to the issue of group effects later. For now, we assume the observed effect results from "permanent" measurement error that has not yet been addressed. If the latent variable model is correct, then the repeated-measure IV still does not correct for this error, since the reliability ratio is $\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2}$.

Braun and Stuhler (2018) point out that if the latent variable model holds, one can identify λ by taking the ratio of direct estimates across generations. If the G4–G1 slope is $\beta_{-3} = \theta_{G3}\lambda^3$ and the G3–G1 slope is $\beta_{-2} = \theta_{G3}\lambda^2$, then their ratio is $\frac{\beta_{-3}}{\beta_{-2}} = \lambda$. This is equivalent to using the great-grandfather's status as an instrument for the father's status (Lindahl et al., 2015). The numerator is the reduced form (G4–G1 slope), and the denominator is the first stage (G3–G1 slope). We refer to this as the "ratio" estimate to distinguish it from the repeated-measure IV estimate.

Table 5 shows that the ratio method estimates a transmission rate of 0.68 for the white population. This is slightly higher than the repeated-measure IV estimate of 0.63, which instruments one father's

[22] We successfully identified death dates for approximately 70% of grandfathers in our dataset. As shown later in Figure 6, this subsample with known death dates has modestly higher persistence estimates (2% increase in the total ancestor effect) compared to the overall sample presented in Figure4.

status with another, and much higher than the OLS estimate of 0.41. Applying the latent variable model to decompose the estimates, these results suggest that 35% of the variation in a single snapshot comes from transitory error, while 4.9% comes from permanent error. We find similar patterns when standardizing the outcomes in each generation and estimate correlations instead of covariances.[23] Thus, in our data, it is transitory error, not permanent error in latent ability, that explains much of the grandparent effect. This result aligns with evidence from Norway (Modalsli and Vosters, 2024), where using an IV eliminates the grandparent effect.

A key advantage of the ratio estimate is that it does not require multiple observations for each father, grandfather, and great-grandfather. This allows us to use data with just one observation per generation, expanding the sample from 2.5 million to 4.2 million. This addresses the possible concern that the strict sample requirements used earlier cause us to have an unrepresentative sample, even after weighting. But the ratio estimate does not change much in the larger sample (0.683 vs. 0.701; see "Expanded Sample" row in Table 5), suggesting that the stricter requirements do not bias the mobility estimate.

We also estimate results using only non-agricultural families and find that they have more permanent measurement error than the full sample. In Table 5, the ratio estimate for non-farmer descendants is nearly identical to that of the full sample (0.684 vs. 0.683). This is surprising because both the OLS and IV estimates are lower for the non-farming group, which would suggest higher mobility. The gap between the ratio estimate and the OLS/IV estimates points to more permanent error in the non-agricultural sample, which pulls both the OLS and IV estimates down more than in the full sample. One reason for this may be that in non-farm settings, latent status is less tightly linked to observable outcomes like occupation due to the rapidly expanding set of occupations in the service sector. If so, this implies that modern-day relative mobility estimates (at least those based on occupational outcomes) may be more affected by measurement error than in the past.

3.7 Comparison to surname method

Clark (2014) proposes estimating intergenerational status transmission (λ) using rare surnames to reduce measurement error. Applying this method to our white-only sample (chosen to address concerns about racial group effects and missing enslaved surnames in the first generation), we estimate a G4-G1 slope of 0.298 (Table A3), translating to an implied AR(1) coefficient of 0.667, which is slightly below our main model's 0.69–0.70 range. While the surname method yields results

[23] The parent-child correlation is 0.38. The ratio of the parent-child correlation to the correlation between two observations of the same parent is 0.58. The ratio of the G4-G1 and G4-G3 correlations is 0.63. Together, these suggest that most of the noise comes from transitory error.

consistent with our main estimates, we prefer our approach using directly linked microdata. We discuss the surname approach in more detail in Appendix D.

4. Including Black families and group effects

We have found a small grandparent effect, or that excess persistence across four generations is relatively small. Besides permanent error, a grandparent effect could be due to persistent group differences, such as a slowly converging Black-white gap (Solon, 2018). However, our results so far have only used the white population to discount such a group effect. Recall the discussion of the latent model with group effects in Section 1.4, which shows that an AR(1) group effect will (1) increase the slope estimates across generations (assuming the transmission of the group effect is stronger than within-group transmission), and (2) increase excess persistence.

When we add Black families to sample, we find results consistent with strong and persistent group effects. That is, enduring racial disparities significantly affect long-run mobility estimates. Using the adjusted Song score, the grandfather-grandson association rises by 54%, from 0.33 to 0.51, when Black families are included (see "white" vs. "Black and white" estimates in Figure 5a).

This increase is driven by strong between-race persistence of about 1.00 throughout this period: according to our data, the Black-white gap in ranks remains at about 40 in both the first and fourth generations. Because the share of between-race variation was also large in the first generation (roughly 25%), a weighted average of 1.00 between-group persistence (multiplied by 0.25) and roughly 0.33 within-group persistence (multiplied by 0.75) yields the G4–G1 slope of 0.51. This result is the multigenerational counterpart to prior findings that show that including Black individuals raises two-generational persistence estimates in American history (Hertz, 2008; Jácome, Kuziemko and Naidu, 2025; Ward, 2023).

Consistent with predictions from our latent variable model with group effects, we also observe significantly stronger excess persistence in the pooled Black-white sample than in the white-only sample. The estimated G4-G1 persistence coefficient of 0.510 for the pooled sample exceeds the predicted value (multiplying each G4-G3, G3-G2 and G2-G1 coefficients) of 0.432 by 18%, a stark contrast to the white-only sample, where the observed value exceeded the predicted value by just 3%. This pattern is consistent with the model's result where if within-group transmission λ differs from the transmission of group effects (γ), excess persistence occurs. In our data, the persistence of group means was about 1.0 (i.e., Black-white gaps did not converge across generations), which is a magnitude far exceeding the 0.67–0.70 transmission rate within the white population. Thus, we find excess persistence, which confirms the importance of group effects as argued by Solon (2018).

These dynamics become even clearer in the AR(3) model (Figure 5b). When including Black families, we detect positive effects from *both* grandparents and great-grandparents—a pattern absent in the white-only analysis. The total ancestor effect (summing father, grandfather, and great-grandfather coefficients) rises to 0.766, exceeding the father-only effect by 21%. This result demonstrates that conventional two-generation mobility estimates do indeed underestimate long-run status transmission in populations with enduring group disparities.[24]

Figure A8 shows results after excluding farmer descendants from the Black and white sample. In the non-agricultural sample, the effects of grandparents and great-grandparents are stronger than in the full sample. The total ancestor effect (0.699) is now 54% larger than the father effect (0.453). The great-grandfather effect remains large and significant, even when using repeated measures and IV. As discussed before, this could reflect higher measurement error in the non-agricultural sector because a given occupation does not capture latent status well. If permanent measurement error has increased over time as the economy shifted away from farming and racial disparities also continue to persist, this suggests that modern-day two-generational estimates in the United States may understate persistence across multiple generations.

5. Conclusions

In this paper, we estimate multigenerational mobility in the United States from 1850 to 1940 using a unique four-generation dataset that links over 2.5 million individuals to their ancestors. In addition to being orders of magnitude larger than most datasets used in previous studies, we observe multiple status observations per generation, providing a deeper understanding of long-run persistence in economic status while also addressing substantial measurement error in historical data.

We present four main findings. First, despite massive structural changes over this period, we document significant persistence. We estimate that 33% of economic differences across white great-grandparents remained in their great-grandchildren. This figure increases to 51% when Black families are included. Similar to recent papers on multigenerational mobility (Stuhler, 2024), these results challenge the conventional wisdom that relative advantages disappear within three generations.

Second, two-generation estimates capture multigenerational mobility fairly well for the white population. In the AR(3) model that includes the father, grandfather, and great-grandfather, the grandfather and great-grandfather effects are quantitatively small. The grandparent effect is about

[24] See Table A4 for the results when instrumenting the father with the grandfather's outcome. Note that this no longer identifies the latent transmission parameter λ in the presence of group effects. See Appendix E.

9% the size of the father effect, suggesting only modest additional persistence. Using records on grandparents who were alive or dead by the time of the grandchild's birth, we further show that direct contact between grandparents and grandchildren is unlikely to explain this positive effect. Instead, the effect may reflect measurement error that repeated observations do not fully fix.

Third, naive estimates of grandparent and great-grandparent effects are substantially inflated by transitory measurement error in historical data. When we address this error through repeated observations and instrumental variables approaches, the apparent direct influence of distant ancestors diminishes significantly for the white population. Our corrected estimates suggest that two-generational data captures long-run mobility relatively well for whites but less well for the pooled Black and white population.

Fourth, group-level inequalities, particularly racial disparities, substantially influence multigenerational mobility patterns. The persistent Black-white gap throughout this period creates stronger multigenerational transmission when both groups are analyzed together, demonstrating how structural barriers limit relative mobility across multiple generations. It also increases the size of grandparent and great-grandparent effects in AR(3) models, indicating that for the whole population, two-generational models do not accurately predict the convergence of gaps across multiple generations.

The results have several important implications for mobility research. One is that strong transmission across multiple generations, coupled with evidence that maternal and paternal ancestral lines equally predict outcomes, suggests that assortative mating was strong in American history. High levels of marital sorting likely helped preserve status across generations, explaining why advantages diminished slowly. Moreover, when between-group differences are large and persist at higher rates than within-group differences, two-generational estimates will likely underestimate mobility across multiple generations (Solon, 2018). Given the persistent lack of Black-white convergence over the past 80 years (Bayer and Charles, 2018) and continuing wide racial disparities, modern two-generational income mobility estimates likely understate the true persistence of longer-run mobility patterns.

References

Abramitzky, Ran, Leah Boustan, Elisa Jácome, and Santiago Pérez. 2021. "Intergenerational mobility of immigrants in the United States over two centuries." *American Economic Review*, 111(2): 580–608.

Adermon, Adrian, Mikael Lindahl, and Daniel Waldenström. 2018. "Intergenerational wealth mobility and the role of inheritance: Evidence from multiple generations." *The Economic Journal*, 128(612): F482–F513.

Adermon, Adrian, Mikael Lindahl, and Mårten Palme. 2021. "Dynastic human capital, inequality, and intergenerational mobility." *American Economic Review*, 111(5): 1523–1548.

Ager, Philipp, Leah Boustan, and Katherine Eriksson. 2021. "The intergenerational effects of a large wealth shock: White southerners after the Civil War." *American Economic Review*, 111(11): 3767–3794.

Althoff, Lukas, Harriet Brookes Gray, and Hugo Reichardt. 2025. "America's Rise in Human Capital Mobility." Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5188024.

Anderson, Lewis R., Paula Sheppard, and Christiaan WS Monden. 2018. "Grandparent effects on educational outcomes: A systematic review." *Sociological Science*, 5: 114–142.

Bailey, Martha, Connor Cole, and Catherine Massey. 2020. "Simple strategies for improving inference with linked data: A case study of the 1850–1930 IPUMS linked representative historical samples." *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 53(2): 80–93.

Bailey, Martha J, Connor Cole, Morgan Henderson, and Catherine Massey. 2020. "How well do automated linking methods perform? Lessons from US historical data." *Journal of Economic Literature*, 58(4): 997–1044.

Bayer, Patrick, and Kerwin Kofi Charles. 2018. "Divergent paths: A new perspective on earnings differences between Black and white men since 1940." *The Quarterly Journal of Economics*, 133(3): 1459–1501.

Becker, Gary S., and Nigel Tomes. 1986. "Human capital and the rise and fall of families." *Journal of Labor Economics*, 4(3, Part 2): S1–S39.

Belloc, Marianna, Francesco Drago, Mattia Fochesato, and Roberto Galbiati. 2024. "Multigenerational Transmission of Wealth: Florence, 1403–1480." *American Economic Journal: Applied Economics*, 16(2): 99–129.

Black, Sandra E, Neil Duzett, Adriana Lleras-Muney, Nolan G Pope, and Joseph Price. 2023. "Intergenerational correlations in longevity." National Bureau of Economic Research.

Black, Sandra E, Paul J Devereux, Petter Lundborg, and Kaveh Majlesi. 2020. "Poor little rich kids? The role of nature versus nurture in wealth and other economic outcomes and behaviours." *The Review of Economic Studies*, 87(4): 1683–1725.

Bleakley, Hoyt, and Joseph Ferrie. 2016. "Shocking behavior: Random wealth in antebellum Georgia and human capital across generations." *The Quarterly Journal of Economics*, 131(3): 1455–1495.

