Measuring multigenerational mobility
What can we learn from cousin correlations?

Ian Lundberg*

Department of Sociology and Office of Population Research

Princeton University

($\approx$ 8,609 words in main text)

Last updated: 30 March 2018

Contents

---

*Draft prepared for presentation at the 2018 Annual Meeting of the Population Association of America. I thank Dalton Conley, Brandon Stewart, Sara McLanahan, Michael Hout, and students in the Joint Degree Program in Social Policy at Princeton University for many useful comments on earlier versions. All errors are my own. Direct all correspondence to Ian Lundberg, 227 Wallace Hall, Princeton, NJ 08544; ilundberg@princeton.edu; 909-561-0531. The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and the National Science Foundation under award number 1157698.

Abstract

For those concerned about economic opportunity, the pattern of social mobility over many generations is a key subject of inquiry. Although this field has long focused on correlations between the attainment of parents and children, this approach may understate the long-run persistence of status if the information about life chances contained in extended families is ignored (Mare, 2011). Recent studies have used a creative empirical lever to assess this possibility: if cousins attain similar socioeconomic outcomes, this suggests that familial advantages persist to a measurable degree for at least three generations. This paper pushes this line of research forward by making explicit the definitions of family background and assumptions required for inferences about a long-run multigenerational process from sibling and cousin correlations. Given these assumptions, I provide researchers tools to translate sibling and cousin correlations, which are difficult to interpret, into comprehensible summaries of the implied persistence of status attainment over many generations. Through new empirical results and a reassessment of published work, I conclude that long-run mobility in the U.S. is reasonably well-approximated by models assuming transmission from parents to children.

(182 words)

1 Introduction

The extent to which social origins are associated with life chances has long played a central role in the sociology of inequality and opportunity. Scholars have typically emphasized correlations between the attainment of parents and children (Blau and Duncan, 1967; Sewell et al., 1969; Hauser and Featherman, 1977; Hout, 1988; Hout and DiPrete, 2006), yet this approach risks missing important associations between life chances and other family background characteristics beyond those of parents, such as the attainment of extended kin. In his presidential address to the Population Association of America, Mare (2011) challenged this literature to take seriously the question of how attainment unfolds over more than two generations. At the core of the argument is a concern about the dynastic persistence of status attainment: if measured parent characteristics only partially capture the full set of family background characteristics that are associated with life chances, then simple correlations between the attainment of parents and children may substantially understate the persistence of socioeconomic status over many generations. Motivated by this concern and empowered by new data sources (Song and Campbell, 2017), a virtual explosion of research has begun analyzing patterns of mobility over multiple generations (Jæger, 2012; Chan and Boliver, 2013; Hertel and Groh-Samberg, 2014; Pfeffer, 2014; Wightman and Danziger, 2014; Huang et al., 2015; Lindahl et al., 2015; Pfeffer and Killewald, 2015; Knigge, 2016; Olivetti et al., 2016; Ziefle, 2016; Hällsten and Pfeffer, 2017; Pfeffer and Killewald, 2017).

These studies share a common style: authors posit a null model, in which the socioeconomic attainment of children is independent of the attainment of extended kin conditional on parent attainment, and assess the evidence against that null model. One promising way to conduct this test is to compare the status attainment of siblings and cousins. If siblings' (cousins') attainment is similar, this suggests that factors common to siblings (cousins) are strongly correlated with life chances. Because siblings share their nuclear and extended family while cousins share only their extended family, a comparison of sibling and cousin correlations provides a useful test of how attainment may be related to factors in the extended family, without requiring the researcher to measure the relevant factors (Jæger, 2012; Hällsten, 2014; Knigge, 2016; Pfeffer et al., 2016).

Although the benefits of sibling and cousin correlations are clear from the perspective of measurement, these approaches face a serious drawback: the resulting estimates are difficult to interpret and their connection to long-run mobility processes is non-obvious. The goal of this paper is to summarize what we learn about multigenerational processes from observed sibling and cousin correlations. Learning anything requires a specific definition of family background as well as several assumptions about the order of the process, the sign of associations, and the reliability with which the outcome is measured. For a researcher willing to defend these assumptions, I provide statistical tools[1] to translate sibling and cousin correlations into interpretable estimates of the long-run mobility of families. I conduct new analysis of data from the Panel Study of Income Dynamics (PSID) and reassess published evidence. Results challenge the substantive conclusions of much existing research and suggest that multigenerational attainment processes in the U.S. are reasonably well-approximated by the classical first-order Markov model in which advantages are transmitted one generation at a time.

2 Multigenerational attainment processes

Classical status attainment models frequently relied on data covering only two generations and therefore limited the study of family background to the correlation between parent and child outcomes (Blau and Duncan, 1967; Sewell et al., 1969). For this to represent the total association of life chances with family background, one would need to assume a first-order Markov process in which status attainment of offspring (Y_t) is conditionally independent of all prior generations $(\dots, Y_{t-3}, Y_{t-2})$ conditional on parent attainment (Y_{t-1}).

Definition 1. In a *first-order Markov process*, attainment in generation t is independent of attainment in all prior generations conditional on attainment in generation $t-1$. Formally, $Y_t \perp\!\!\!\perp \{Y_{t-2}, Y_{t-3}, \dots\} \mid Y_{t-1}$.

The solid black lines in Figure 1 represent this first-order Markov process. Under this

[1]All software used in this paper will be open-sourced upon publication for use by future researchers.

process, β_1 represents the intergenerational correlation in Y, and β_1^k captures the correlation of the attainment of ancestors and descendants separated by k generations. Because $\beta_1 < 1$, the process decays geometrically toward zero, effectively wiping out the benefit of having high-status ancestors in only a few generations. As Becker and Tomes (1986, p. S28) concluded from this assumed model, "practically all the advantages or disadvantages of ancestors tend to disappear in only three generations."

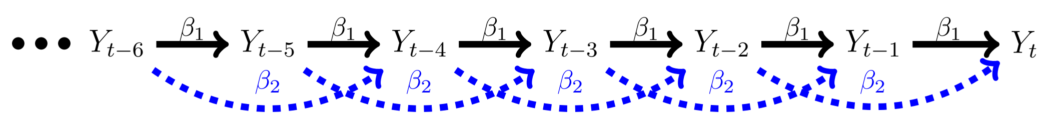

Fig. 1. First-order (solid black) and second-order (all lines) Markov transmission processes

This conclusion is not warranted, however, if the true mobility process involves a higher-order Markov process, such as the second-order process represented by the solid black and dashed blue lines in Fig. 1. Under this model, the correlation of offspring attaiment with attainment in prior generations is more complicated: the correlation between a child and their parent is $\frac{\beta_1}{1-\beta_2}$, and the correlation between a child and their grandparent is $\frac{\beta_1^2}{1-\beta_2} + \beta_2$ (see Appendix A).

Definition 2. In a *second-order Markov process*, attainment in generation t is independent of attainment in all prior generations conditional on attainment in generation $t-1$ and $t-2$. Formally, $Y_t \perp\!\!\!\perp \{Y_{t-3}, Y_{t-4}, \dots\} \mid \{Y_{t-1}, Y_{t-2}\}$.

A brief example clarifies that the distinction between these processes is not a mathematical curiosity but a fundamental research question with implications for the long-term well-being of families. Suppose we observe that parents' incomes and the incomes of their children are correlated 0.5. Two scenarios (among many) that could produce this intergenerational correlation are a first-order Markov process with $\beta_1 = 0.5$ or a second-order Markov process with $\beta_1 = 0.4$ and $\beta_2 = 0.2$. The distinction between these two processes becomes most pronounced when we extrapolate to the implied correlation between attainment in generation t and attainment many generations later. In this scenario, the correlation between income attained by those separated by five generations would be $0.5^5 \approx 0.03$ under the first-order model, compared with 0.13 under the second-order model (see Appendix A for derivation of these values). In fact, it would take a gap of nine generations before the second-order process would produce a multigenerational correlation below 0.03, the value achieved by the first-order process in only five generations. As this example shows, the distinction between a first- or second-order Markov process has implications for the dynastic quality of multigenerational mobility (see Fig. 2).

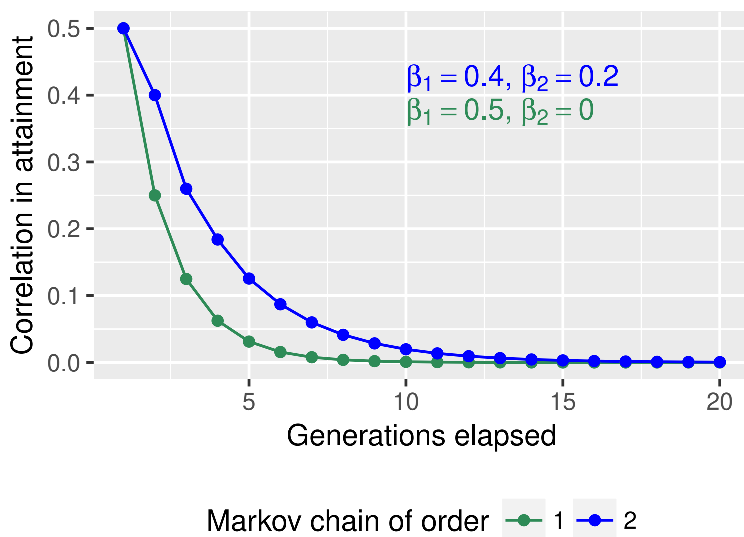

Fig. 2. Multigenerational correlations in socioeconomic attainment under two regimes which both produce a parent-child correlation of 0.5. Socioeconomic status exhibits more long-run persistence if transmission follows a Markov chain of order 2.

Although a second-order parameter β_2 may substantially increase the dynastic persistence of socioeconomic status over many generations, an early evaluation of the two models found no substantial association between grandparent status and grandchildren's education or occupational attainment net of parent characteristics in a sample of Wisconsin high school graduates from the class of 1957 (Warren and Hauser, 1997). In other words, β_2 in Fig. 1 was approximately zero. But might it be non-zero in other contexts?

2.1 Theories of complex multigenerational associations

Numerous theories suggest that extended families might play critical roles in children's lives through pathways including resource compensation, wealth transmission, role modeling, or direct involvement in the socialization process. Under each of these theories, a second-order Markov transmission process is plausible. While I will discuss these theories from the perspective of grandparents for concreteness, they represent associations that could arise from any family member shared by cousins, such as an aunt or uncle.

Theories of resource compensation suggest that extended kin may provide material resources when parents fall short (Jæger, 2012; Ziefle, 2016). As a concrete example of the resource compensation perspective, many families formed multigenerational households during the Great Recession (Mykyta and Macartney, 2011), often relying on extended family as a private safety net (Pilkauskas et al., 2014). Even before the Great Recession, an estimated 24% of American children born in 2001 lived with their grandparents at some point between birth and age 5 (Pilkauskas and Martinson, 2014). If the safety net provided by extended family members is more available to children whose grandparents have high levels of status attainment, this could produce a violation of the Markov assumption such that children of downwardly mobile parents experience upward counter-mobility (Chan and Boliver, 2013).