Braun, Sebastian Till, and Jan Stuhler. 2018. "The transmission of inequality across multiple generations: testing recent theories with evidence from Germany." *The Economic Journal*, 128(609): 576–611.

Buckles, Kasey, Adrian Haws, Joseph Price, and Haley E.B. Wilbert. 2024*a*. "Breakthroughs in historical record linking using genealogy data: The Census Tree Project." National Bureau of Economic Research Working Paper 31671.

Buckles, Kasey, Joseph Price, Zachary Ward, and Haley E.B. Wilbert. 2024*b*. "Family Trees and Falling Apples: Historical Intergenerational Mobility Estimates for Women and Men." National Bureau of Economic Research Working Paper 31918.

Celhay, Pablo, and Sebastian Gallegos. 2025. "Schooling mobility across three generations in six Latin American countries." *Journal of Population Economics*, 38(1): 1–35.

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2020. "Race and economic opportunity in the United States: An intergenerational perspective." *The Quarterly Journal of Economics*, 135(2): 711–783.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

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

Clark, Gregory. 2023. "The inheritance of social status: England, 1600 to 2022." *Proceedings of the National Academy of Sciences*, 120(27): e2300926120.

Clark, Gregory, and Neil Cummins. 2015. "Intergenerational wealth mobility in England 1858–2012: Surnames and social mobility." *The Economic Journal*, 125(582): 61–85.

Colagrossi, Marco, Béatrice d'Hombres, and Sylke V Schnepf. 2020. "Like (grand) parent, like child? Multigenerational mobility across the EU." *European Economic Review*, 130: 103600.

Collado, M Dolores, Ignacio Ortuño-Ortín, and Jan Stuhler. 2023. "Estimating intergenerational and assortative processes in extended family data." *The Review of Economic Studies*, 90(3): 1195–1227.

Collins, William J, and Marianne H Wanamaker. 2022. "African American intergenerational economic mobility since 1880." *American Economic Journal: Applied Economics*, 14(3): 84–117.

Corak, Miles. 2013. "Income inequality, equality of opportunity, and intergenerational mobility." *Journal of Economic Perspectives*, 27(3): 79–102.

Davis, Jonathan MV, and Bhashkar Mazumder. 2024. "The decline in intergenerational mobility after 1980." *Review of Economics and Statistics*, 1–47.

Derenoncourt, Ellora, Chi Hyun Kim, Moritz Kuhn, and Moritz Schularick. 2024. "Wealth of two nations: The US racial wealth gap, 1860–2020." *The Quarterly Journal of Economics*, 139(2): 693–750.

Dribe, Martin, and Jonas Helgertz. 2016. "The lasting impact of grandfathers: Class, occupational status, and earnings over three generations in Sweden 1815–2011." *The Journal of Economic History*, 76(4): 969–1000.

Engzell, Per, Carina Mood, and Jan O. Jonsson. 2020. "It's All about the Parents: Inequality Transmission across Three Generations in Sweden." *Sociological Science*, 7: 242–267.

Espín-Sánchez, José-Antonio, Joseph Ferrie, and Chris Vickers. 2025. "The Mismeasure of Women: Intergenerational Mobility and the Econometrics of Family Trees." Manuscript, March 2025.

Ferrie, Joseph, Catherine Massey, and Jonathan Rothbaum. 2016. "Do grandparents and great-grandparents matter? Multigenerational mobility in the US, 1910-2013." National Bureau of Economic Research.

Ferrie, Joseph, Catherine Massey, and Jonathan Rothbaum. 2021. "Do grandparents matter? Multigenerational mobility in the United States, 1940–2015." *Journal of Labor Economics*, 39(3): 597–637.

Goldstein, Ezra G. 2025. "A Mighty Toll: Mine Accidents and the Long-Run Effect of Losing a Father Among Sons." *Journal of Human Resources*.

Gray, Rowena, Siobhan M O'Keefe, Sarah Quincy, and Zachary Ward. 2024. "Tasks and Black-white inequality over the long twentieth century." National Bureau of Economic Research.

Haider, Steven, and Gary Solon. 2006. "Life-cycle variation in the association between current and lifetime earnings." *American Economic Review*, 96(4): 1308–1320.

Hällsten, Martin, and Martin Kolk. 2023. "The shadow of peasant past: Seven generations of inequality persistence in Northern Sweden." *American Journal of Sociology*, 128(6): 1716–1760.

Helgertz, Jonas, and Martin Dribe. 2022. "Do grandfathers matter for occupational and earnings attainment? Evidence from Swedish register data." *European Sociological Review*, 38(1): 54–72.

Hertz, Tom. 2008. "A group-specific measure of intergenerational persistence." *Economics Letters*, 100(3): 415–417.

Hodge, Robert W. 1966. "Occupational mobility as a probability process." *Demography*, 3: 19–34.

Jácome, Elisa, Ilyana Kuziemko, and Suresh Naidu. 2025. "Mobility for all: Representative intergenerational mobility estimates over the twentieth century." *Journal of Political Economy*, 133(1): 306–354.

Kalsi, Priti, and Zachary Ward. 2025. "The Gilded Age and beyond: The persistence of elite wealth in American history." National Bureau of Economic Research.

Knigge, Antonie. 2016. "Beyond the parental generation: The influence of grandfathers and great-grandfathers on status attainment." *Demography*, 53(4): 1219–1244.

Lindahl, Mikael, Mårten Palme, Sofia Sandgren Massih, and Anna Sjögren. 2015. "Long-term intergenerational persistence of human capital an empirical analysis of four generations." *Journal of Human Resources*, 50(1): 1–33.

Lindert, Peter H., and Jeffrey G. Williamson. 2016. *Unequal Gains: American Growth and Inequality since 1700.* Princeton University Press.

Long, Jason, and Joseph Ferrie. 2018. "Grandfathers Matter(ed): Occupational Mobility Across Three Generations in the US and Britain 1850–1911." *The Economic Journal*, 128(612): F422–F445.

Mare, Robert, and Xi Song. 2014. "Social mobility in multiple generations." *UCLA CCPR Population Working Papers*.

Mare, Robert D. 2011. "A Multigenerational View of Inequality." *Demography*, 48(1): 1–23.

Margo, Robert A. 2016. "Obama, Katrina, and the persistence of racial inequality." *The Journal of Economic History*, 76(2): 301–341.

Mazumder, Bhashkar. 2005. "Fortunate sons: New estimates of intergenerational mobility in the United States using Social Security earnings data." *Review of Economics and Statistics*, 87(2): 235–255.

Modalsli, Jørgen. 2023. "Multigenerational persistence: Evidence from 146 years of administrative data." *Journal of Human Resources*, 58(3): 929–961.

Modalsli, Jørgen, and Kelly Vosters. 2024. "Spillover bias in multigenerational income regressions." *Journal of Human Resources*, 59(3): 743–776.

Neidhöfer, Guido, and Maximilian Stockhausen. 2019. "Dynastic inequality compared: Multigenerational mobility in the United States, the United Kingdom, and Germany." *Review of Income and Wealth*, 65(2): 383–414.

Nybom, Martin, and Jan Stuhler. 2017. "Biases in Standard Measures of Intergenerational Income Dependence." *Journal of Human Resources*, 52(3): 800–825.

Olivetti, Claudia, M. Daniele Paserman, and Laura Salisbury. 2018. "Three-generation mobility in the United States 1850–1940: The role of maternal and paternal grandparents." *Explorations in Economic History*, 70: 73–90.

Pfeffer, Fabian T, and Alexandra Killewald. 2018. "Generations of advantage. Multigenerational correlations in family wealth." *Social Forces*, 96(4): 1411–1442.

Price, Joseph, Kasey Buckles, Adrian Haws, and Haley Wilbert. 2023. "The Census Tree." Distributed on 2023-08-10.

Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren. 2024. "IPUMS Ancestry Full Count Data: Version 4.0." [dataset].

Saavedra, Martin, and Tate Twinam. 2020. "A machine learning approach to improving occupational income scores." *Explorations in Economic History*, 75: 101304.

Santavirta, Torsten, and Jan Stuhler. 2024. "Name-based estimators of intergenerational mobility." *The Economic Journal*, ueae035.

Shiue, Carol H. 2019. "Social mobility in the long run: a temporal analysis of China from 1300 to 1900."

Shiue, Carol H, and Wolfgang Keller. 2024. "Elite Strategies for Big Shocks: The Case of the Fall of the Ming." National Bureau of Economic Research Working Paper 33121.

Solon, Gary. 1992. "Intergenerational income mobility in the United States." *The American Economic Review*, 393–408.

Solon, Gary. 2014. "Theoretical Models of Inequality Transmission Across Multiple Generations." *Research in Social Stratification and Mobility*, 35: 13–18.

Solon, Gary. 2018. "What Do We Know So Far About Multigenerational Mobility?" *The Economic Journal*, 128(612): F340–F352.

Song, Xi, Catherine G Massey, Karen A Rolf, Joseph P Ferrie, Jonathan L Rothbaum, and Yu Xie. 2020. "Long-term decline in intergenerational mobility in the United States since the 1850s." *Proceedings of the National Academy of Sciences*, 117(1): 251–258.

Stuhler, Jan. 2012. "Mobility Across Multiple Generations: The Iterated Regression Fallacy." Institute of Labor Economics (IZA).

Stuhler, Jan. 2024. "Multigenerational inequality." *Research Handbook on Intergenerational Inequality*, 100–121.

Torche, Florencia, and Alejandro Corvalan. 2018. "Estimating intergenerational mobility with grouped data: A critique of Clark's The son also rises." *Sociological Methods & Research*, 47(4): 787–811.

Vosters, Kelly, and Martin Nybom. 2017. "Intergenerational persistence in latent socioeconomic status: Evidence from Sweden and the United States." *Journal of Labor Economics*, 35(3): 869–901.

Wanamaker, Marianne H. 2017. "150 Years of economic progress for African American men: Measuring outcomes and sizing up roadblocks." *Economic History of Developing Regions*, 32(3): 211–220.

Ward, Zachary. 2020. "The Not-So-Hot Melting Pot: The Persistence of Outcomes for Descendants of the Age of Mass Migration." *AEJ: Applied Economics*.

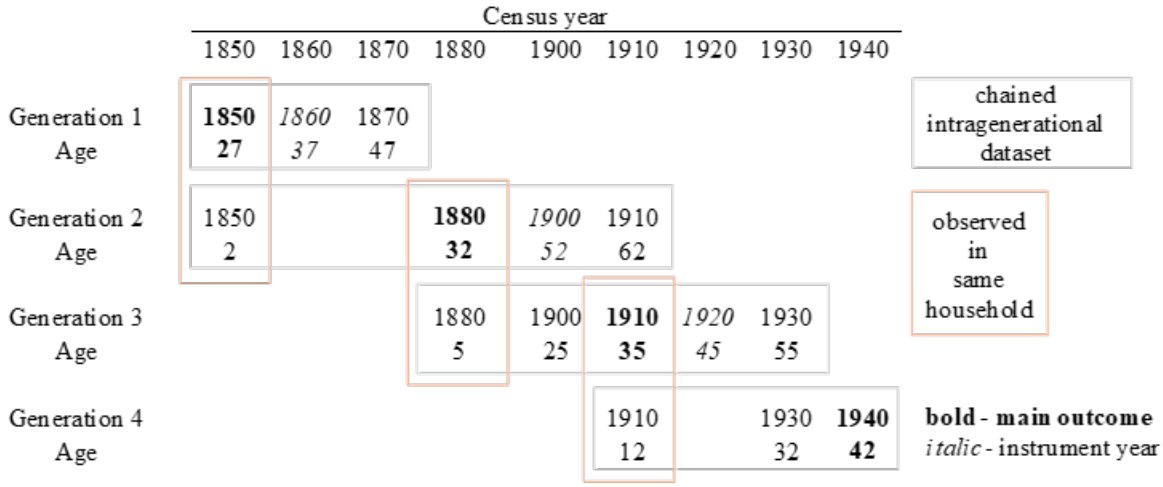
Ward, Zachary. 2023. "Intergenerational mobility in American history: Accounting for race and measurement error." *American Economic Review*, 113(12): 3213–48.

Wright, Gavin. 1986. *Old South, new South: Revolutions in the southern economy since the Civil War.* Basic Books.

Xie, Yu, and Alexandra Killewald. 2013. "Intergenerational occupational mobility in Great Britain and the United States since 1850: Comment." *American Economic Review*, 103(5): 2003–2020.

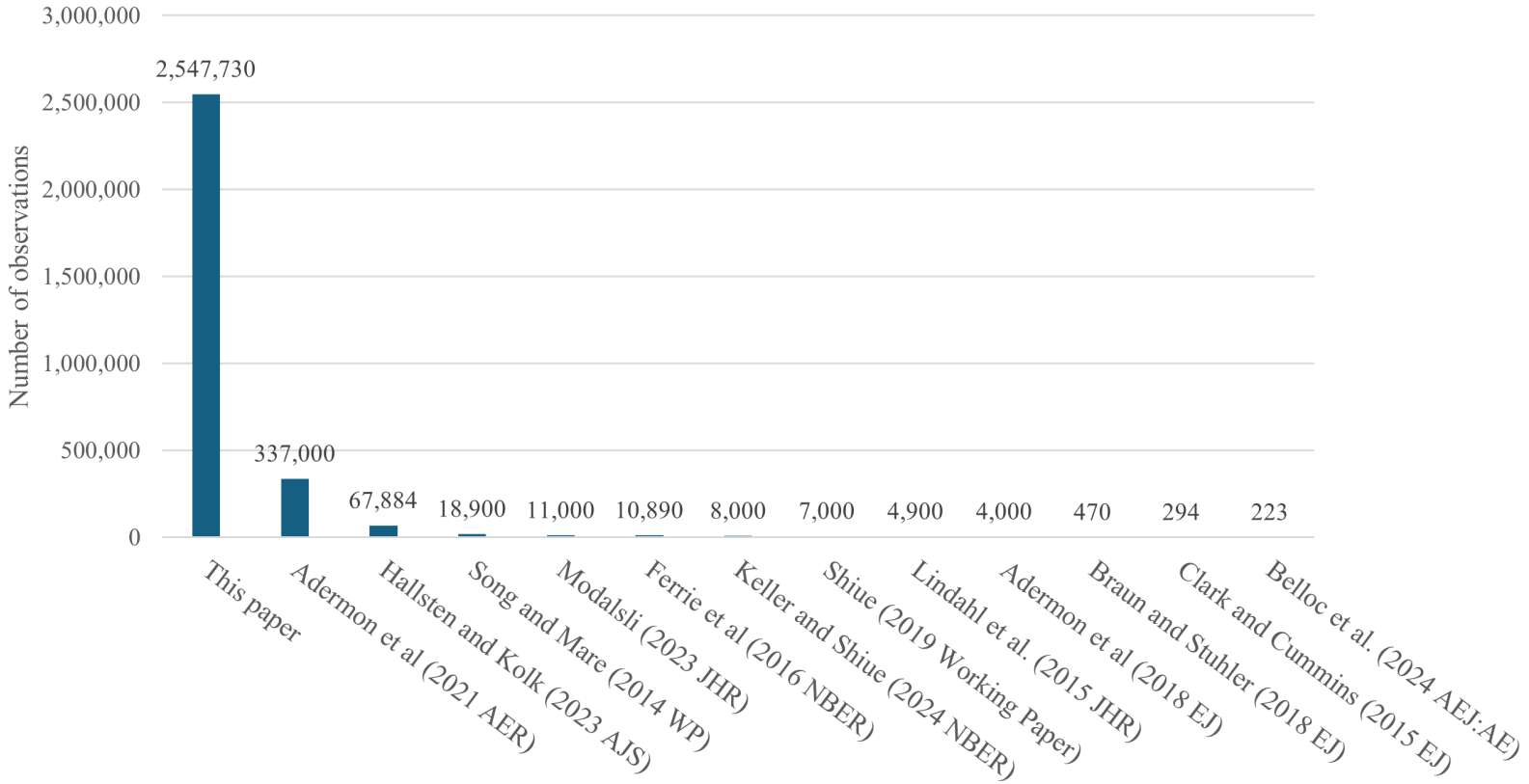
Figures and Tables

Figure 1: Building multigenerational samples



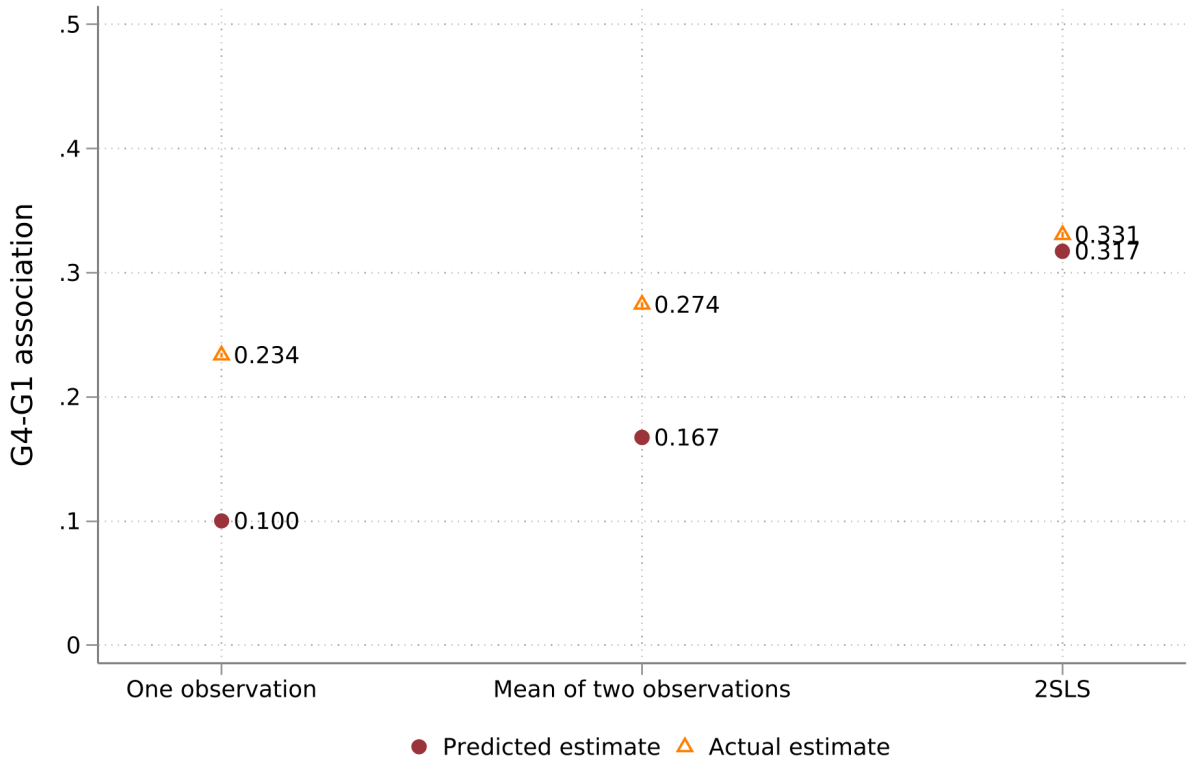
Notes: The figure shows the process of linking, using an example family. The links are from the Census Tree (Price et al., 2023). The census data is from (Ruggles et al., 2024). First, we chain links from the Census Tree to build the "chained intragenerational dataset," which is shown with the gray box. This dataset tracks an individual over their whole life. For individuals in this dataset, we attach parental outcomes when they are at most 14 years old (orange box). We then merge the parent into the intragenerational sample to observe both (1) their childhood household and (2) a second adult observation for an instrumental variables strategy. The 3-generational dataset are G3 parents who we can additionally track back to their childhood home to attach a G2 parent. The 4-generational sample repeats this process to find the G1 great-grandparent. When estimating mobility in a regression, we use the final generation's outcome when they are closest to age 40. For ancestor's generations, we use the outcome when they are observed with the child (bold observations). We also use an instrumental variables strategy to account for measurement error. For this strategy, we use the nearest observation up to 20 years away (italicized years). The 1890 Census is missing because the manuscripts were lost in a fire. While the figure shows a single parent, grandparent and great-grandparent, our process finds all ancestors on the maternal and paternal side, causing the dataset to observe up to 8 great grandparents, 4 grandparents and 2 parents. When measuring economic status, we always use the occupational status of the male. If the great-grandchild is a woman, we use the husband's spouse's occupation.

Figure 2: Our four-generational sample is large relative to the literature



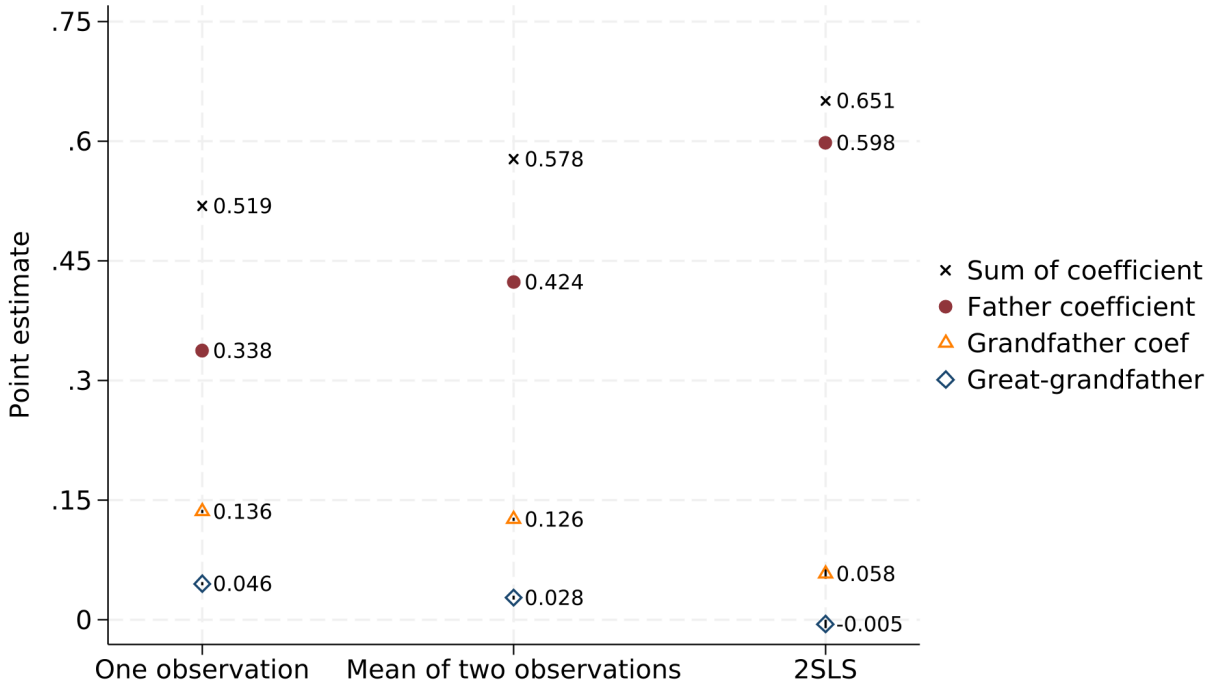
Notes: Our sample is large relative to the literature. The papers referenced are Adermon, Lindahl and Palme (2021); Hällsten and Kolk (2023); Mare and Song (2014); Modalsli (2023); Ferrie, Massey and Rothbaum (2016); Shiue and Keller (2024); Shiue (2019); Lindahl et al. (2015); Adermon, Lindahl and Waldenström (2018); Braun and Stuhler (2018); Clark and Cummins (2015); Belloc et al. (2024). The sample sizes are based on AR(3) regressions where each generation is included.

Figure 3: Predicted v. actual great-grandparent-great-grandchild association for white families



Notes: Data are from a linked sample of white families (see Figure 1) (Ruggles et al., 2024; Price et al., 2023). The status measure is the adjusted Song score, which ranks occupation, race, and region cells by their human capital (Song et al., 2020; Ward, 2023). The "actual estimate in data" points are from regressing the great-grandson's outcome on the great-grandfather's outcome. The "predicted estimate" points are the product of the associations for the great-grandfather-grandfather, grandfather-father, and father-child, which are estimated separately. The first column shows results when using a single observation of the father, grandfather, and great-grandfather. The second one averages two observations. The third uses an IV estimate where the first observation is instrumented with a second. The sample is the same across all estimates, and is weighted to be representative on observables.