Wealth transmission may confer numerous advantages on offspring regardless of whether parents have been downwardly or upwardly mobile. Wealth typically must be accumulated over a lifetime, so that even well-off parents may not have the requisite wealth to confer benefits on their children. Those with wealthy grandparents may benefit from transfers to pay for college or contribute toward a down-payment on a home, for instance, or may be more inclined to marry given the financial security that wealthy kin can provide (Pfeffer and Killewald, 2017).

Cultural pathways may also link the outcomes of offspring with those of their grandparents if descendants look toward ancestors for a role model or a reference frame. The Swedish education system is highly egalitarian so that it is difficult to "purchase" achievement, yet Hällsten and Pfeffer (2017) find that grandparent wealth predicts the academic achievement of offspring net of parent controls. They argue that those descended from wealthy grandparents know college is possible and feel a normative push toward high academic achievement. In other words, grandparents may serve as a reference frame when their descendants make decisions relevant to mobility (Hertel and Groh-Samberg, 2014). The reference frame of grandparent attainment may also influence how parents perceive their own status; parents with similar socioeconomic circumstances may view their own class position through different cultural lenses if they arrived at that position via upward or downward mobility from the grandparent generation, potentially influencing their own childrearing practices (Wightman and Danziger, 2014).

Perhaps most directly, grandparents may be directly involved in the socialization of offspring in some settings. For instance, multigenerational coresidence is common in China (Zhang, 2004), providing a site in which interactions between grandparents and grandchildren may be numerous. Zeng and Xie (2014) find that the education of Chinese grandparents is associated with offspring educational attainment net of parents only when grandparents and grandchildren live together, providing opportunities for direct contact and socialization. In the American context, the role of direct socialization may be more relevant today than in the past. Given increased life expectancy and increasing rates of multigenerational coresidence (Taylor et al., 2010), more grandparents can know and interact with their grandchildren through a longer portion of development than they did in the past (Bengtson, 2001; Uhlenberg, 2004; Seltzer and Bianchi, 2013). Direct socialization may therefore represent an increasingly important cultural pathway even in societies like the U.S. in which multigenerational coresidence has not historically been common.

2.2 Linking intragenerational correlations to multigenerational processes

Despite the theoretical appeal of a component of life chances shaped by extended family members, this quantity is empirically difficult to estimate within an intergenerational framework because it is very difficult to net out all association of offspring attainment with parent characteristics in order to discover the residual association with grandparent characteristics. If grandparent attainment is associated with offspring attainment net of measured parent characteristics, this could represent a direct association or could represent an indirect association of grandparent and offspring attainment through unmeasured parent characteristics.

Sibling and cousin correlations may solve this problem by serving as an "omnibus" measure to capture the total influence of family background (Solon et al., 2000), sidestepping the need to measure characteristics in the parent and grandparent generations. Inference about the underlying multigenerational process, however, remains several steps removed from what the data can show: it requires a definition of family background which may or may not be useful as well as three assumptions about the order, sign, and measurement precision of the multigenerational process.

2.2.1 Defining family background

Sibling and cousin correlations are informative about the association between life chances and family of origin only under a very particular definition of family background.

Definition 3. For sibling correlations to be informative about intergenerational transmission, *family background* must be defined as all characteristics shared by siblings.

This definition accords with a quantity of theoretical interest insofar as the fluidity of society is related to the ability of individuals to move away from the attainment that would be expected given the characteristics they share with their siblings. In many theories of intergenerational processes, the relevant aspects of family background are in fact shared by siblings. Blau and Duncan's (1967, p. 170) basic model of stratification, for instance, included only father's education and occupation as the key variables from one's family background associated with one's own attainment. To the extent to which father's education and occupation are constant over the period in which children are raised, these factors are incorporated into sibling correlations. Alternatively, father's occupation may not be constant over time, especially in recent decades characterized by precarious work and reduced opportunities for stable, long-term careers (Kalleberg, 2009). Siblings born many years apart could be exposed to substantially different occupations held by the same father, yet these different exposures would not be considered part of "family background" by the definition above because they are not shared by siblings. Sibling correlations may therefore be informative about a quantity that omits aspects that we would theoretically consider central a meaningful definition of family background.

As an extreme example, Conley (2008) describes a world in which parents have tremendous influence over the life chances of two children, forcing the oldest child to sacrifice everything to support the youngest child. In this world, the oldest child in every family might attain no education and no income, producing a correlation of zero between the positive attainment of the youngest child and the constant lack of any attainment for the oldest child. By the definition above, the relationship between offspring attainment and family background would also be 0; siblings do not share any relevant inputs to the attainment process. Our intuition that family background is very influential in this case suggests a shortcoming of the definition of family background for which sibling correlations are informative.

Likewise, the definition of family background above includes factors that researchers may *not* wish to include in a theoretically-motivated definition of the concept. Suppose parents have no influence over the life chances of children, but all children follow closely in the footsteps of the eldest sibling. The first-born sibling goes to college, and subsequent children follow. The first-born chooses a particular profession, and subsequent children follow suit. Here, all siblings share in common the first sibling's attainment decisions, yet this factor would not fall under the umbrella that we would usually define as family background. In this case, it is possible that child outcomes would be completely independent of parent outcomes, yet the sibling correlation could be close to one.

To draw inferences about a second-order process from sibling and cousin correlations, one must adopt a similar definition of extended family background.

Definition 4. For cousin correlations to be informative about multigenerational transmission, *extended family background* must be defined as all characteristics shared by cousins.

The concerns that applied to sibling correlations carry over to cousin correlations. If grandparents worked full-time while early-born cousins were raised but retired and became deeply involved in the lives of later-born cousins, this would fall outside the scope of "extended family background" in the definition above and would be missed by inferences relying on cousin correlations. If all cousins follow the career trajectory set in place by the first-born cousin, this could produce sizable cousin correlations despite no influence of prior generations on these shared trajectories. As these examples demonstrate, the definition of family background required for multigenerational inference from intragenerational correlations, which requires counting as extended family background all variables shared by cousins, may be suboptimal.

Given the gap between the influence of family background and the quantity that is estimable by sibling correlations, prior research disagrees on the utility of sibling and cousin correlations for the study of family background. Some authors find the approach reasonable and write that a sibling correlation "primarily reflects the influence of. . . background factors" (Blau and Duncan, 1967, p. 317), captures the share of inequality that "may be attributed to family background," (Mazumder, 2008, p. 686), summarizes the "global impact of family background" [p. 109](Conley et al., 2007), or suggests that "unmeasured parent characteristics also exert considerable impact on son's earnings," (Corcoran et al., 1976, p. 435). Others have made limited claims such as that within-group similarities provide "an upper bound on the influence" of membership in the group under study (Solon et al., 2000, p. 383). Still others find the definition so absurd that sibling similarities "should not be interpreted, except tautologically, as reflecting the force of family influences," (Griliches, 1979, p. S59).

I take a middle position of treating sibling and cousin correlations as imperfect proxies for the total influence of family background. For the most part, parent characteristics may be constant over the course of child development; dramatic mid-life career changes, for instance, may be the exception rather than the norm. It is possible that parents invest differentially in their children, but differential investment may be relatively minor compared with common benefits bestowed all children by well-resourced parents and grandparents. Siblings may influence each other, but perhaps not to the same degree as parents influence them jointly. If these assumptions approximately hold, then it may be reasonable to proceed to draw inferences about multigenerational processes from intragenerational correlations, albeit with caution.

2.2.2 Three key assumptions: Order, sign, and measurement

Accepting the definitions of family background above, the connection between sibling and cousin correlations and multigenerational processes still requires three key assumptions about the order of the process, the sign of associations, and the precision with which the outcome is measured.

Assumption 1. Researchers must assume the order of the multigenerational transmission process.

Any inference about a multigenerational process first requires an assumption about the order of that process: that offspring attainment is independent of attainment more than k generations in the past given attainment the preceding k generations. This assumption is typically constrained by the data at hand. Data including siblings but not cousins, for instance, are uninformative about the association of attainment with extended as opposed to nuclear family characteristics because siblings share both their nuclear and extended families. This makes it impossible to distinguish between a first and second-order model with these data. With data on siblings and cousins, one can assume a second-order model in which grandparents' attainment is associated with offspring attainment independent of parents. The second-order model contains the first-order model as a special case, so the distinction between these is estimable. One must, however, assume away a third- or higher-order model by assuming that great-grandparent attainment is independent of offspring attainment net of parents and grandparents. As with the definition of family background, it is unlikely that the multigenerational process perfectly corresponds to a first- or second-order model, but these might be reasonable approximations to the extent to which the attainment of great-grandparents and earlier kin are associated with offspring attainment net of the intervening generations in only negligible ways.

Assumption 2. Researchers must assume the values of structural parameters among several candidates that are equally consistent with the data. In a second-order Markov process, this involves assuming the sign of the first-order parameter and selecting among up to three possible values for the second-order parameter.

Suppose the multigenerational mobility process can be adequately summarized by a second-order Markov model with unknown parameters β_1 and β_2. Suppose also that income is reported with reliability η, which represents the correlation between true and reported income (see Assumption 3). Equations 1 and 2 state the sibling and cousin correlations implied by this model. Proofs are provided in the Appendix, and the discussion here is only intended to build intuition. If we randomized the incomes of parents and grandparents, each child's income would be correlated β_1 with parent income and β_2 with grandparent income. For siblings who share their parents and grandparents, the resulting sibling correlation would be $\beta_1^2 + \beta_2^2$. Because parent and grandparent incomes are not randomized, these two inputs are themselves correlated producing an additional adjustment term $2\frac{\beta_1^2\beta_2}{1-\beta_2}$. Finally, reporting error reduces the sibling correlation in reported attainment Y by the multiplicative factor η^2 (where $\eta < 1$), because each sibling provides a report that is only correlated η with the truth. In general, η is not estimable from the data (see Assumption 3).

$$\rho_{\text{Siblings}} = \text{Cor}\left(\tilde{Y}_{t,a}, \tilde{Y}_{t,b}\right) = \eta^2\left(\beta_1^2 + \beta_2^2 + 2\frac{\beta_1^2\beta_2}{1-\beta_2}\right) \tag{1}$$

Equation 2 for the cousin correlation is similar except that the β_1^2 term is multiplied by ρ_{Siblings} to account for the fact that cousins do not have the same parents but instead have parents whose attainment is correlated ρ_{Siblings} because they are siblings.

$$\rho_{\text{Cousins}} = \text{Cor}\left(\tilde{Y}_{t,aa}, \tilde{Y}_{t,bb}\right) = \eta^2\left(\beta_1^2\rho_{\text{Siblings}} + \beta_2^2 + 2\frac{\beta_1^2\beta_2}{1-\beta_2}\right) \tag{2}$$

Reversing the process, now suppose we observe the sibling and cousin correlations and want to infer the structural parameters β_1 and β_2 that could have produced those correlations under the model above. Equations 1 and 2 can be rearranged to yield the following formulas for the structural parameters (proofs are provided in Appendix C).

$$\beta_1 = \pm\sqrt{\frac{\frac{1}{\eta^2}\rho_{\text{Siblings}} - \frac{1}{\eta^2}\rho_{\text{Cousins}}}{1 - \frac{1}{\eta^2}\rho_{\text{Siblings}}}} \begin{array}{l} \leftarrow \text{Difference in sibling and cousin correlations in } Y \\ \leftarrow \text{Proportion of variance in } Y \text{ that is between families} \end{array} \tag{3}$$

$$0 = \beta_2^3 - \beta_2^2 - \left(\beta_1^2 + \frac{1}{\eta^2}\rho_{\text{Siblings}}\right)\beta_2 + \left(\frac{1}{\eta^2}\rho_{\text{Siblings}} - \beta_1^2\right) \tag{4}$$

Equations 3 and 4 are noteworthy in that the solution is not unique. First, an observed set of sibling and cousin correlations is completely uninformative about the sign of β_1; this term enters Eq. 1 and 2 only in squared form. Substantively, a positive sibling correlation could be generated by either a positive or a negative correlation β_1 between child attainment and a shock to parent attainment. One might reasonably assume that this correlation would be positive, as children reproduce the socioeconomic outcomes of their parents. This assumption, however, requires theory or additional data beyond the observed intragenerational correlation. The choice is non-obvious; compelling theories also exist to suggest that the intergenerational correlation is *negative*. For instance, children may systematically rebel against their parents, vowing to take a different life trajectory. The children of a CEO may achieve correlated outcomes because both are spoiled and end up in the working class. Likewise, the children of working-class parents may develop an especially strong work ethic and ascend the hierarchy to attain correlated outcomes as business executives. In the absence of additional analyses, each of these stories is equally consistent with the evidence provided by an observed sibling correlation.

Even after assuming a sign on β_1, an observed set of sibling and cousin correlations may be generated by up to three possible values of β_2: the roots of the polynomial in Eq. 4. To make the possibilities concrete, suppose we estimate a sibling correlation of 0.25 and a cousin correlation of 0.1. Figure 3 shows the possible values of β_1 and β_2 that could have generated these correlations for a given level of assumed measurement reliability η. Under perfect measurement ($\eta = 1$), the blue dot in the upper right quadrant indicates that the observed intragenerational correlations could have been generated by $\beta_1 = 0.45$ and $\beta_2 = 0.09$. The blue dot in the upper left quadrant indicates that the same correlation could have been generated by an equivalent *negative* correlation between child attainment and a shock to parent attainment. Under either scenario, the correlations could also have been generated by strong negative values of -.41 for β_2. Theory or additional analyses are required to distinguish these possibilities.

Assumption 3. Researchers must assume the precision with which the outcome is measured as well as how measurement error is correlated within families.

Measurement error has formed the core of numerous debates in stratification research. As early as the 1960s, it was known that interpretation of correlations in socioeconomic status depends on the "assumptions one is willing to defend" about the relationship between measured variables and true values (Siegel and Hodge, 1968). For instance, while Blau and Duncan (1967) found substantial returns to education net of family background variables, Bowles (1972, p. S222) called this conclusion "seriously misleading" due to reporting error in family background variables that downwardly biased their association with the outcome and produces an overstatement of the returns to education. Others noted that the debate becomes more complex if measurement error varies by race (Bielby et al., 1977). In the present-day debate about multigenerational mobility, Ferrie et al. (2016) argue that apparent correlations between grandparent and offspring incomes conditional on parent income appear solely because parent income is measured with error. Central conclusions in the field of mobility research often hinge on assumptions about measurement.

Sibling and cousin correlations are no exception to this rule. Although this approach avoids the pitfalls of measuring family background characteristics, results are sensitive to measurement error in the outcome variable. As first noted in the context of sibling correlations (Solon et al., 1991), intragenerational correlations in measured attainment $\tilde{Y}$ will be less than correlations in true attainment Y to the extent to which attainment is reported with classical measurement error. On the other hand, non-classical measurement error such that siblings and cousins tend to over- or under-report their incomes together may produce sibling correlations in measured attainment that exceed correlations in true attainment. In many cases, it may be reasonable to assume that measurement error is classical and uncorrelated within families. By assuming this structure, one can estimate the sibling and cousin correlations as a function of a range of possible values of measurement reliability η by which true and reported attainment are correlated. This is depicted graphically in Fig. 3 by the shading of the line, covering the range of η from 0.7 to 1.0, approximately the range of reliability for which many status attainment outcomes are measured (Angrist and Krueger, 1999, p. 1346). In cases where there is only one observation per person, η must be assumed because the data are completely uninformative about this quantity.[2]

With repeated observations, one can reduce this problem by estimating sibling and cousin correlations in permanent income, netting out all within-person variation. As Solon et al. (1991) show, sibling correlations are substantially higher with this strategy. This approach, however, requires two caveats. First, the estimand changes when one moves from annual income to permanent income; these quantities are theoretically distinct. Even with perfect measurement, true sibling correlations in permanent and annual income would be different. Second, permanent income solves measurement problems only if measurement errors are idiosyncratic. If a given respondent tends to over-report their income in all years, for instance, this would produce classical measurement error in permanent income. For this reason, assumptions about the reliability η in the annual income setting translate naturally to similar assumptions about η in the setting of permanent income. Because one might reasonably assume that η is close to one in the permanent income setting, the permanent income correlation approach developed by Solon et al. (1991) represents a reasonable solution if repeated observations are available.

2.2.3 Sensitivity of conclusions to assumptions

One might hope that the long-run persistence of attainment within families is robust to the assumptions made, but this is not the case. Fig. 4 plots the multigenerational correlation in attainment between an ancestor and their descendants separated by between one and ten generations,

[2] That the measurement reliability η is not observable in the data has a long history in social science theory; Northrop (1947) referred to this type of unobservable correlation as an "epistemic correlation" which connects what we observe in the world with an unobservable theoretical concept.

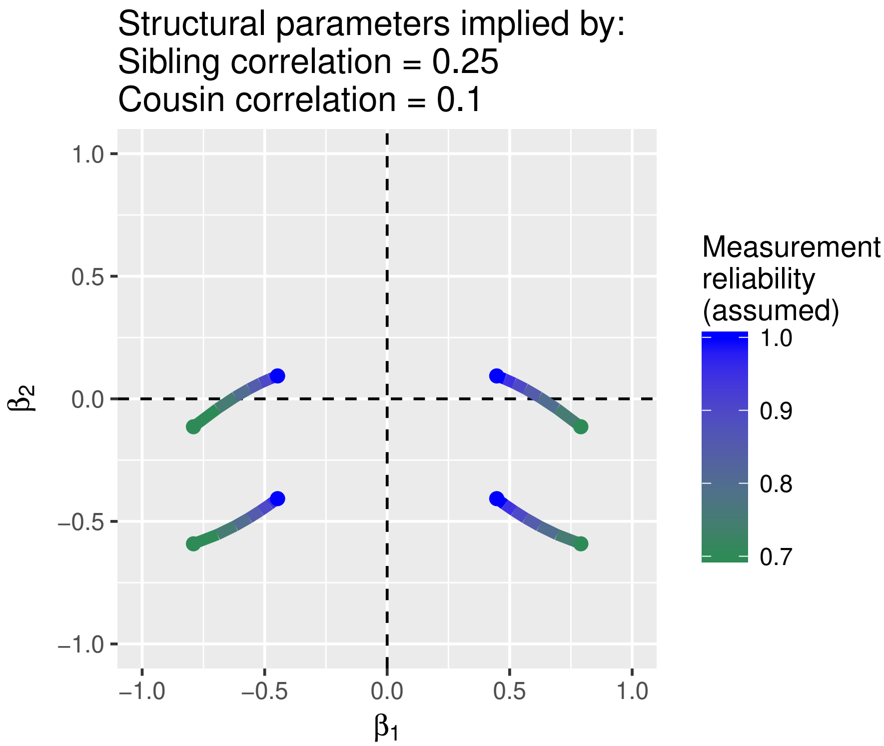

Fig. 3. Structural parameters implied by $\rho_{\text{Siblings}} = 0.25$ and $\rho_{\text{Cousins}} = 0.1$.

under the structural parameters presented in Fig. 3. Every line in the plot is equally consistent with the observed data under a certain set of assumptions. The upper-right quadrant again shows the familiar case for which $\beta_1 \geq 0$ and $\beta_2 \geq 0$: the multigenerational correlation decays as the number of generations increases. The other quadrants, however, show that multigenerational correlations may oscillate between positive and negative for certain combinations involving a negative value of β_1 or β_2. In all cases, the multigenerational correlations depend on the assumed measurement reliability η because this affects the structural parameters β_1 and β_2 implied by the observed set of sibling and cousin correlations.

To summarize, researchers who use intragenerational correlations to infer multigenerational processes must accept a very specific definition of family background and then assume (1) the order of the transmission process, (2) the sign of the underlying associations, and (3) the degree of measurement precision. Although these represent many decisions, theory may inform Assumptions 1 and 2, and prior literature can inform Assumption 3. The next section demonstrates how to reason about these decisions in a specific empirical example.