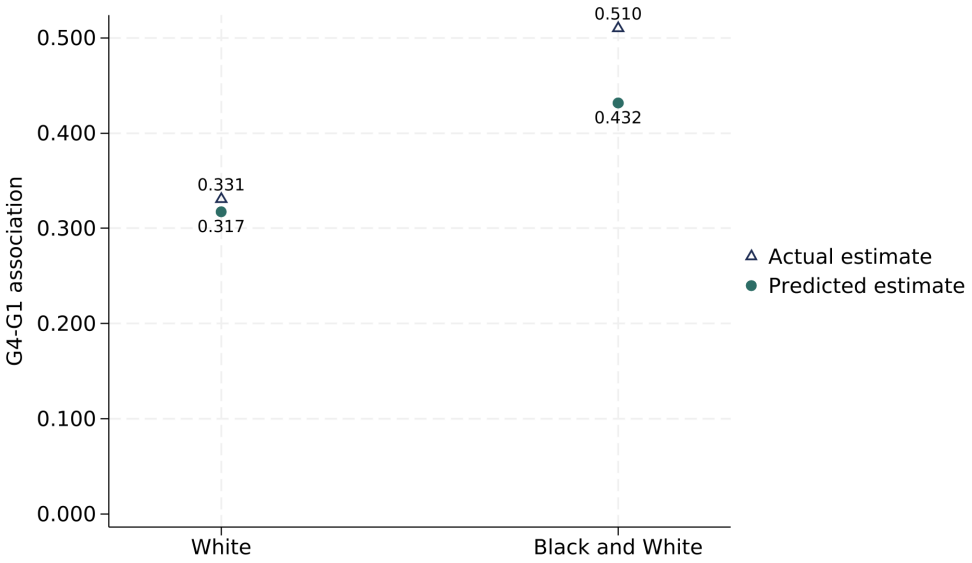
Figure 4: An AR(3) model of multigenerational mobility for white families



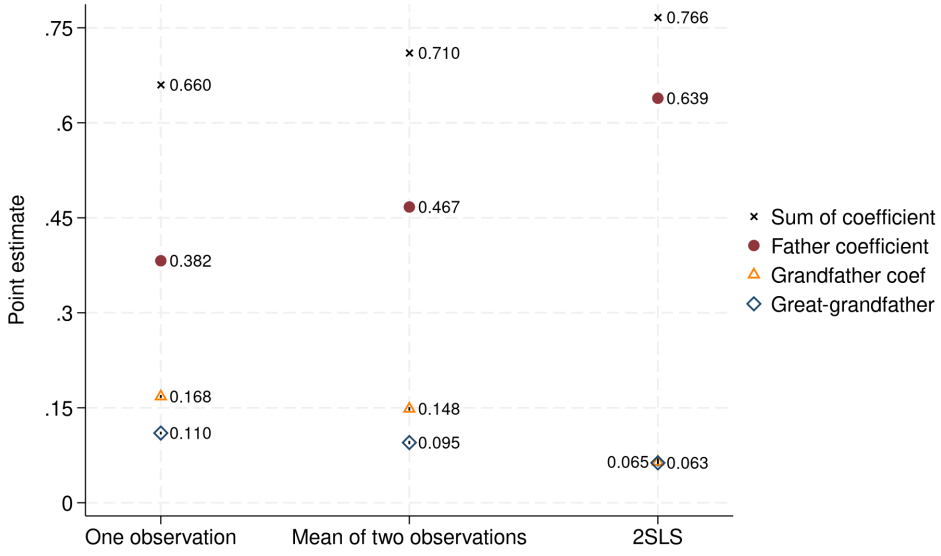
Notes: Data are from a linked sample of great-grandchildren to their great-grandfathers (see Figure 1) (Ruggles et al., 2024; Price et al., 2023). The sample includes white families. The status measure is the adjusted Song score, which ranks occupation, race, and region cells by their human capital (Song et al., 2020; Ward, 2023). The one observation column uses a single G1 outcome, a single G2 outcome, and a single G3 outcome. The mean of two observations uses the mean of the two G1, G2, and G3 outcomes. The 2SLS column instruments the first G1, G2, and G3 outcome with the second observation. The sample is the same across all estimates, and is weighted to be representative on observables.

Figure 5: Racial disparities are important for multigenerational estimates

(a) Adding Black families increases both inter- and multi-generational transmission estimates

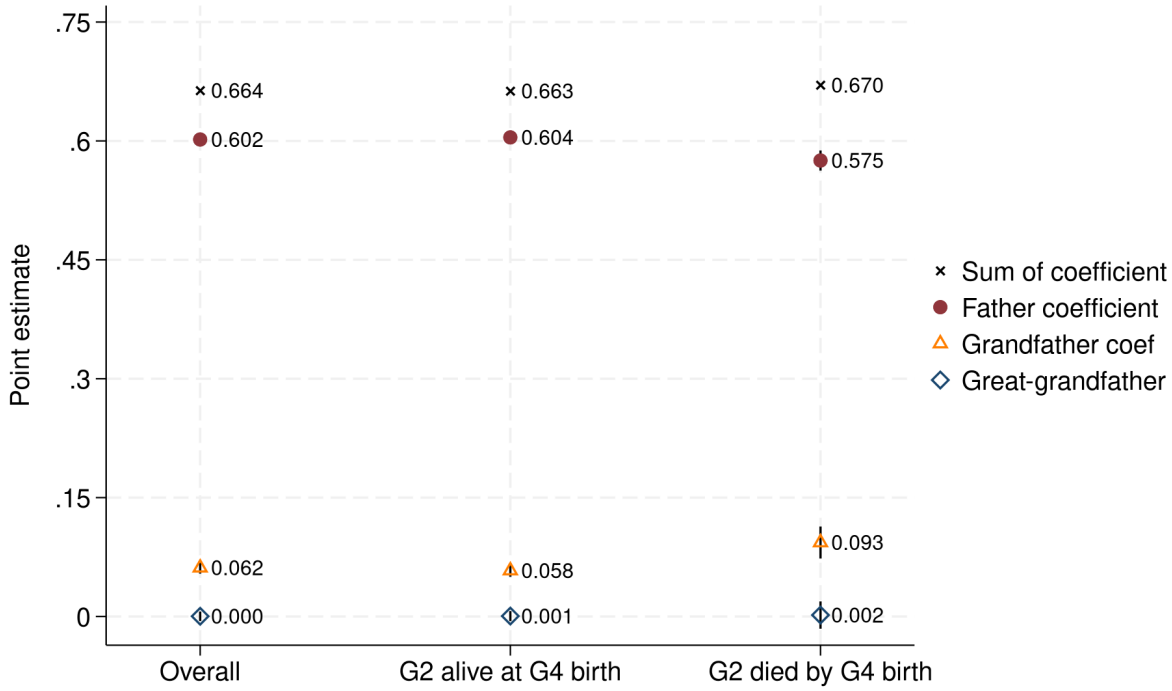


(b) Adding Black families increases evidence for excess persistence



Notes: Panel (a) shows the great-grandparent-great-grandchild association with the white-only sample, and then after adding in Black families. See Figure 3 for additional details. Panel (b) shows AR(3) estimates after moving from a single observation to an average of two observations to an IV estimate. See Figure 4 for additional details.

Figure 6: Grandfather effects (G2) are not weaker for those who died before the grandchild's (G4) birth



Notes: Data are from the 1850-1940 Censuses (Ruggles et al., 2024), linked by (Price et al., 2023). This figure shows the AR(3) estimates after splitting the sample by grandparents who died before the grandchild's birth ("no overlap") or after ("overlap"). The sample is only for those who we can find death dates on FamilySearch.org, and therefore differ from the main effects presented in Figure 4. About 9 percent of great-grandsons (G4) had their grandfathers (G2) die by their birth. The samples are weighted to be representative on observables.

Table 1: Descriptive statistics for the analytic sample

	Mean	Std. Dev.	Min	Max	N
Age	35	6	25	52	2,547,730
Birth Year	1904	7	1888	1915	2,547,730
Male	0.52	0.50	0.00	1.00	2,547,730
Black	0.11	0.31	0.00	1.00	2,547,730
Adjusted Song score	53.08	28.96	0.01	100.00	2,547,730
Professional	0.17	0.38	0.00	1.00	2,547,730
Farmer	0.14	0.34	0.00	1.00	2,547,730
Clerical/Sales	0.14	0.35	0.00	1.00	2,547,730
Craftsmen	0.15	0.35	0.00	1.00	2,547,730
Low-skill manual/service	0.40	0.49	0.00	1.00	2,547,730
Father observables:					
Age	35	8	20	60	1,653,301
Birth Year	1874	9	1817	1900	1,653,301
Adjusted Song score	47.67	27.81	0.02	100.00	1,653,301
Professional	0.10	0.31	0.00	1.00	1,653,301
Farmer	0.44	0.50	0.00	1.00	1,653,301
Clerical/Sales	0.06	0.24	0.00	1.00	1,653,301
Craftsmen	0.13	0.34	0.00	1.00	1,653,301
Low-skill manual/service	0.26	0.44	0.00	1.00	1,653,301
Grandfather observables:					
Age	37	8	20	60	1,018,493
Birth Year	1845	12	1791	1880	1,018,493
Adjusted Song score	46.54	26.84	-0.06	100.00	1,018,493
Professional	0.08	0.26	0.00	1.00	1,018,493
Farmer	0.58	0.49	0.00	1.00	1,018,493
Clerical/Sales	0.02	0.15	0.00	1.00	1,018,493
Craftsmen	0.11	0.32	0.00	1.00	1,018,493
Low-skill manual/service	0.21	0.41	0.00	1.00	1,018,493
Father's father	0.50	0.50	0.00	1.00	1,018,493
Mother's father	0.50	0.50	0.00	1.00	1,018,493
Great-Grandfather observables:					
Age	39	8	20	60	704,561
Birth Year	1819	11	1790	1860	704,561
Adjusted Song score	52.24	24.71	0.00	100.00	704,561
Professional	0.07	0.25	0.00	1.00	704,561
Farmer	0.62	0.49	0.00	1.00	704,561
Clerical/Sales	0.01	0.09	0.00	1.00	704,561
Craftsmen	0.15	0.36	0.00	1.00	704,561
Low-skill manual/service	0.15	0.36	0.00	1.00	704,561
Father's Father's father	0.25	0.43	0.00	1.00	704,561
Mother's Father's father	0.26	0.44	0.00	1.00	704,561
Father's Mother's father	0.23	0.42	0.00	1.00	704,561
Mother's Mother's father	0.27	0.44	0.00	1.00	704,561

Notes: Data are from the 1850-1940 Censuses (Ruggles et al., 2024), linked by (Price et al., 2023). The table is split by generation, and the sample is weighted to be representative of the population on observables.

Table 2: Four-generational transmission based on other measures of status

G4 Measure	G1 Measure	Actual G4-G1 estimate	Implied AR(1)	N
White sample				
Adj. Song Score	Adj. Song Score	0.331	0.692	2,505,954
Song Score	Song Score	0.335	0.695	2,505,954
Rank 1940 Years of Educ.	Adj. Song Score	0.248	0.628	2,183,595
Rank Neighborhood Status	Adj. Song Score	0.371	0.719	2,484,341
Rank 1940 Wage Income (>0)	Adj. Song Score	0.203	0.588	1,171,662
Rank 1940 Wage Income (>0)	Rank Real Property (>0)	0.201	0.585	591,733
Rank Home Value (>0)	Rank Real Property (>0)	0.347	0.703	571,818
Rank Home Value	Rank Real Property	0.234	0.616	1,182,191
Black and white sample				
Adj. Song Score	Adj. Song Score	0.510	0.799	2,547,766
Song Score	Song Score	0.347	0.703	2,547,766
Rank 1940 Years of Educ.	Adj. Song Score	0.333	0.693	2,216,072
Rank Neighborhood Status	Adj. Song Score	0.432	0.756	2,500,853
Rank 1940 Wage Income (>0)	Adj. Song Score	0.291	0.662	1,192,550
Rank 1940 Wage Income (>0)	Rank Real Property (>0)	0.203	0.588	595,744
Rank Home Value (>0)	Rank Real Property (>0)	0.344	0.701	573,378
Rank Home Value	Rank Real Property	0.306	0.674	1,199,614

Notes: Data are from the 1850-1940 Censuses (Ruggles et al., 2024), linked by Price et al. (2023). Each row is from a different regression of of the great-grandchild's outcome (G4) on the great-grandfathers (G1). Each row shows a different slope, varying depending on the measure of economic status. Adj. Song Score is our main status measure. Rank 1940 Years of Education uses the 1940 Census variable "educational attainment" and percentile ranks it. Neighborhood status is based on the average adjusted Song score of the 10 household heads above and below, where non-relative receive their own score (see Buckles et al. (2024*b*)). 1940 Wage income is only reported for wage workers. Home value is observed in 1930 and 1940 only for those who owned homes. Wealth is real property wealth in 1850, and the sum of real estate and property wealth in 1860 and 1870, and then percentile ranked in the cumulative distribution. The Implied AR(1) estimate takes the cube root of the Actual G4-G1 estimate. The samples are weighted to be representative on observables.