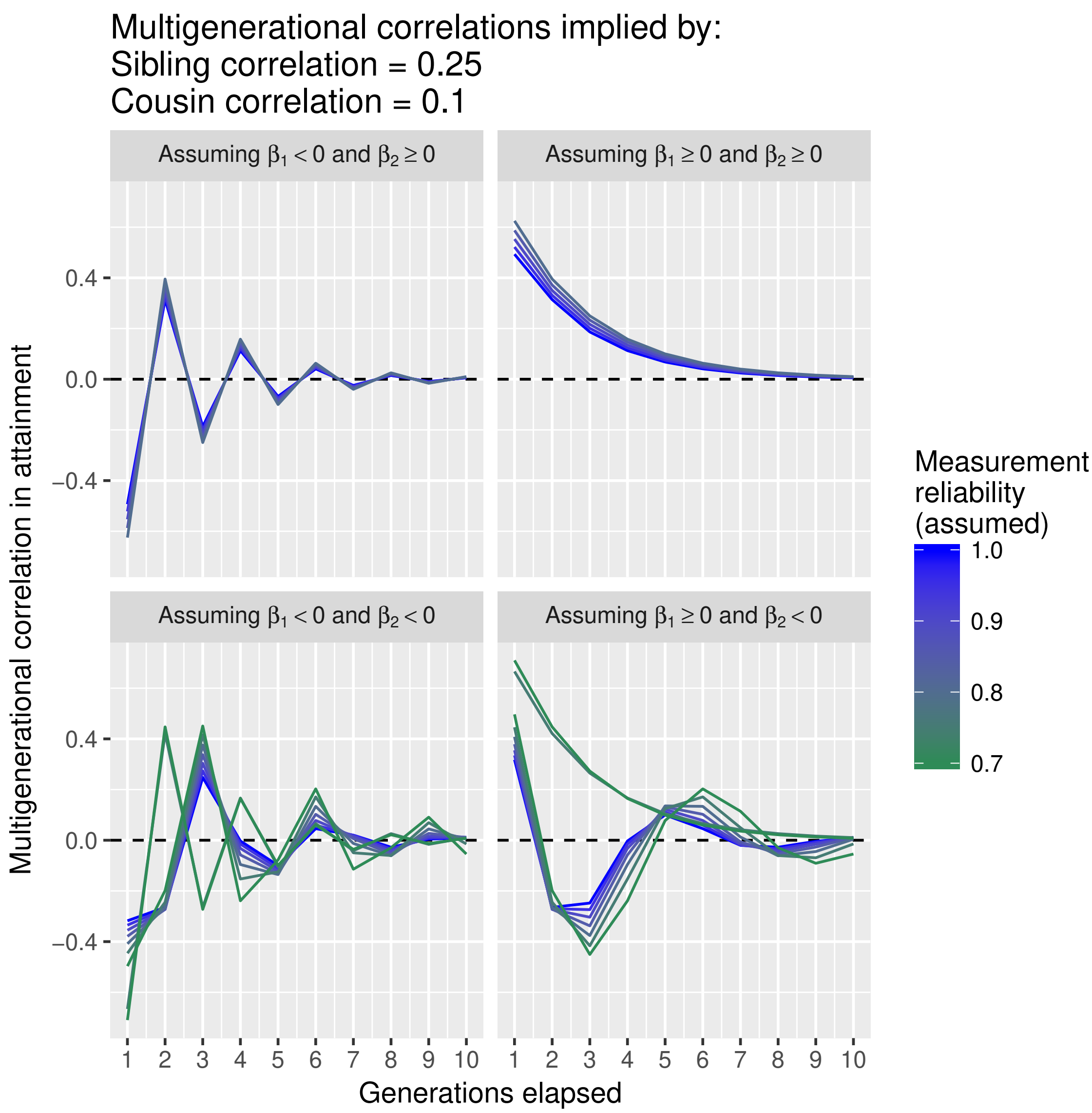

Fig. 4. Multigenerational correlations implied by $\rho_{\text{Siblings}} = 0.25$ and $\rho_{\text{Cousins}} = 0.1$.

3 Empirical example: Drawing inferences in the PSID

Applying the formal relationships derived above, this section reports sibling and cousin correlations in age-adjusted log family income in the U.S. and summarizes their implications for long-run multigenerational mobility processes.

3.1 Data

The Panel Study of Income Dynamics (PSID, 2017) selected a sample of 1968 households and interviewed descendants from those households annually through 1999 and biannually thereafter through 2015, producing a sample of individuals nested in family lineages. I restrict to the Survey Research Center (SRC) sample, for which all households in the contiguous U.S. in 1968 had an equal probability of selection (for a discussion of other samples, see Appendix D). The SRC sample initially selected 3,000 households, of whom 731 produced descendants two generations later who were interviewed at ages 25-45. This reduction in the number of families reflects the combined influence of survey nonresponse, childlessness, delayed fertility, and survey attrition (see Appendix D). The analytic sample includes 9,076 observations on 2,008 respondents nested in 1,101 nuclear families in 703 extended families. A total of 1,547 respondents have at least one cousin in the sample and 1,220 respondents have at least one sibling in the sample. The outcome variable of interest is log total family income, adjusted to 2014 dollars using the Consumer Price Index. To reduce the influence of very low family incomes when logged, I bottom-code inflation-adjusted family incomes at the first percentile ($2,603).

3.2 Estimation

Estimation of sibling and cousin correlations requires a decomposition of the variance of log family incomes into proportions within and between family groups. The PSID data are structured in four nested levels: observations, respondents, nuclear families, and extended families. I denote the proportion of the variance at each level by the vector $\vec{\pi}$.

$$\vec{\pi} = \begin{bmatrix} \pi_{\text{Extended}} & \pi_{\text{Nuclear}} & \pi_{\text{Person}} & \pi_{\text{Observation}} \end{bmatrix} \tag{5}$$

Siblings share both their nuclear and extended family of birth, so the sibling correlation (ρ_{Siblings}) corresponds to the proportion by which the variance of the outcome is reduced by conditioning on both extended and nuclear family of birth. The cousin correlation (ρ_{Cousins}) corresponds to the proportion by which the variance of the outcome is reduced by conditioning on extended family of birth.

$$\rho_{\text{Siblings}} = \pi_{\text{Extended}} + \pi_{\text{Nuclear}} \tag{6}$$

$$\rho_{\text{Cousins}} = \pi_{\text{Extended}} \tag{7}$$

The goal of the analysis is to estimate the partition $\vec{\pi}$, calculate the implied sibling and cousin correlations, and use these to calculate estimands of interest related to multigenerational mobility, such as the expected correlation of an ancestor and a descendant separated by k generations. Because these estimands of interest are related to the sibling and cousin correlations in nonlinear ways, translating uncertainty about $\vec{\pi}$ into uncertainty about the estimands of interest is difficult. To facilitate reporting of uncertainty, I adopt a Bayesian framework, assume prior distributions on the parameters, and estimate the model by sampling from the posterior. This procedure makes it easy to calculate uncertainty around quantities of interest; one can simply transform the posterior samples and summarize the resulting distribution.

The most central choice of prior is the prior distribution on $\vec{\pi}$, for which I assume a Dirichlet(1) distribution that places a uniform density over the simplex of possible partitions. While the Dirichlet distribution is "flat" in this sense, the marginal distribution for each component of $\vec{\pi}$ is not flat; each is a priori expected to be closer to 0 than to 1 because the components must sum to 1. As an additional complication, the sibling correlation involves the sum of two components and is by definition no smaller than the cousin correlation. The implied priors on the sibling and cousin correlations are depicted in Appendix Fig. A1 and place the highest prior density on the 75% of the distributions covering sibling correlations from 0.22 to 0.78 and cousin correlations from 0 to 0.37. Although not "flat", any reasonable value for sibling and cousin correlations falls roughly within these ranges, and the priors include additional support over the full (0,1) interval that would be utilized if the data strongly suggest a different value.

$$\underbrace{\vec{\pi} \equiv \begin{bmatrix} \pi_{\text{Extended}} & \pi_{\text{Nuclear}} & \pi_{\text{Person}} & \pi_{\text{Observation}} \end{bmatrix}}_{\text{Partition of V}(Y \mid \text{Age}) \text{ across levels of nested data}} \sim \text{Dirichlet}\left(\vec{1}\right) \tag{8}$$

As discussed previously, one solution to measurement error is to estimate sibling and cousin correlations in permanent income, treating the year-to-year observation-level variance as a nuisance (Solon et al., 1991). I estimate the sibling and cousin correlations in permanent income to be the proportion of the between-person variance component $(1 - \pi_{\text{Observation}})$ that is between nuclear and extended families, respectively.

$$\rho^{\text{Permanent}}_{\text{Cousins}} = \frac{\pi_{\text{Extended}}}{1 - \pi_{\text{Observation}}} \tag{9}$$

$$\rho^{\text{Permanent}}_{\text{Siblings}} = \frac{\pi_{\text{Extended}} + \pi_{\text{Nuclear}}}{1 - \pi_{\text{Observation}}} \tag{10}$$

The remainder of the model is not of substantive interest, but is necessary to estimate the partition $\vec{\pi}$. I assume a weakly informative Half-Cauchy prior (Gelman et al., 2008) on the total variance of age-adjusted log family incomes, which is the absolute value of a Cauchy distribution. The scale parameter of the Half-Cauchy corresponds to the median of the distribution.

$$\underbrace{\sigma^2}_{\text{V}(Y|\text{Age})} \sim \text{Half-Cauchy}\,(\text{Scale} = 1) \tag{11}$$

The variance components are connected to the observed outcome $Y_{i[j[k]],t}$ for the k-th person in the j-th nuclear family in the i-th extended family, observed at time t, by a set of normal priors on the expected age-adjusted log family income at each level. At the level of the data, the model treats each observation $Y_{i[j[k]],t}$ as drawn from a normal distribution centered at a person-specific expectation $\alpha^{\text{Person}}_{i[j[k]]}$ plus an adjustment β^{Age}_t for the reporting of incomes at different ages. The variance term $\sigma^2\pi_{\text{Observation}}$ corresponds to the within-person variance in incomes; a society characterized by substantial income insecurity such that families' incomes changed rapidly from year to year would have a high value of $\pi_{\text{Observation}}$.

$$Y_{i[j[k]],t} \sim \text{Normal}(\alpha^{\text{Person}}_{i[j[k]]} + \beta^{\text{Age}}_t,\ \ \sigma^2\pi_{\text{Observation}}) \tag{12}$$

The person-specific permanent income $\alpha^{\text{Person}}_{i[j[k]]}$ itself involves a systematic component shared within nuclear families and a normal stochastic component with variance $\sigma^2\pi_{\text{Person}}$ which corresponds to opportunities for mobility of individuals away from the family mean. In a highly fluid society in which individual attainment was stable over the life course but was only weakly associated with family background, π_{Person} would be large.

$$\alpha^{\text{Person}}_{i[j[k]]} \sim \text{Normal}(\alpha^{\text{Nuclear}}_{i[j]},\ \ \sigma^2\pi_{\text{Person}}) \tag{13}$$

The component $\alpha^{\text{Nuclear}}_{i[j]}$ of income that is constant within nuclear families likewise represents the combination of a systematic component shared within extended families and a normal stochastic component with variance $\sigma^2\pi_{\text{Nuclear}}$, which corresponds to opportunities for mobility of nuclear families away from the extended family mean. A society in which siblings' outcomes were very similar but cousins' outcomes were very different would have a high value of π_{Nuclear}.

$$\alpha^{\text{Nuclear}}_{i[j]} \sim \text{Normal}(\alpha^{\text{Extended}}_{i},\ \ \sigma^2\pi_{\text{Nuclear}}) \tag{14}$$

Finally, the component $\alpha^{\text{Extended}}_{i}$ of income that is constant within extended families represents a stochastic normal draw with variance $\sigma^2\pi_{\text{Extended}}$ and corresponds to the extent to which the status attainment of extended families look very different. A society in which advantage persisted dynastically over three generations with minimal decay would be characterized by a high value of π_{Extended}.

$$\alpha^{\text{Extended}}_{i} \sim \text{Normal}(0,\ \ \sigma^2\pi_{\text{Extended}}) \tag{15}$$

The age adjustment in Eq. 12 is needed so that respondents who happen to be observed at older ages do not appear to have higher incomes than those observed at younger ages simply as a result of their age. To flexibly capture the age-income association, I assume a multivariate normal prior on the age-specific terms $\vec{\beta}$ with a structured variance matrix with autocorrelation ρ. Intuitively, this prior regularizes toward a fit in which expected family incomes are similar at nearby ages, but allows the data to inform if the shape is more complex. In order to achieve a relatively high degree of smoothness, I rule out negative autocorrelations and place greater prior density on autocorrelations near 1 with a Beta(1,3) prior on ρ (see graphical depiction in Appendix Fig. A1). Because I do not have strong prior beliefs about the potential steepness of the association between age and log family income, I assume a weakly informative Half-Cauchy prior on the marginal variance σ^2_{Age}.

$$\sigma^2_{\text{Age}} \sim \text{Half-Cauchy}(0, 1) \tag{16}$$

$$\rho_{\text{Age}} \sim \text{Beta}(1, 3) \tag{17}$$

$$\vec{\beta}^{\text{Age}} \sim \text{Normal}\left(\vec{0}, \quad \sigma^2_{\text{Age}} \begin{bmatrix} 1 & \rho & \cdots & \rho^T \\ \rho & 1 & \cdots & \rho^{T-1} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^T & \rho^{T-1} & \cdots & 1 \end{bmatrix}\right) \tag{18}$$

The fitted age-income trajectory is plotted in Appendix Figure A2.

I estimate the model using the `rstan` package in R (Stan Development Team, 2017) which simulates from the posterior distribution using Hamiltonian Monte Carlo in Stan. I simulate 10,000 burn-in draws and 10,000 posterior samples from a single chain.[3] Point estimates reported are posterior means. Credible intervals represent the 5th and 95th percentiles of posterior draws. Trace plots for key parameters of interest are provided in Appendix Figure A3.

3.3 Results

Figure 5 summarizes the estimated sibling and cousin correlations in age-adjusted log family income. In Panel A, I assume perfect measurement ($\eta = 1$ in Eq. 1 and 2). Under the null of a first-order Markov model, the estimated sibling correlation of 0.351 would imply an expected

[3]This procedure took approximately half an hour on a Windows cluster computer with 512 GB of RAM and an Intel Xeon CPU E7-4850 v3 @2.20 GHz processor.

cousin correlation of 0.123, calculated by squaring each draw from the posterior distribution and taking the posterior mean. The observed cousin correlation of 0.158 is slightly higher than expected under the null model, suggesting the possibility of a non-Markovian process, although the 95% credible interval for the difference includes zero. These results provide non-significant but suggestive evidence of a small deviation from a first-order Markov process.

Allowing even a modest amount of measurement error in Y, however, yields results that are consistent with a first-order model. If we assume reported incomes are correlated 0.88 with true incomes due to measurement error, then the sibling and cousin correlations in true incomes (Panel B) are higher than those in reported incomes (Panel A), and the observed cousin correlation exactly equals that expected under the null model. Reviews of measurement in economics often place the reliability of log earnings in the range of 0.7 to 0.9 (Angrist and Krueger, 1999, p. 1346), so this amount of measurement error is reasonable. Therefore, the weak evidence against the null first-order Markov model in Panel A is very weak when measurement error is considered.

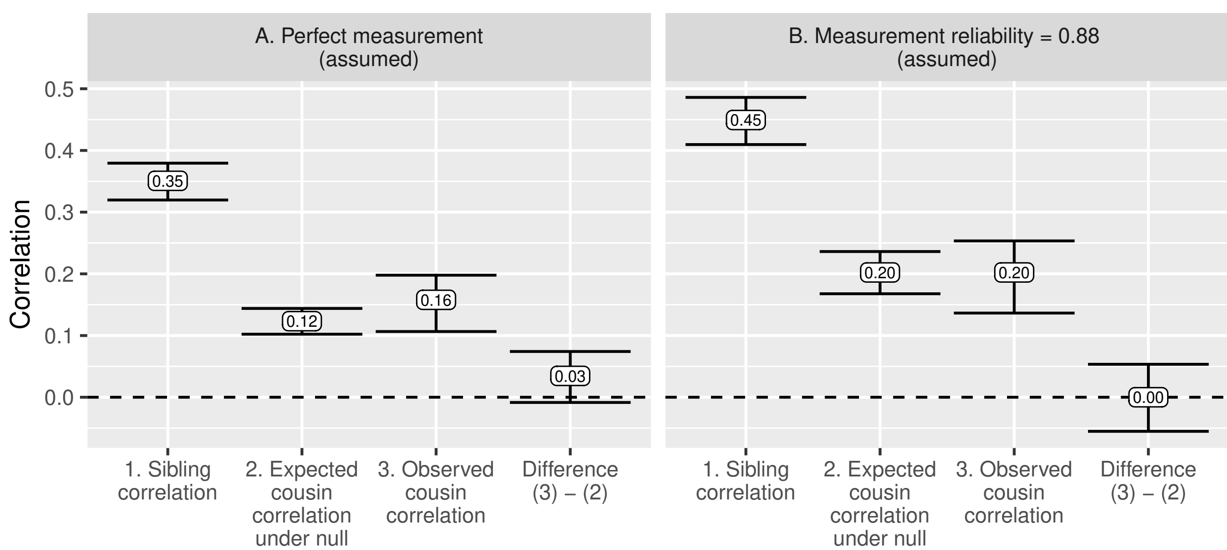

Fig. 5. Sibling and cousin correlations in log family income. Error bars represent 95% credible intervals.

While the results in Figure 5 follow the established approach in the literature of comparing to a null model, they do not directly reveal the underlying structural parameters of the second-order process that would be most consistent with the data. This is important because failure to reject the null model does not imply that the results are null; our estimate of β_2 may simply be imprecise. Figure 6 translates the estimated sibling and cousin correlations to the structural parameters β_1 and β_2, using the formulas presented in Eq. 3 and 4. Assuming perfect measurement, the observed sibling and cousin correlations are consistent with four possible parameter values represented by the blue dots: $\beta_1 = \pm 0.66$ and $\beta_2 = \{0.07, -0.50\}$. While any combination of these parameters could have generated the observed data, one might reasonably assume that the shocks to parent income are beneficial for children ($\beta_1 > 0$) and shocks to grandparent income are beneficial, even if an intervention held parent income constant ($\beta_2 > 0$). These assumptions restrict the set of possible parameters to those in the upper right quadrant of Figure 6. As in Figure 5, one must assume the reliability with which the outcome was measured. If the outcome is measured with even a small amount of error, we would underestimate β_1 and overestimate β_2. This case is represented by the green dot in the upper right quadrant, which corresponds to measurement reliability of 0.88 as in Figure 5 Panel B. Under this assumption, the estimated violation of the first-order Markov assumption is $\beta_2 = 0$ (95% CI: -.12, .11).

Although sibling correlations and their implied structural parameters may be interesting in their own right, in many cases the real object of theoretical interest is the implied pattern of multigenerational mobility. Under the second-order Markov model implied by the data, the dashed green lines in Figure 7 plot the implied multigenerational correlation in family incomes between an ancestor in generation 0 and a descendant in generations 1 through 10. The solid blue line represents the analogous pattern which we would estimate by assuming a first-order Markov model and estimating the structural parameter β_1 from the sibling correlations alone. The left panel assumes perfect measurement of Y and shows persistence is only slightly higher under the second-order estimates than under the first-order estimates. The correlation of family incomes separated by 5 generations, for instance, is 0.07 (95% CI: 0.06, 0.09) under the first-order model compared with 0.11 (95% CI: 0.06, 0.15) under the second-order model. As discussed previously, the distinction could easily be explained by measurement error; assuming measurement reliability of 0.88 yields slightly higher estimates of multigenerational correlations that are substantively the same under

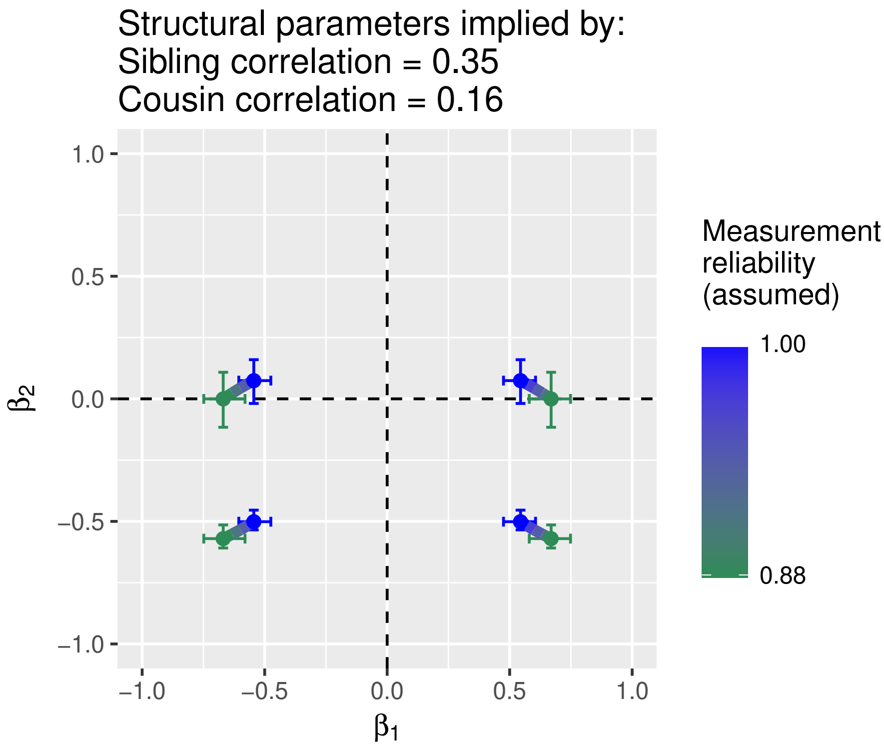

Fig. 6. Structural parameters for log family income transmission implied by sibling and cousin correlations under a second-order Markov model assuming the most positive β_1 and β_2 consistent with the observed data. Error bars represent 95% credible intervals.

either model (Panel B).

While the results above yielded minimal evidence against a first-order Markov model, they also revealed the possibility that in some applications this conclusion could depend on the assumed measurement reliability. To make an alternative set of assumptions about measurement reliability, one might shift the target of inference to intragenerational correlations in permanent age-adjusted log family income, defined as correlations in the $\vec{\alpha}^{\text{Person}}$ parameters that pool information across multiple time periods (Solon et al., 1991). To the extent to which measurement error is not correlated within individuals, this would fully address the measurement problem. Results are presented in Figure 8. Point estimates suggest that cousins' permanent incomes are slightly *less* positively correlated under the second-order model than the first-order model (Panel A), though estimates of the structural parameter are highly uncertain (β_2 = -.08, 95% CI = [-.22, 0.05], Panel B). The point estimates suggest that the multigenerational persistence of permanent incomes is less than expected