Table 3: G4-G1 slope is similar regardless of which grandparent is used

Regression	Coefficient	Std. Error	Obs.
Father's father father	0.334	0.002	840958
Father's mother's father	0.343	0.002	843654
Mother's father's father	0.340	0.002	887592
Mother's mother's father	0.350	0.002	910196

Notes: Data are from the 1850–1940 Censuses (Ruggles et al., 2024; Price et al., 2023). Each row uses IV from a second observation of the specific relative. The samples are weighted to be representative on observables.

Table 4: Similar "transmission" for great-grandchildren and great-grandchildren-in-law suggests strong assortative mating

Sample	Coef.	SE	N
Direct assortative mating regression			
G4 education rank on spouse's education rank	0.577	(0.001)	1,750,395
Evidence of high assortment across four generations			
G4 men (Great-Grandson Adj. Song Score) on G1	0.351	(0.002)	1,121,059
G4 women (Great-Grand-son-in-law Adj. Song Score) on G1	0.342	(0.002)	895,080
G4 education on G1 Adj. Song Score	0.251	(0.001)	1,750,395
G4 spouse's education on G1 Adj. Song Score	0.234	(0.001)	1,750,395

Notes: Data come from the 1850-1940 Censuses (Ruggles et al., 2024), linked using (Price et al., 2023). Each row reports a different regression. The top row is the great-grandchild's percentile rank of years of education regressed on their spouses. The next set of rows are regressions of a great-grandchild's outcome (G4) on their great-grandfather's (G1) status. The slope varies depending on the measure of economic status and whose status is used. The first row regresses a male G4's own adjusted Song score on his great-grandfather's score. The second row regresses a married female G4's spouse's adjusted Song score on her great-grandfather's score. The third row regresses a G4's own years of education (percentile ranked) in 1940 on their great-grandfather's adjusted Song Score. The fourth row regresses a married G4's spouse's years of education (percentile ranked) in 1940 on their great-grandfather's adjusted Song score. The samples are weighted to be representative on observables. The point of these sets of analyses is that the predictive power of a great-grandparent does not decrease by much when using the spouse's outcome, suggesting that spouses match on similar latent characteristics not observed in the direct regression.

Table 5: Underlying two-generational transmission estimates from latent model, white families

	OLS $\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2+\sigma_\varepsilon^2}\lambda$	IV $\frac{\sigma_x^2}{\sigma_x^2+\sigma_u^2}\lambda$	Ratio λ	Transitory σ_ε^2	Permanent σ_u^2	N
Adjusted occupation score						
Baseline sample	0.408	0.631	0.683	0.354	0.049	2505954
	(0.001)	(0.001)	(0.003)			
Expanded Sample	0.405	0.628	0.701	0.355	0.068	4129381
	(0.001)	(0.001)	(0.002)			
Non-agricultural families	0.304	0.489	0.684	0.378	0.178	310647
	(0.002)	(0.003)	(0.010)			
Occupation-only score						
Baseline sample	0.364	0.662	0.870	0.451	0.131	2505954
	(0.001)	(0.002)	(0.009)			
Expanded Sample	0.370	0.662	0.856	0.442	0.127	4129381
	(0.001)	(0.001)	(0.007)			
Non-agricultural families	0.337	0.551	0.728	0.389	0.148	310647
	(0.002)	(0.003)	(0.012)			

Notes: Data come from the 1940 Census (Ruggles et al., 2024) and (Price et al., 2023). Each cell in the first three columns represents a different regression. The OLS estimate comes from a regression of G4's status on a snapshot of G3's status. The IV column instruments one observation of G3's status with a second observation. The Ratio column represents the ratio of the G4-G1 slope to the G3-G1 slope, estimated using IV, where G3's status is instrumented with G1's status. The transitory column estimates the share of total variation in the father's outcome due to transitory error, while the permanent column estimates the share due to permanent error. N represents the number of observations. The baseline sample requires at least two observations for G3, G2, and G1. The expanded sample relaxes this requirement, allowing only one observation for G2 and G1 while still requiring two for G3 (for the IV estimation). The non-agricultural families sample excludes families in which any G1, G2, or G3 member is listed as a farmer. The samples are weighted to be representative on observables.

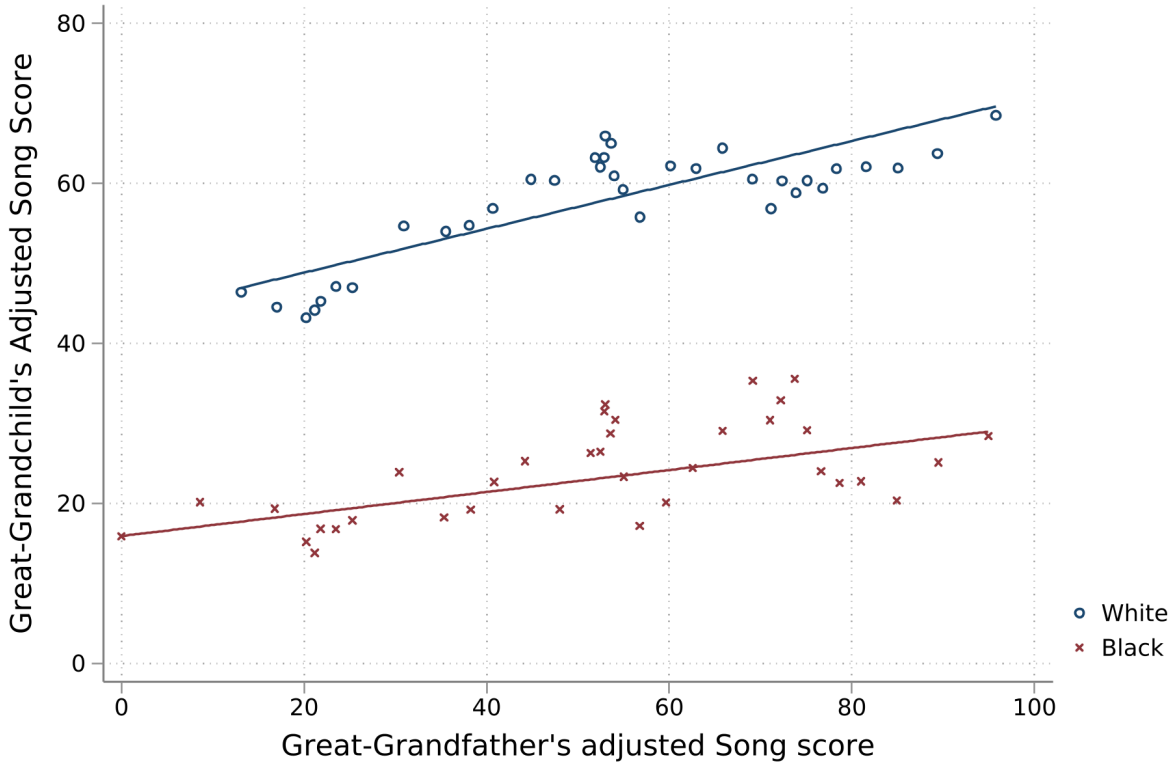
Online Appendix for

Like Great-Grandparent, Like Great-Grandchild?

Multigenerational Mobility in American History

August 20, 2025

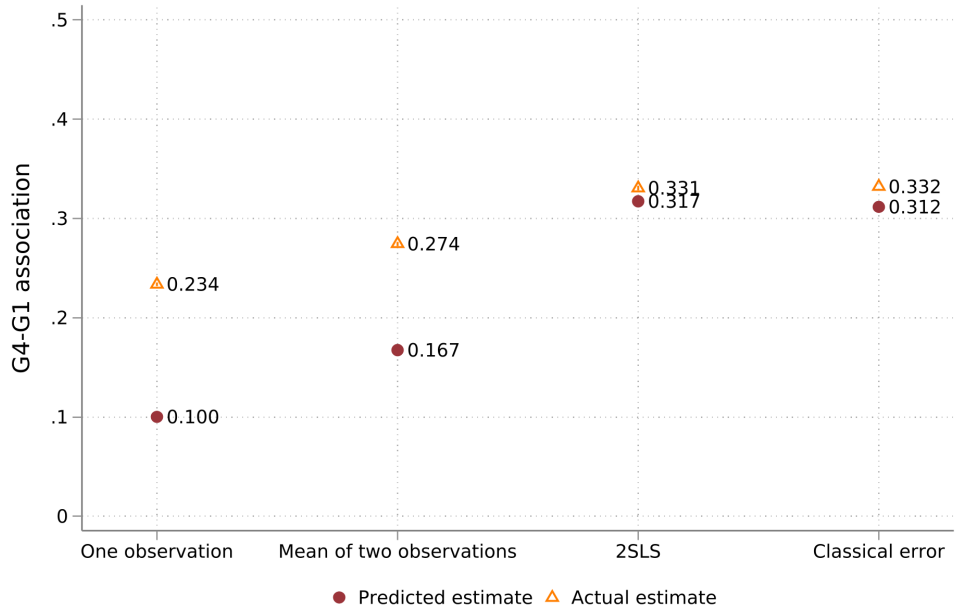
Figure A1: Black-white differences in mobility



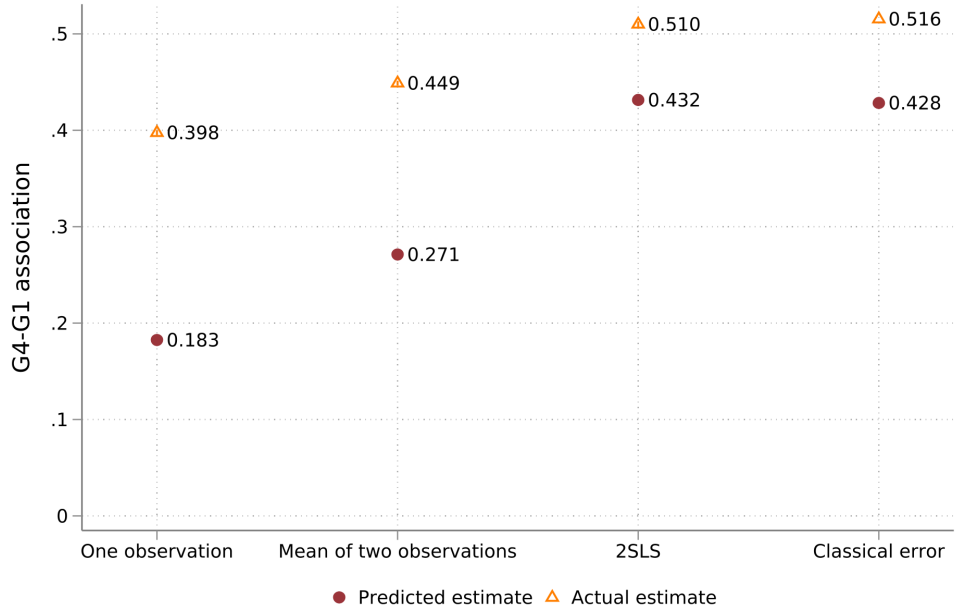
Notes: Data come from the 1850–1940 Censuses (Ruggles et al., 2024), linked using (Price et al., 2023). Because we cannot show the IV relationship, we plot the binscatter relationship based on the average of two great-grandfather's measures. The samples are weighted to be representative on observables.