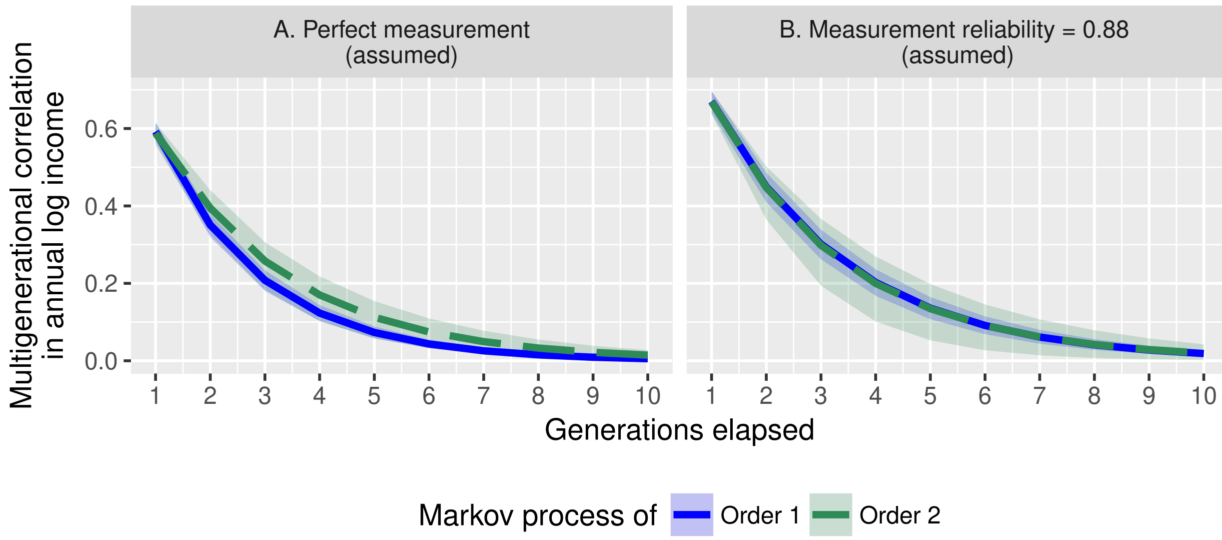

Fig. 7. Multigenerational correlations in log family income implied by sibling and cousin correlations under a second-order Markov model assuming the most positive β_1 and β_2 consistent with the observed data. Error bars represent 95% credible intervals.

under a Markov model (Panel C), though results are too uncertain to draw definitive conclusions. If reporting errors are correlated within individuals such that some individuals consistently over- or under-report their incomes, this would produce measurement error in permanent income ($\eta < 1$). In this case, results would suggest that β_2 is even more negative. Overall, results assessing correlations in permanent incomes reaffirm that any evidence that multigenerational correlations are greater than would be expected under a Markov model is very weak.

4 Re-evaluating prior research

The preceding sections have outlined a procedure for using sibling and cousin correlations to draw inferences about multigenerational mobility: estimate intragenerational correlations, compare to the null hypothesis of a first-order Markov model to assess statistical significance, evaluate the extent of measurement error that would undermine non-Markovian conclusions, and then assess the substantive significance of the results. Prior research, in contrast, has only conducted the hypothesis test, typically without a formal test of significance and without considering the role

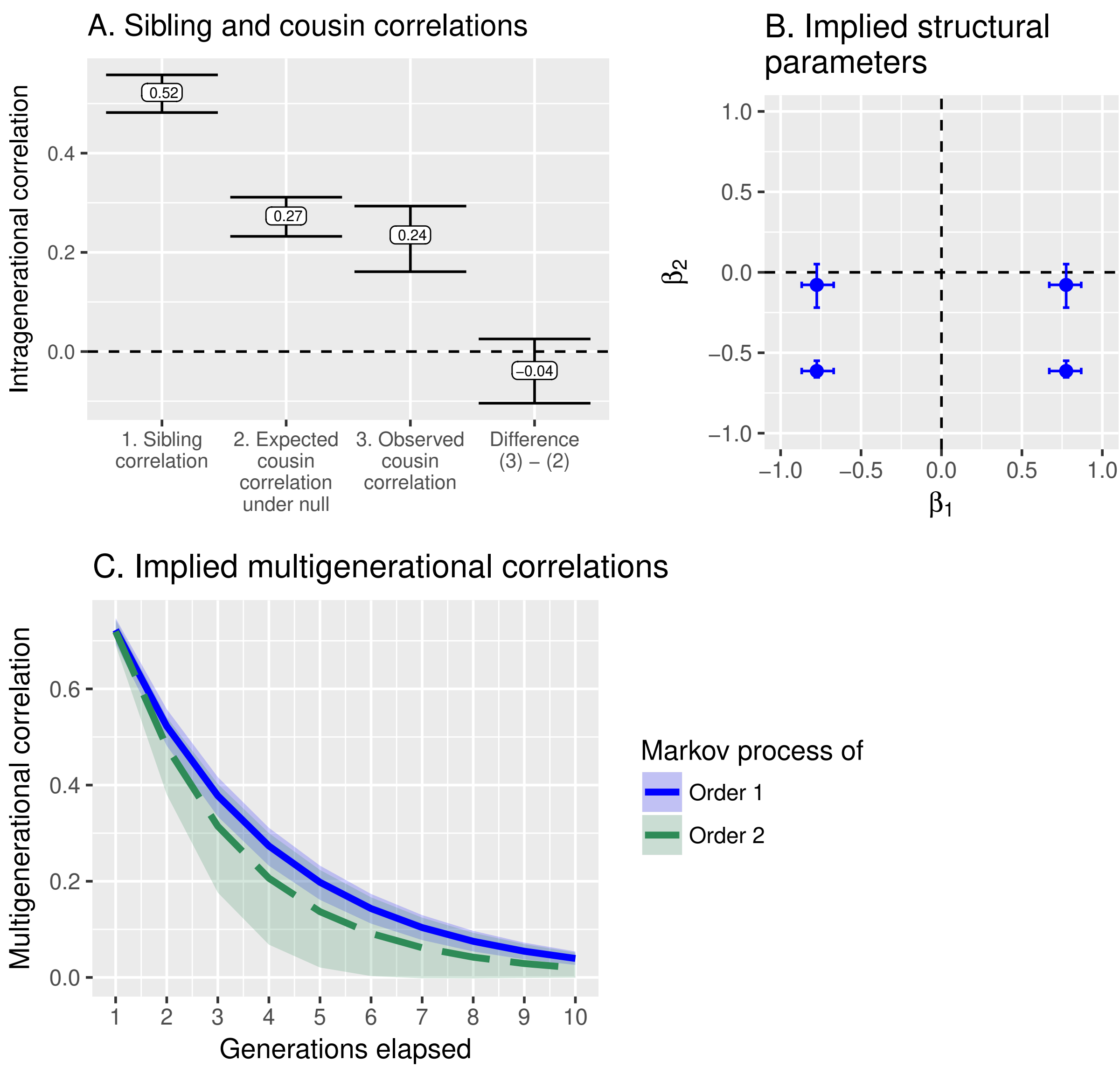

Fig. 8. Sibling and cousin correlations in permanent age-adjusted log family incomes and the implied mobility regimes with which they are compatible. Error bars and shaded bands represent 95% credible intervals.

of measurement error. This section re-evaluates prior research in light of the procedure outlined above.

What degree of measurement precision is required for published cousin correlations to imply a non-Markovian status transmission process? Table 1 presents published sibling and cousin correlations and the degree of measurement precision required for them to provide evidence against a Markov process. In the text, I highlight one estimate from each paper. Overall, published estimates demonstrate that familial advantages persist over three generations; the total association of cousins' attainment is substantial. However, only a few estimates provide evidence against a Markov transmission process, and only under an assumption of very precise measurement.

Analyzing educational attainment in the Wisconsin Longitudinal Study, Jæger (2012) estimates that 14.4% of the variance in years of education is between extended families, and 26.4% of the variance is between nuclear families within extended families. These estimates indicate a sibling correlation of $0.144 + 0.264 = 0.407$. The estimated cousin correlation of 0.166 does suggest important *total* grandparent effects, thereby supporting the author's claim that "factors in the extended family contribute to the total effect of family background on G_3 educational success." Estimates do not, however, reject a Markov model in which grandparents transmit these advantages indirectly via parents. Under perfect measurement ($\eta = 1$), this sibling correlation would imply a cousin correlation of $\rho_Y^{\text{Cousin, implied}} = \frac{1}{\eta^2}\rho_Y^{\text{Sibling}^2} = \frac{1}{1} \times 0.407^2 = 0.166$. This prediction is very close to the estimated cousin correlation of 0.144, providing no evidence against a Markov model. Nevertheless, Jaeger concludes, "a two-generation Markov process does not represent the total effect of family background on educational success," (p. 913).

Using Swedish register data, Hällsten (2014) estimates that cousins' years of education are correlated 0.15, which the author compares to an estimate by Björklund and Jäntti (2012) indicating that the sibling correlation is 0.39. This cousin correlation is approximately the sibling correlation squared as predicted under a Markov model without measurement error ($\rho_Y^{\text{Cousin, implied}} = \frac{1}{\eta^2}\rho_Y^{\text{Sibling}^2} = \frac{1}{1} \times 0.39^2 \approx 0.15$). Nevertheless, on the basis of the inability to explain away the cousin correlation with measured covariates, the author concludes, "unless *unobserved* characteristics of parents account for *all* of the 1st and 2nd cousin correlations, the estimated and adjusted correlations are clearly incompatible with a Markov process," (p. 31, emphasis in original). This conclusion depends on one's beliefs about whether unobserved parent characteristics might matter for children's life chances; if you believe that unobserved parent characteristics might be quite important (i.e. embodied cultural capital, Bourdieu 1986), then the author's estimates are compatible with a Markov process. To argue otherwise, one would have to explain why the cousin correlation

Author	Context	Outcome	Sibling correlation ρ_Y^{Sibling}	Cousin correlation ρ_Y^{Cousin}	Expected cousin correlation $\rho_Y^{\text{Cousin, implied}} = \frac{1}{\eta^2}\left(\rho_Y^{\text{Sibling}}\right)^2$	Violates Markovian process?
Jæger (2012)	Wisconsin	Years of education	0.41	0.14	$\frac{1}{\eta^2} \times 0.17$	No
		Log odds of college completion	0.56	0.21	$\frac{1}{\eta^2} \times 0.31$	No
	NLSY-CYA[a]	PIAT math score	0.38	0.23	$\frac{1}{\eta^2} \times 0.14$	Only if $\lvert\eta\rvert > 0.79$
		PIAT reading score	0.37	0.19	$\frac{1}{\eta^2} \times 0.14$	Only if $\lvert\eta\rvert > 0.85$
		Log odds of high school completion	0.37	0.14	$\frac{1}{\eta^2} \times 0.14$	No
		Years of education	0.37	0.26	$\frac{1}{\eta^2} \times 0.14$	Only if $\lvert\eta\rvert > 0.73$
		Log odds of college completion	0.45	0.29	$\frac{1}{\eta^2} \times 0.20$	Only if $\lvert\eta\rvert > 0.83$
Hällsten (2014)	Sweden	Years of education	0.39[b]	0.15	$\frac{1}{\eta^2} \times 0.15$	No
		Occupational prestige	0.29	0.11	$\frac{1}{\eta^2} \times 0.08$	Only if $\lvert\eta\rvert > 0.87$
		GPA (9th grade)	0.51[b]	0.19	$\frac{1}{\eta^2} \times 0.26$	No
		Cognitive ability	0.47[b]	0.16	$\frac{1}{\eta^2} \times 0.22$	No
Knigge (2016)	Netherlands	Occupation status	0.50	0.32	$\frac{1}{\eta^2} \times 01$	Only if $\lvert\eta\rvert > 0.88$
Pfeffer et al. (2016)	U.S.	Wealth	0.34	0.19	$\frac{1}{\eta^2} \times 0.12$	Only if $\lvert\eta\rvert > 0.78$

Note: In each row above, the assumed correlation between the measured outcome Y and the true value purged of measurement error is η, which is unobserved but must lie between -1 and 1. An assumption of no measurement error corresponds to an assumption of $\eta = 1$. [a] The original author notes that the NLSY-CYA sample was very young and highly selected at the time of analysis, so results should be interpreted with caution. This is especially true of the models of college completion and years of schooling, which appear to violate the Markov assumption but which are estimated on a highly selected sample of respondents whose mothers gave birth at young ages, such that the respondent was at least 25 years old by the time of data collection. [b] With the exception of occupational prestige, Hällsten (2014) does not directly estimate sibling correlations but cites the sibling correlations reported by Björklund and Jäntti (2012).

Table 1. Prior estimates of cousin correlations in socioeconomic and academic outcomes

is not higher relative to the sibling correlation.

Knigge (2016) studies occupational status attainment in the Netherlands in the late 19th and early 20th century and estimates a sibling correlation of 0.50 and a cousin correlation of 0.32. The author notes that this cousin correlation is slightly higher than the 0.25 estimate that would be predicted by a Markov model. However, if the measure of occupational status is only correlated $\eta = 0.88$ with latent status attainment due to reporting or construct error, this would yield a Markov prediction of $\rho_Y^{\text{Cousin, implied}} = \frac{1}{\eta^2}\rho_Y^{\text{Sibling}^2} = \frac{1}{0.88^2}0.50^2 = 0.32$. The author's conclusion that "these results are not congruent with the Markovian model" (p. 1233) thus depends on an unstated assumption that measurement of status attainment is very precise.

Finally, using PSID data from the U.S., Pfeffer et al. (2016) estimate a sibling correlation in log net worth of 0.34, which implies a cousin correlation of $0.34^2 = 0.12$. The authors estimate a cousin correlation of 0.19, which is slightly higher than expected and warrants the authors' conclusion that "19% of individuals' wealth attainment can be traced to the common origins of cousins (i.e. grandparent environments), reflecting concentration of family wealth within lineages beyond just two generations." The paper makes no claims about whether this multigenerational persistence violates Markovian predictions, but the estimate is consistent with Markovian predictions if the correlation between wealth and latent status attainment is 0.78, so that $\rho_Y^{\text{Cousins, implied}} = \frac{1}{\eta^2}\rho_Y^{\text{Siblings}^2} = \frac{1}{0.78^2}0.34^2 = 0.19$.

Each of the authors discussed above presents additional evidence based on intergenerational correlations to support their claims, and the purpose of this review is not to discredit the contribution of these authors. Their estimates have brought a novel framework to bear on an old question, yielding an important result: grandparent advantages continue to be associated with life chances two generations later. However, the estimates provide less compelling evidence with respect to direct grandparent associations: cousin correlations are not substantially greater than would be expected if the status transmission process operates from parent to child.

5 Limitations

Whether the process of status transmission can be represented by a first-order Markov process may be context-dependent. The context of mid-20th century Wisconsin, for instance, might have been a period in which transmission operated under this regime (Warren and Hauser, 1997). The transmission process in rural China, on the other hand, may be quite different (Zeng and Xie, 2014). It is likewise possible that in the U.S. grandparents may play a direct role in the attainment of other indicators of socioeconomic status, such a wealth (Pfeffer and Killewald, 2017), even if family incomes follow a first-order Markov transmission process. Finally, Mare (2011) proposed that attainment may be especially rigid at the extremes of the distribution, such as among those living for multiple generations in poor neighborhoods (Sharkey and Elwert, 2011) or those who benefit from legacy admissions to elite colleges (Karabel, 2006). Future research will need to evaluate transmission in other contexts, with other outcomes, and in various subpopulations; I hope that the framework presented in this paper will be useful for this task.

Importantly, conclusions from intragenerational correlations are sensitive to the assumptions required and require acceptance of a very particular definition of family background. If family background affects siblings differently, as in the example of parents investing all resources in the youngest child and forcing the oldest child to sacrifice for the success of the youngest, results may be misleading. While this extreme violation may be far-fetched, qualitative evidence suggests that some resource-constrained families do invest more heavily in one sibling (Conley, 2004). Future work should take seriously the possibilities of complex intra-family dynamics that would make the variables that are associated with the outcomes of all siblings in common a poor proxy for the association between family background and life chances.

6 Discussion

Sibling and cousin correlations provide a promising source of information about the association between family background and life chances, yet previous literature has not explicitly formalized the relationship between these correlations and the underlying multigenerational process. This paper clarifies the assumptions under which intragenerational correlations are informative about multigenerational processes, outlining a procedure to report substantively meaningful estimates rather than relying on hypothesis testing alone. Fresh analysis of PSID data and a reanalysis of published research suggest that, once a reasonable degree of measurement error is assumed, cousin correlations do not deviate substantially from what would be expected in a first-order Markov transmission process that operates from parent to child.

This paper contributes intragenerational evidence to join recent intergenerational evidence (Ferrie et al., 2016) indicating that apparent direct grandparent effects that skip from grandparents to offspring are likely a statistical artifact arising from measurement error. However, readers are cautioned against interpreting this as a claim that the U.S. is a particularly fluid society: *total* grandparent effects (direct + indirect) remain substantial and important, even if they operate through the parent generation. Cousin correlations in the U.S. and other settings are sizable and reflect the persistent association between familial advantages and life chances over multiple generations. In agreement with early work on sibling correlations (Solon et al., 1991), my results support the notion that the total association between family background on life chances is in many cases larger than previously believed due to measurement error in socioeconomic outcomes that reduces estimated sibling and cousin correlations in cross-sectional surveys.

In summary, what do cousin correlations in socioeconomic outcomes tell us? When paired with sibling correlations, cousin correlations imply a range of possible multigenerational processes that are consistent with the data under a variety of assumptions. Cousin correlations are also a clear reminder that processes involving only two generations at a time can still have long-term correlations with the life chances of future generations, even if these associations may operate indirectly through the intervening generations. In the case of the U.S. in the late 20th century, sibling and cousin correlations that account for measurement error demonstrate that the remarkable extent to which opportunity is constrained at birth reaches far beyond what we might find with observed measures of socioeconomic advantage, but operates one generation at a time.

References

Angrist, J. D. and A. B. Krueger 1999. Empirical strategies in labor economics. *Handbook of Labor Economics*, 3:1277–1366.

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

Bengtson, V. L. 2001. Beyond the nuclear family: The increasing importance of multigenerational bonds. *Journal of Marriage and Family*, 63(1):1–16.

Bielby, W. T., R. M. Hauser, and D. L. Featherman 1977. Response errors of black and non-black males in models of the intergenerational transmission of socioeconomic status. *American Journal of Sociology*, 82(6):1242–1288.

Björklund, A. and M. Jäntti 2012. How important is family background for labor-economic outcomes? *Labour Economics*, 19(4):465–474.

Blau, P. M. and O. D. Duncan 1967. *The American Occupational Structure.* New York: Wiley.

Bourdieu, P. 1986. The forms of capital. In *Handbook of Theory and Research for the Sociology of Education*, J. G. Richardson, ed., Pp. 241–258. New York: Greenwood Press.

Bowles, S. 1972. Schooling and inequality from generation to generation. *Journal of Political Economy*, 80(3, Part 2):S219–S251.

Chan, T. W. and V. Boliver 2013. The grandparents effect in social mobility: Evidence from british birth cohort studies. *American Sociological Review*.

Conley, D. 2004. *The Pecking Order: Which Siblings Succeed and Why*. New York: Pantheon.

Conley, D. 2008. Bringing sibling differences in: Enlarging our understanding of the transmission of advantage in families. In *Social Class: How Does it Work?*, A. Lareau and D. Conley, eds., Pp. 179–200. New York: Russell Sage Foundation.

Conley, D., K. M. Pfeiffer, and M. Velez 2007. Explaining sibling differences in achievement and behavioral outcomes: The importance of within-and between-family factors. *Social Science Research*, 36(3):1087–1104.

Corcoran, M., C. Jencks, and M. Olneck 1976. The effects of family background on earnings. *The American Economic Review*, 66(2):430–435.

Ferrie, J., C. Massey, and J. Rothbaum 2016. Do grandparents and great-grandparents matter? multigenerational mobility in the U.S., 1910-2013. Technical report, National Bureau of Economic Research.

Gelman, A., A. Jakulin, M. G. Pittau, and Y.-S. Su 2008. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, Pp. 1360–1383.

Griliches, Z. 1979. Sibling models and data in economics: Beginnings of a survey. *Journal of Political Economy*, 87(5, Part 2):S37–S64.

Hällsten, M. 2014. Inequality across three and four generations in egalitarian sweden: 1st and 2nd cousin correlations in socio-economic outcomes. *Research in Social Stratification and Mobility*, 35:19–33.

Hällsten, M. and F. T. Pfeffer 2017. Grand advantage: Family wealth and grandchildren?s educational achievement in sweden. *American Sociological Review*, 82(2):328–360.

Hauser, R. M. and D. L. Featherman 1977. *The Process of Stratification: Trends and Analyses*. New York: Academic Press.

Hertel, F. R. and O. Groh-Samberg 2014. Class mobility across three generations in the US and Germany. *Research in Social Stratification and Mobility*, 35:35–52.

Hout, M. 1988. More universalism, less structural mobility: The American occupational structure in the 1980s. *American Journal of Sociology*, 93(6):1358–1400.

Hout, M. and T. A. DiPrete 2006. What we have learned: RC28's contributions to knowledge about social stratification. *Research in Social Stratification and Mobility*, 24:1–20.

Huang, J. Y., A. R. Gavin, T. S. Richardson, A. Rowhani-Rahbar, D. S. Siscovick, and D. A. Enquobahrie 2015. Are early-life socioeconomic conditions directly related to birth outcomes? Grandmaternal education, grandchild birth weight, and associated bias analyses. *American Journal of Epidemiology*, 182(7):568–578.

Jæger, M. M. 2012. The extended family and children's educational success. *American Sociological Review*, 77(6):903–922.

Kalleberg, A. L. 2009. Precarious work, insecure workers: Employment relations in transition. *American Sociological Review*, 74(1):1–22.

Karabel, J. 2006. *The Chosen: The Hidden History of Admission and Exclusion at Harvard, Yale, and Princeton*. Boston: Houghton Mifflin.

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

Lindahl, M., M. Palme, S. S. Massih, and A. Sjögren 2015. Long-term intergenerational persistence of human capital an empirical analysis of four generations. *Journal of Human Resources*, 50(1):1–33.

Mare, R. D. 2011. A multigenerational view of inequality. *Demography*, 48(1):1–23.

Mazumder, B. 2008. Sibling similarities and economic inequality in the US. *Journal of Population Economics*, 21(3):685–701.

Mykyta, L. and S. Macartney 2011. The effects of recession on household composition: "Doubling up" and economic well-being. *U.S. Census Bureau SEHSD Working Paper Number 2011-4*.

Northrop, F. S. C. 1947. *The Logic of the Sciences and the Humanities*. New York: Macmillan.