Figure A2: Predicted 4-generational coefficients using classical measurement error

(a) White population

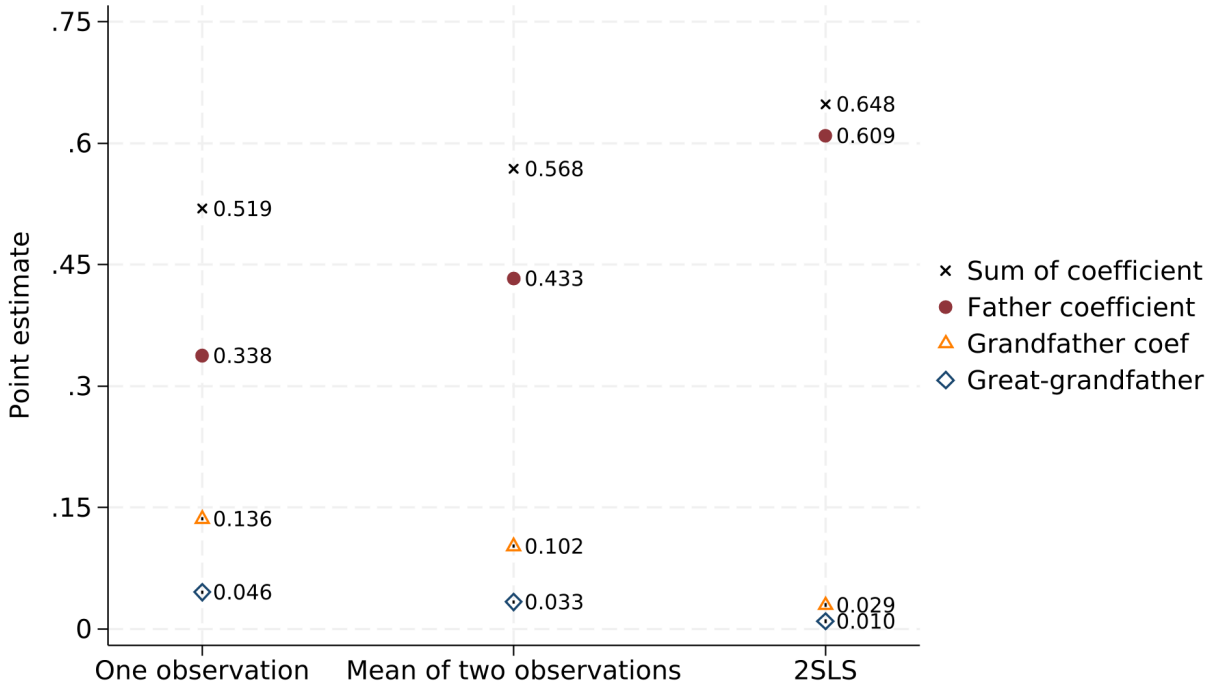


(b) Black and white



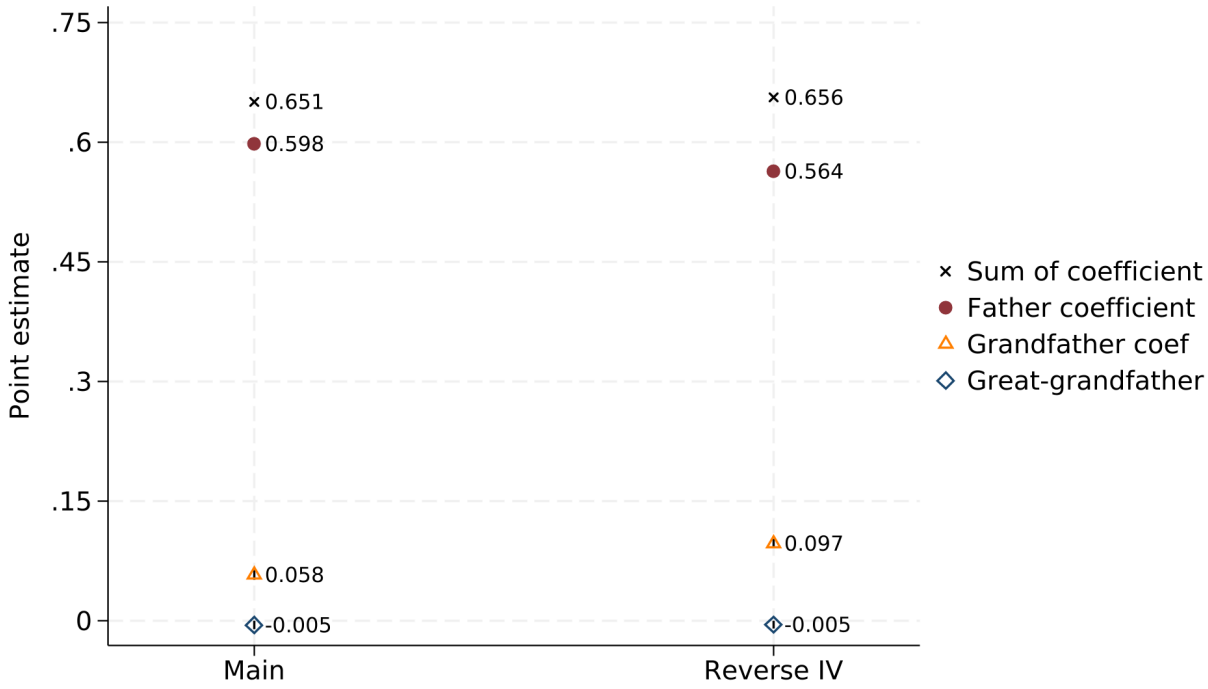
Notes: Data come from the 1850–1940 Censuses (Ruggles et al., 2024), linked using (Price et al., 2023). These mimic Figure 3 but also show the predictions when using the classical measurement error formula. The samples are weighted to be representative on observables.

Figure A3: An AR(3) model of multigenerational mobility for white families, IV for father only



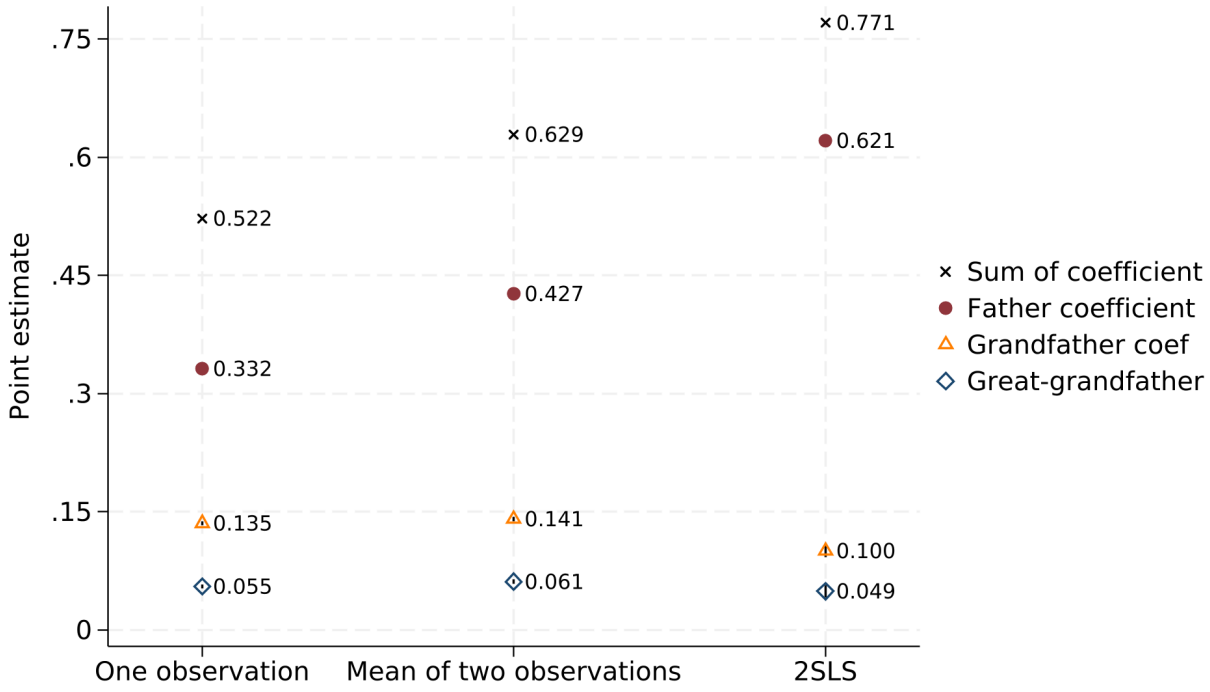
Notes: Data are from the 1850-1940 US censuses (Ruggles et al., 2024), linked by the Census Tree (Price et al., 2023). The samples are weighted to be representative on observables. The status measure is the adjusted Song score, which ranks occupation, race, and region cells by their human capital (Song et al., 2020; Ward, 2023). The one observation column uses a single G1 outcome, a single G2 outcome, and a single G3 outcome. The mean of two observations uses the mean of the two G3 outcomes, but a single of the G2 and G1. The last column only instrumnts G3 with a second observation.

Figure A4: An AR(3) model of multigenerational mobility for white families, reversing the IV



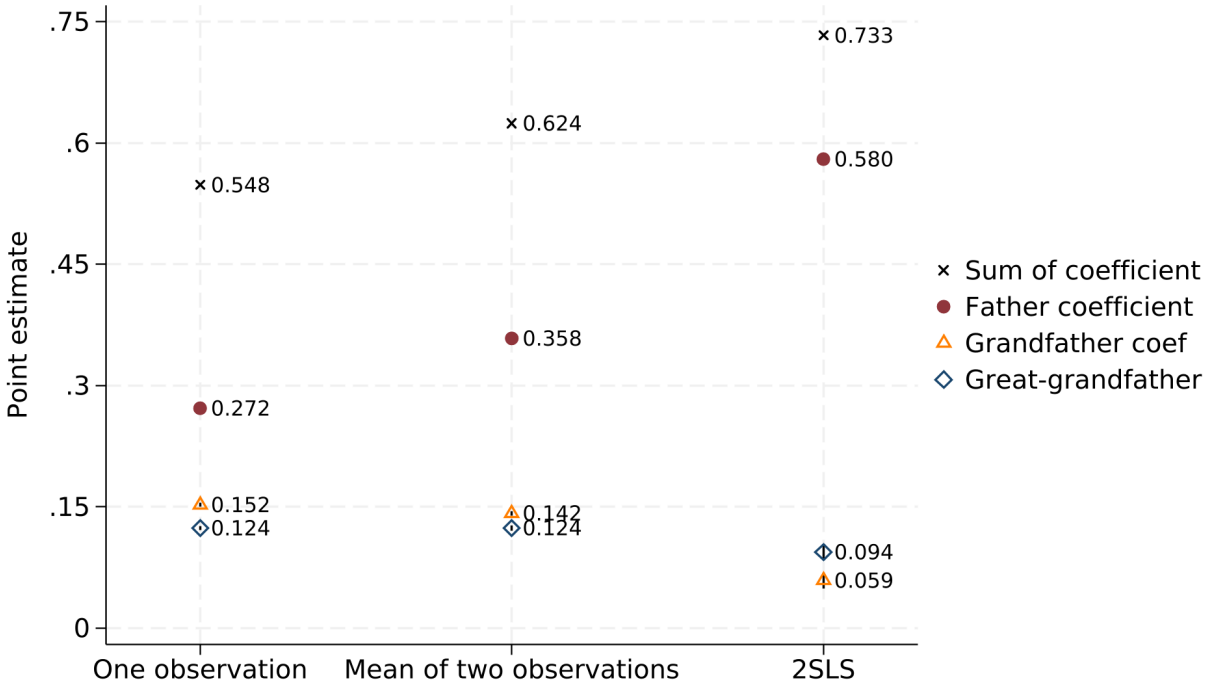
Notes: Data are from a linked sample of great-grandchildren to their great-grandfathers (see Figure 1) (Price et al., 2023; Ruggles et al., 2024). The samples are weighted to be representative on observables. The status measure is the adjusted Song score, which ranks occupation, race, and region cells by their human capital (Song et al., 2020; Ward, 2023). This shows what happens if you reverse the endogenous and exogenous variables in the IV method.

Figure A5: An AR(3) model of multigenerational mobility for white families, occupation-only score (Song et al., 2020)



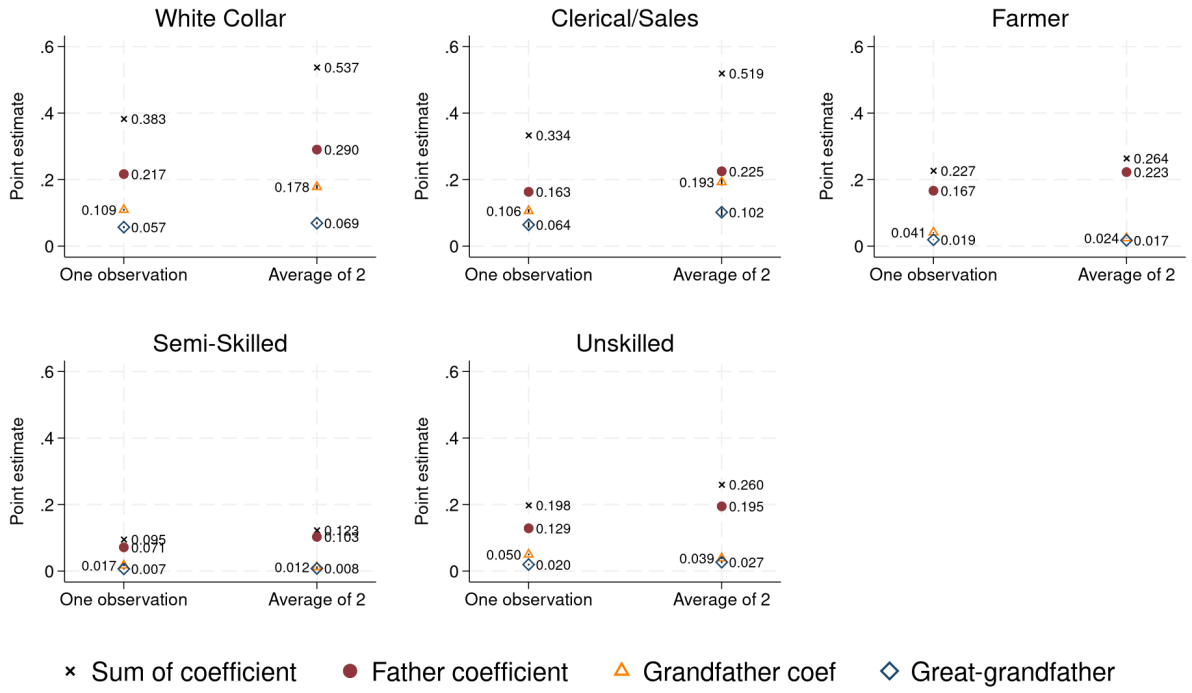
Notes: Data are from a linked sample of great-grandchildren to their great-grandfathers (see Figure 1) (Price et al., 2023; Ruggles et al., 2024). The samples are weighted to be representative on observables. The status measure is the adjusted Song score, which ranks occupation cells by their human capital (Song et al., 2020). The one observation column uses a single G1 outcome, a single G2 outcome, and a single G3 outcome. The mean of two observations uses the mean of the two G1, G2, and G3 outcomes. The 2SLS instrumnted the G1, G2 and G3 outcome with the second observations.

Figure A6: Predicted 4-generational coefficients using predicted income from 1940 Census



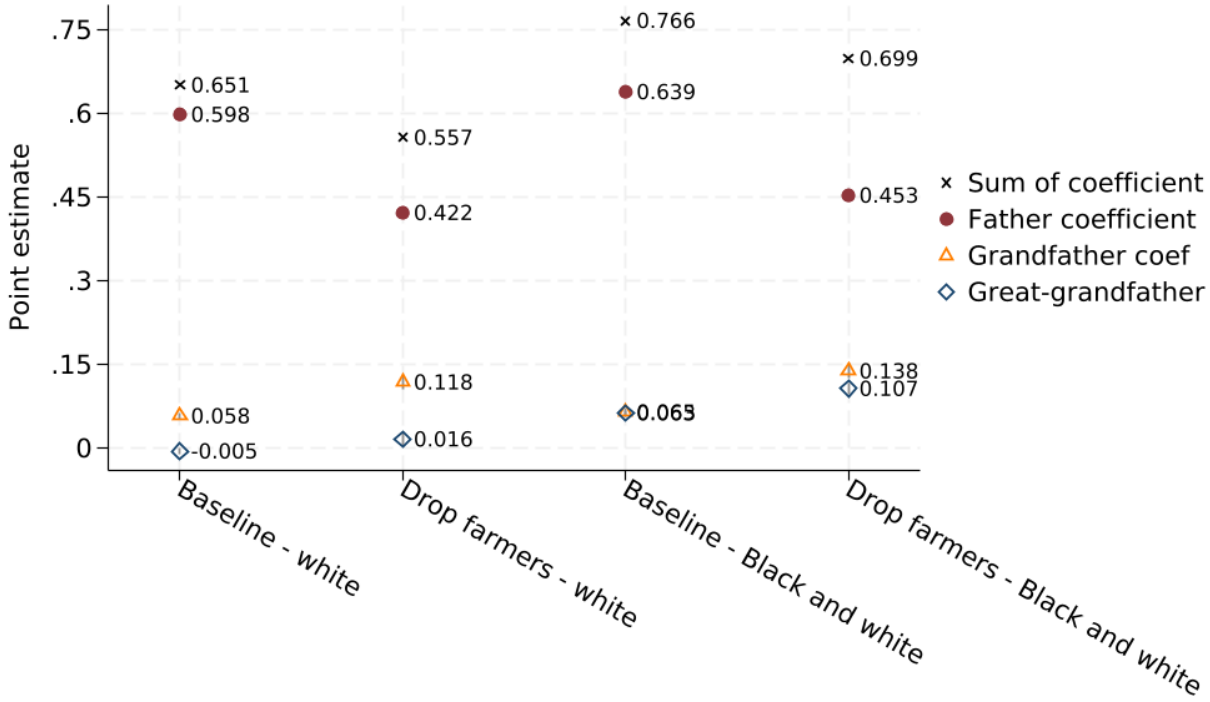
Notes: Data come from the 1850–1940 Censuses (Ruggles et al., 2024), linked using (Price et al., 2023). The samples are weighted to be representative on observables. These mimic Figure A5 but also show the predictions when using predicted income from the 1940 Census, similar to (Collins and Wanamaker, 2022).

Figure A7: An AR(3) model of multigenerational mobility for white families using occupational categories



Notes: Data are from a linked sample of great-grandchildren to their great-grandfathers (see Figure 1) (Price et al., 2023; Ruggles et al., 2024). The samples are weighted to be representative on observables. White collars are occupation codes *OCC*1950 that start with 0 or 2; Clerical sales start with 3 or 4; Farmer start with 1; Semi-skilled start with 5; Unskilled start with 6, 7, 8 or 9 (below 970).

Figure A8: Multigenerational mobility when dropping farmers



Notes: Data are from the 1850-1940 US censuses (Ruggles et al., 2024), linked by the Census Tree (Price et al., 2023). The samples are weighted to be representative on observables. This figure shows the AR(3) estimates (Equation 3) for the entire sample and when dropping farmer fathers, grandfathers and great-grandfathers. The point is that the father effect drops, while the grandparent and great-grandparent effects increase.

Table A1: Representativeness of 4-gen Linked Sample in 1930

		Linked Sample		
	Full	Unweighted	Weighted	Full=Weight
Age	33.018	35.099	33.115	0.000
Male	0.494	0.535	0.515	0.000
White	0.883	0.972	0.886	0.026
Black	0.117	0.028	0.114	0.026
Literate	0.970	0.985	0.972	0.006
Unskilled	0.237	0.217	0.263	0.000
Farmer	0.077	0.124	0.079	0.000
Skilled	0.084	0.093	0.091	0.000
Urban	0.588	0.442	0.595	0.000
Ever married	0.730	0.794	0.721	0.000
Head of household	0.376	0.430	0.380	0.000
Northeast	0.246	0.174	0.237	0.000
Midwest	0.329	0.352	0.330	0.370
South	0.328	0.373	0.331	0.002
West	0.097	0.101	0.102	0.000
Owns home	0.302	0.299	0.303	0.372
Sample size	27,687,562	331,694	331,694	

Notes: Data are from the 1930 Census (Ruggles et al., 2024), linked by Price et al. (2023). The columns are split by those linked four-generations back, as far as 1850, and the cross-sectional 1930 Census. The "full" column is the full-count census, the "unweighted" sample are those successfully linked back four generations, and the "weighted" applies inverse probability weights. The last column shows p-values for tests of the null hypothesis that the means for the full and weighted samples are the same.

Table A2: Representativeness of 4-gen Linked Sample in 1940

		Linked Sample		
	Full	Unweighted	Weighted	Full=Weight
Age	36.960	33.540	35.235	0.000
Male	0.496	0.579	0.516	0.000
White	0.895	0.985	0.895	0.467
Black	0.105	0.015	0.105	0.467
Educ. (years)	9.347	10.232	9.615	0.000
White Collar	0.208	0.231	0.223	0.000
Unskilled	0.270	0.280	0.290	0.000
Farmer	0.064	0.097	0.066	0.000
Skilled	0.077	0.089	0.081	0.000
Urban	0.578	0.430	0.569	0.000
Ever married	0.731	0.804	0.755	0.000
Head of household	0.409	0.452	0.406	0.000
Northeast	0.253	0.160	0.242	0.000
Midwest	0.314	0.347	0.313	0.092
South	0.327	0.377	0.336	0.000
West	0.106	0.116	0.109	0.000
Sample size	46,330,226	2,216,072	2,216,072	

Notes: Data are from the 1940 Census (Ruggles et al., 2024), linked by Price et al. (2023). The columns are split by those linked four-generations back, as far as 1850, and the cross-sectional 1940 Census. The "full" column is the full-count census, the "unweighted" sample are those successfully linked back four generations, and the "weighted" applies inverse probability weights.

Table A3: Using the surname method for the white population

Regression	OLS	OLS	IV	IV	Ratio
Sample	Rare surname	Rare surname	Rare surname	All	Rare surname
G1 surname average	0.298				
	(0.0109)				
G1 actual		0.212	0.314	0.331	
		(0.00525)	(0.00777)	(0.00131)	
$\widehat{\lambda}$					0.748
					(0.0195)
Observations	66,995	66,995	66,995	2,505,954	66,995
R-squared	0.016	0.036	0.027	0.034	0.023

Notes: Data are from the 1850-1940 US censuses (Ruggles et al., 2024), linked by the Census Tree (Price et al., 2023). See Appendix D for details on how the estimates are constructed. Column 1 regresses the G4 Adjusted Song Score on the average score within the same rare surname. The sample includes only surnames with fewer than 10 people ("rare" surnames). Column 2 directly regresses the G4 Adjusted Song Score on the G1 Adjusted Song Score using OLS. Column 3 uses an IV approach for the G1 score (repeated measures), again restricting to rare surnames. Column 4 expands the sample to include all surnames. Column 5 instruments the father's G1 score using the great-grandfather's score. The samples are weighted to be representative on observables.

Table A4: Underlying transmission estimates from latent model

	OLS	IV	Ratio	N
	$\lambda(\frac{\sigma_x^2}{\sigma_y^2}) + \rho(\frac{\sigma_{\eta_{\text{group}}}^2}{\sigma_y^2})$	$\lambda \frac{\sigma_x^2}{\sigma_x^2+\sigma_\eta^2+\sigma_\omega^2} + \rho \frac{\sigma_\eta^2}{\sigma_x^2+\sigma_\eta^2+\sigma_\omega^2}$	$\frac{\lambda^3\sigma_x^2+\rho^3\sigma_{\eta_{\text{group}}}^2}{\lambda^2\sigma_x^2+\rho^2\sigma_{\eta_{\text{group}}}^2}$	
Baseline sample	0.518	0.722	0.826	2547730
One required 1 obs. per ancestor	0.529	0.722	0.837	4197192
Non-agriculutural families	0.431	0.618	0.827	318092

Notes: Data come from the 1940 Census (Ruggles et al., 2024) and (Buckles et al., 2024*a*). Each cell in the first three columns represents a different regression. The OLS estimate comes from a regression of G4's status on a snapshot of G3's status. The IV column instruments one observation of G3's status with a second observation. The Ratio column represents the ratio of the G4-G1 slope to the G3-G1 slope, estimated using IV, where G3's status is instrumented with G1's status. The transitory column estimates the share of total variation in the father's outcome due to transitory error, while the permanent column estimates the share due to permanent error. N represents the number of observations. The baseline sample requires at least two observations for G3, G2, and G1. The expanded sample relaxes this requirement, allowing only one observation for G2 and G1 while still requiring two for G3 (for the IV estimation). The non-agricultural families sample excludes families in which any G1, G2, or G3 member is listed as a farmer. The samples are weighted to be representative on observables.

A. Data construction process

We build the data in the following steps.[1] First, we "chain" all 10-, 20-, 30-, and 40-year links from the Census Tree together in order to track an individual over their entire lifespan. This is an *intra*-generational dataset. (See the gray boxes in Figure 1.) The purpose of this within-generational dataset is twofold: first, to track an individual from their childhood home to their adult outcome. Second, we will use this dataset to retrieve multiple occupational observations of parents, grandparents, and great-grandparents, which we will use to address measurement error in an instrumental variables strategy.

Second, we create an "intergenerational" dataset by simply merging in parental information when a person in the "intragenerational" dataset is in their childhood home in the 1850 to 1920 Censuses. (See orange boxes in Figure 1). Due to concern of selection into leaving the household at older ages, we keep intergenerational links where a child is between ages 0 and 14 when in the household (Xie and Killewald, 2013). After we attach a parent to a child observation, we then check whether the parent is the intragenerational dataset such that we can track the parent over his or her own lifetime. There are two reasons why we do this. First is to obtain multiple occupation observations of the parent in order to address measurement error. Second is that it allows us to track the childhood home of the G2 parents, allowing us to build a multigenerational dataset, which we describe in the next step.

The third step is to take this intergenerational dataset and make it multigenerational. We do this by taking fathers and mothers in the intergenerational dataset, and then checking whether they are also children in a different row of the same dataset. For example, "Zachary Ward" may be observed as a father of "Elliot Ward" in one row, but then "Zachary Ward" may be observed as a child to "Jeffrey Ward" in a separate row. For this group, we can create three-generational links between grandparents and grandchild. Recall that the Census Tree includes women, and so for some links we can observe all four grandparents. This only occurs if both the husband and wife are successfully linked from childhood and adulthood.

Finally, we simply iterate this process again to obtain four-generational links. Our analysis will primarily rely on both 3-generation and 4-generational samples. Note that one only ends up in this sample if their grandparents or great-grandparents were in the United States. The descendants of immigrants will not be included as we cannot link outside of the United States. This will likely bias

[1] See description of chaining the sample in the Appendix of Kalsi and Ward (2025) for the more information.

us toward finding less mobility, as the descendants of immigrants were often more upwardly mobile than their counterparts who descended from US-born parents (Abramitzky et al., 2021).

B. Details on Sample Construction and Weighting Procedure

Our sample effectively represents individuals from 1930 and 1940 for whom we can directly observe parents', grandparents', and great-grandparents' outcomes, rather than relying on recollection. However, this creates sample selection concerns since successful linkage across four generations depends on factors like name uniqueness, birthplace, and age (Bailey et al., 2020).

To improve representativeness, we implement a comprehensive weighting procedure. We pool our linked sample of fourth-generation individuals with the broader census population (either 1930 or 1940, depending on when the fourth generation is observed). We then estimate a probit model predicting successful four-generation linkage based on:

- Age
- Education (in 1940, categorized as: 0-3 years; 4-6 years; 7-8 years; 9-11 years; 12 years; and more than 12 years) or literacy (in 1850-1930)
- Race (binary indicator for being white)
- Marital status (based on having a spouse in the same household)
- Household relationship
- Urban status (using IPUMS-USA variable URBAN, indicating residence in incorporated areas with more than 2,500 people)
- Gender
- Region (defined by nine Census divisions)

We calculate weights as $\frac{1-p}{p} * \frac{q}{1-q}$, where p is the predicted probability from the probit model, and q is the share of linked records (Bailey, Cole and Massey, 2020). To minimize the influence of extreme values, particularly for Black individuals who are less likely to be traced back, we Winsorize weights at the top and bottom 1 percent. However, this adjustment reduces the weighted Black share below the population share. Therefore, we make a final adjustment by multiplying weights to achieve an accurate share of Black individuals, resulting in Black individuals receiving approximately five times the weight of white individuals.

Tables A1 and A2 demonstrate the effectiveness of our weighting procedure by comparing characteristics of our linked sample to the general population in 1930 and 1940. Without weights, our sample overrepresents individuals who are white, live in rural areas, and reside in the South and Midwest. In 1940, our unweighted linked sample has 0.9 more years of education on average. These differences may partly reflect historical migration patterns rather than linking bias, as individuals need at least one great-grandparent who lived in the United States to be included in our sample. Applying weights substantially reduces all these observable differences, though we acknowledge potential remaining concerns about unobservable factors. As a robustness check, we show that using a sample with fewer required links produces similar four-generation results (see "Expanded Sample" in Table 5).

C. Assortative mating

To understand the importance of assortative mating in a latent variable model, following Stuhler (2024), suppose the latent transmission from the father (one parent) is a misspecified model where true transmission occurs from both parents. Let a child's latent ability be a weighted average of the parents':

$$x_{ig} = \tilde{\lambda}\left(\frac{x^{f}_{i,g-1} + x^{m}_{i,g-1}}{2}\right) + v_{ig} \tag{D1}$$

This equation assumes that the father and mother contribute equally to the child's status. If mother's matter more, then one may expect that transmission from maternal ancestors is stronger than with paternal ancestors. However, we find that they are similar in magnitude (Table 4).

Based on this model, the slope coefficient β_{-1} across two generations based on the *father's* observed outcome $y_{i,g-1,s}$ is

$$\frac{\text{Cov}(y_{i,g,t}, y^{f}_{i,g-1,s})}{\text{Var}(y^{f}_{i,g-1,s})} = \theta_{g-1}\tilde{\lambda}\left(\frac{1+m}{2}\right) \tag{D2}$$

where θ_{g-1} is attenuation bias from measurement error in the father's generation, and m is the slope coefficient from regressing the mother's latent factor on the father's latent factor. This parameter m captures the degree of assortative mating. .

When estimating transmission across three generations, the slope coefficient is:

$$\frac{\text{Cov}(y_{igt}, y_{i,g-3,s})}{\text{Var}(y_{i,g-3,s})} = \theta_{g-3}\tilde{\lambda}^{3}\left(\frac{1+m}{2}\right)^{3} \tag{D3}$$

This equation shows that when assortative mating is weak, outcomes fade quickly across generations. For example, if the assortative mating coefficient is 0.5, status gaps should shrink fast. In that case, the impact of assortative mating would be $\left(\frac{1+0.5}{2}\right)^{3} = 0.42$. Combining this with a high transmission estimate of 0.75 gives a four-generation association of 0.22 (assuming no measurement error). If assortative mating is stronger, say, 0.85, and the transmission coefficient is 0.75, then four-generation slope would be about 0.34, which is close to what we find for white families.

D. Surname estimates

We can compare our estimates to the method where surnames are used to estimate mobility. Clark (2014) argues that one can estimate multigenerational persistence using a grouping estimator based on rare surnames. Essentially, instead of using the great-grandfather's status measure on the right-hand side, you replace it with the average status of those with the same surnames. This grouping estimator has been strongly criticized by several scholars, as the surname estimator may capture some group-level persistence (γ) that differs from the parent-child transmission (λ) (Chetty et al., 2014; Torche and Corvalan, 2018; Vosters and Nybom, 2017). Indeed, the surname estimator has unique statistical properties that make it difficult to compare to directly linked data, including the possibility that surnames have informational content beyond the ancestor's status that positively biases transmission estimates (Santavirta and Stuhler, 2024).

Because we have names in our data, we can estimate the association between the first and fourth generations where, instead of using the direct observation of the grandparent's status, we use the average of surnames in the dataset. Because the data is directly linked, not surname-linked, there is 100 percent overlap between children's and parents' outcomes, an issue that attenuates surname estimates when comparing cross-sections generations apart (Santavirta and Stuhler, 2024). We calculate the surname average only for rare surnames (those with fewer than 10 observations per surname), following the methodology of Gregory Clark. Because others have raised concerns over group effects, and in particular, racial group effects (Chetty et al., 2014), we limit the sample to white families.

Specifically, we take out four-generational linked sample and then merge in last names for each generation using the restricted-access names from IPUMS-USA (Ruggles et al., 2024). We then clean the surnames using the NYSIIS algorithm to address potential phonetic mistakes when enumerators wrote names. We then take the average of all great-grandfathers with the same surname.[2] We then count how many unique great-grandfathers we have int he first generation, limiting the sample of rare surnames to those with fewer than 10 observations.

Our surname-based regression is then:

$$y_{i,\text{surname},G4} = \beta_0 + \beta_1 \overline{y}_{\text{surname},G1} + \varepsilon_{i,G4} \tag{D4}$$

[2]Our data structure observes the same grandfather multiple times because we can trace out to second cousins. Although the same great-grandfather is in different "rows" of the dataset based on the G4 observation, when taking the average, we use only one observation per great-grandfather.

where we take the G4 outcome and regress is on the surname-average outcome of the first generation. We only do this for G4 and G1 with the same (NYSIIS-cleaned) surnames, as they otherwise would not be matched in Clark's method.

The results are in Table A3. The first column shows that the surname average in the first generation has a 0.298 predicted slope for the great-grandchild's outcome. If we assume this is the direct transmission parameter to the third power (λ^3), this backs out to a 0.668 slope. This number is surprisingly close to the estimate using the ratio method in Table 5, where it was 0.684. However, there is a difference across these samples, as the rare-surname sample is limited to individuals with fewer than 10 occurrences in the dataset, which amounts to only about 67,000 in our data. If we use the ratio method to estimate λ in this sample, it is 0.748 (see last column of Table A3), which is higher than the backed-out surname estimate. The magnitude of the surname estimate is also similar to the IV estimate between the first and fourth generations, when we instrument one great-grandfather observation with a second, which is 0.314. Overall, it does not appear that the surname method *overstates* persistence, at least for our sample, and it produces similar results.

E. Ratio estimate, or great-grandfather IV, with group effects

In this section, we demonstrate that the "ratio method" can produce biased estimates of the intergenerational transmission of the latent factor. This method, which uses great-grandfather's outcome as an instrument for the father's outcome, may incorrectly account for both within-group mobility patterns and between-group differences. This occurs because the ratio method incorrectly combines and weights the separate transmission mechanisms that operate within groups and between groups.

The ratio estimate can be rewritten as the G4-G1 slope over the G3-G1 slope. After cancelling out the variance of the G1 coefficient, this can be written as:

$$\beta_{-1}^{ratio} = \frac{\beta_{-3}}{\beta_{-2}} = \frac{\lambda^3 \sigma_{x,G1}^2 + \gamma^3 \sigma_{\eta_{\text{group, G1}}}^2}{\lambda^2 \sigma_{x,G1}^2 + \gamma^2 \sigma_{\eta_{\text{group,G1}}}^2}. \tag{D5}$$

This ratio may either overstate or understate the "true" weighted transmission across generations, depending on the relative magnitudes of λ and γ. If within-group transmission is weaker than between-group transmission ($\lambda < \gamma$)—which was almost certainly the case for Black-white differences before World War II—then the ratio is below the true weighted average. At the same time, if the variance of the group effect shrinks over time (e.g., as Black-white distributions converge), then the ratio estimate could overstate the true transmission because the numerator over-weights group effect transmission γ. Either way, since the ratio method relies on data spanning multiple generations to estimate transmission across two, it cannot be used to test for excess persistence.

To see this more clearly, one can rewrite the ratio estimate as a weighted average of the within-group transmission λ and the between-group transmission γ. If we factor out $\lambda^2\sigma_x^2 + \gamma^2\sigma_\eta^2$ in both numerator and denominator (variances for the G1 generation), then this yields

$$\beta_{-1}^{ratio} = \lambda \frac{\lambda^2 \sigma_x^2}{\lambda^2 \sigma_x^2 + \gamma^2 \sigma_\eta^2} + \gamma \frac{\gamma^2 \sigma_\eta^2}{\lambda^2 \sigma_x^2 + \gamma^2 \sigma_\eta^2}.$$

If we define the "weight" to be

$$w = \frac{\lambda^2 \sigma_x^2}{\lambda^2 \sigma_x^2 + \gamma^2 \sigma_\eta^2},$$

then the ratio estimate is a weighted average of the within-group and between-group effects

$$\beta_{-1}^{ratio} = w\lambda + (1 - w)\gamma.$$

However, the true weighted average is

$$\beta_{-1}^{*} = \lambda \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\eta^2} + \gamma \frac{\sigma_\eta^2}{\sigma_x^2 + \sigma_\eta^2}.$$

In this expression, the weight on λ is $\sigma_x^2/(\sigma_x^2 + \sigma_\eta^2)$, but in the ratio formula, the weight is $\lambda^2\sigma_x^2/(\lambda^2\sigma_x^2 + \gamma^2\sigma_\eta^2)$. If $\lambda < \gamma$, the ratio method puts too much emphasis on γ.

If we define the ratio $r = \frac{\lambda}{\gamma}$, then we can rewrite the weight on the within-group effect as

$$w = \frac{r^2\sigma_x^2}{r^2\sigma_x^2 + \sigma_\eta^2}$$

Thus, if $r < 1$ (i.e. $\lambda < \gamma$), then $r^2 < 1$, and

$$w = \frac{r^2\sigma_x^2}{r^2\sigma_x^2 + \sigma_\eta^2} < \frac{\sigma_x^2}{\sigma_x^2 + \sigma_\eta^2}.$$

In that case, the ratio estimate places more weight on the group-effect persistence γ than on the direct transmission λ. Since γ is larger, the ratio estimate can exceed the true value.

To see this in a numerical example, Let $\lambda = 0.70$, $\gamma = 0.95$, $\sigma_x^2 = 0.75$, and $\sigma_\eta^2 = 0.25$. Compute the true weighted average:

$$\beta_{-1}^{*} = 0.70 \times \frac{0.75}{0.75 + 0.25} + 0.95 \times \frac{0.25}{0.75 + 0.25} = 0.762.$$

The ratio estimate is:

$$\beta_{-1}^{ratio} = \frac{0.70^3 \times 0.75 + 0.95^3 \times 0.25}{0.70^2 \times 0.75 + 0.95^2 \times 0.25} = 0.80.$$

which is larger than 0.763. Because γ is bigger than λ, the ratio method places more weight on group persistence and overshoots the true weighted average.