Olivetti, C., M. D. Paserman, and L. Salisbury 2016. Three-generation mobility in the United States, 1850-1940: The role of maternal and paternal grandparents. Technical report, National Bureau of Economic Research.

Pfeffer, F. T. 2014. Multigenerational approaches to social mobility: A multifaceted research agenda. *Research in Social Stratification and Mobility*, 35:1.

Pfeffer, F. T. and A. Killewald 2015. How rigid is the wealth structure and why? inter-and multigenerational associations in family wealth. *Population Studies Center Research Report*, Pp. 15–845.

Pfeffer, F. T. and A. Killewald 2017. Generations of advantage: Multigenerational correlations in family wealth. *Social Forces*, Online first.

Pfeffer, F. T., A. Killewald, and A. Siliunas 2016. The concentration of wealth within family lineages and intergenerational transfers. In *Conference on New Directions in Intergenerational Transfers and Time Use in Later Life*.

Pilkauskas, N. V., I. Garfinkel, and S. S. McLanahan 2014. The prevalence and economic value of doubling up. *Demography*, 51(5):1667–1676.

Pilkauskas, N. V. and M. L. Martinson 2014. Three-generation family households in early childhood: Comparisons between the United States, the United Kingdom, and Australia. *Demographic Research*, 30:1639.

PSID 2017. *Panel Study of Income Dynamics, Public Use Dataset*. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI (2017).

Seltzer, J. A. and S. M. Bianchi 2013. Demographic change and parent-child relationships in adulthood. *Annual Review of Sociology*, 39:275–290.

Sewell, W. H., A. O. Haller, and A. Portes 1969. The educational and early occupational attainment process. *American Sociological Review*, Pp. 82–92.

Sharkey, P. and F. Elwert 2011. The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability. *American Journal of Sociology*, 116(6):1934–81.

Si, Y., N. S. Pillai, and A. Gelman 2015. Bayesian nonparametric weighted sampling inference. *Bayesian Analysis*, 10(3):605–625.

Siegel, P. M. and R. W. Hodge 1968. A causal approach to the study of measurement error. *Methodology in Social Research. New York: McGraw-Hill*, 28(59):35.

Solon, G., M. Corcoran, R. Gordon, and D. Laren 1991. A longitudinal analysis of sibling correlations in economic status. *The Journal of Human Resources*, Pp. 509–534.

Solon, G., M. E. Page, and G. J. Duncan 2000. Correlations between neighboring children in their subsequent educational attainment. *Review of Economics and Statistics*, 82(3):383–392.

Song, X. and C. D. Campbell 2017. Genealogical microdata and their significance for social science. *Annual Review of Sociology*, 43:75–99.

Stan Development Team 2017. *RStan: the R interface to Stan*. R package version 2.16.2.

Taylor, P., J. Passel, R. Fry, R. Morin, W. Wang, G. Velasco, and D. Dockterman 2010. The return of the multi-generational family household. Technical report, Pew Research Center.

Uhlenberg, P. 2004. Historical forces shaping grandparent-grandchild relationships: Demography and beyond. *Annual Review of Gerontology and Geriatrics*, 24:77–97.

Walker, G. 1931. On periodicity in series of related terms. *Philosophical Transactions of the Royal Statistical Society A: Mathematical, Physical, and Engineering Sciences*, 131:518–532.

Warren, J. R. and R. M. Hauser 1997. Social stratification across three generations: New evidence from the Wisconsin Longitudinal Study. *American Sociological Review*, Pp. 561–572.

Wightman, P. and S. Danziger 2014. Multi-generational income disadvantage and the educational attainment of young adults. *Research in Social Stratification and Mobility*, 35:53–69.

Yule, G. U. 1927. On a method of investigating the periodicities of a disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Statistical Society A: Mathematical, Physical, and Engineering Sciences*, 226:267–298.

Zeng, Z. and Y. Xie 2014. The effects of grandparents on children?s schooling: Evidence from rural china. *Demography*, 51(2):599–617.

Zhang, Q. F. 2004. Economic transition and new patterns of parent-adult child coresidence in urban china. *Journal of Marriage and Family*, 66(5):1231–1245.

Ziefle, A. 2016. Persistent educational advantage across three generations: Empirical evidence for Germany. *Sociological Science*, 3:1077–1102.

Appendix

A Yule-Walker equations for multigenerational correlations

This section derives the ρ_k in status attainment Y (standardized to mean 0 and variance 1) for two individuals separated by k generations under a second-order Markov process. These formulas were originally proposed by Yule (1927) and Walker (1931) and are standard in time series analysis. The structural regression model for attainment in generation t as a function of attainment in the prior two generations is given by Eq. 19. The term ϵ_t represents an independent error.

$$Y_t = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \epsilon_t \tag{19}$$

Multiply both sides of Eq. 19 by Y_{t-1} and take the expected value to yield a formula for ρ_1, the correlation in the attainment of parents and children.

$$\begin{aligned}
Y_t Y_{t-1} &= \beta_1 Y_{t-1}^2 + \beta_2 Y_{t-1} Y_{t-2} + Y_{t-1}\epsilon_t \\
\mathbf{E}\left(Y_t Y_{t-1}\right) &= \mathbf{E}\left(\beta_1 Y_{t-1}^2 + \beta_2 Y_{t-1} Y_{t-2} + Y_{t-1}\epsilon_t\right) \\
\underbrace{\mathbf{E}\left(Y_t Y_{t-1}\right)}_{=\text{Cor}(Y_t, Y_{t-1})=\rho_1} &= \beta_1 \underbrace{\mathbf{E}\left(Y_{t-1}^2\right)}_{=\text{V}(Y_{t-1})=1} + \beta_2 \underbrace{\mathbf{E}\left(Y_{t-1} Y_{t-2}\right)}_{=\text{Cor}(Y_{t-1}, Y_{t-2})=\rho_1} + \underbrace{\mathbf{E}\left(Y_{t-1}\epsilon_t\right)}_{=0 \text{ since independent}} \\
\rho_1 &= \beta_1 + \beta_2 \rho_1 \\
\rho_1 &= \frac{\beta_1}{1-\beta_2}
\end{aligned} \tag{20}$$

Multiply both sides of Eq. 19 by Y_{t-2} and take the expected value to yield a formula for ρ_2, the correlation in the attainment of grandparents and children.

$$\begin{aligned}
\mathbf{E}\left(Y_t Y_{t-2}\right) &= \mathbf{E}\left(\beta_1 Y_{t-1} Y_{t-2} + \beta_2 Y_{t-2}^2 + Y_{t-2}\epsilon_t\right) \\
\underbrace{\mathbf{E}\left(Y_t Y_{t-2}\right)}_{=\text{Cor}(Y_t, Y_{t-2})=\rho_2} &= \beta_1 \underbrace{\mathbf{E}\left(Y_{t-1}, Y_{t-2}\right)}_{=\text{Cor}(Y_{t-1}, Y_{t-2})=\rho_1} + \beta_2 \underbrace{\mathbf{E}\left(Y_{t-2}^2\right)}_{=\text{V}(Y_{t-2})=1} + \underbrace{\mathbf{E}\left(Y_{t-2}\epsilon_t\right)}_{=0 \text{ since independent}} \\
\rho_2 &= \beta_1 \rho_1 + \beta_2 = \frac{\beta_1^2}{1-\beta_2} + \beta_2
\end{aligned} \tag{21}$$

To generalize the formula to the correlation between two individuals separated by an arbitrary number of generations k, multiply both sides of Eq. 19 by Y_{t-k} and take the expected value.

$$\begin{aligned}
\mathbf{E}\left(Y_t Y_{t-k}\right) &= \mathbf{E}\left(\beta_1 Y_{t-1} Y_{t-k} + \beta_2 Y_{t-2} Y_{t-k} + Y_{t-k}\epsilon_t\right) \\
\underbrace{\mathbf{E}\left(Y_t Y_{t-k}\right)}_{=\text{Cor}(Y_t, Y_{t-k})=\rho_k} &= \beta_1 \underbrace{\mathbf{E}\left(Y_{t-1}, Y_{t-k}\right)}_{=\text{Cor}(Y_{t-1}, Y_{t-k})=\rho_{k-1}} + \beta_2 \underbrace{\mathbf{E}\left(Y_{t-2} Y_{t-k}\right)}_{=\text{Cor}(Y_{t-2}, Y_{t-k})=\rho_{k-2}} + \underbrace{\mathbf{E}\left(Y_{t-2}\epsilon_t\right)}_{=0 \text{ since independent}} \\
\rho_k &= \beta_1 \rho_{k-1} + \beta_2 \rho_{k-2}
\end{aligned} \tag{22}$$

B Derivation of sibling and cousin correlations from structural parameters

Building on the results in Appendix A, this section derives the sibling and cousin correlations that result from a second-order Markov transmission process with structural parameters β_1 and β_2.

Denoting attainment of siblings A and B in generation t by $Y_{t,a}$ and $Y_{t,b}$ and using Eq. 19,

$$Y_{t,a} = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \epsilon_{t,a} \tag{23}$$

$$Y_{t,b} = \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \epsilon_{t,b} \tag{24}$$

Multiplying Eq. 36 and 37 and taking the expected value,

$$\rho^{Y}_{\text{Siblings}} = \mathbf{E}\left(Y_{t,a}Y_{t,b}\right) \tag{25}$$

$$= \mathbf{E}\left[\left(\beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \epsilon_{t,a}\right)\left(\beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \epsilon_{t,b}\right)\right] \tag{26}$$

$$= \beta_1^2 \underbrace{\mathbf{E}\left(Y_{t-1}^2\right)}_{=\text{V}(Y_{t-1})=1} + \beta_2^2 \underbrace{\mathbf{E}\left(Y_{t-2}^2\right)}_{=\text{V}(Y_{t-2})=1} + 2\beta_1\beta_2 \underbrace{\mathbf{E}\left(Y_{t-1}Y_{t-2}\right)}_{=\text{Cor}(Y_{t-1},Y_{t-2})=\rho_1} \tag{27}$$

$$+ \underbrace{\mathbf{E}\left[\left(\beta_1 Y_{t-1} + \beta_2 Y_{t-2}\right)\epsilon_{t,b}\right] + \mathbf{E}\left[\left(\beta_1 Y_{t-1} + \beta_2 Y_{t-2}\right)\epsilon_{t,a}\right]}_{=0 \text{ since error terms independent}}$$

$$= \beta_1^2 + \beta_2^2 + 2\beta_1\beta_2\rho_1 \tag{28}$$

Denoting the observed sibling outcomes $\tilde{Y}_{t,a}$ and $\tilde{Y}_{t,b}$ and following similar steps,

$$\tilde{Y}_{t,a} = \eta Y_{t,a} + \delta_{t,a} \tag{29}$$

$$\tilde{Y}_{t,b} = \eta Y_{t,b} + \delta_{t,b} \tag{30}$$

$$\rho_{\text{Siblings}} = \mathbf{E}\left(\tilde{Y}_{t,a}\tilde{Y}_{t,b}\right) \tag{31}$$

$$= \mathbf{E}\left[\left(\eta Y_{t,a} + \delta_{t,a}\right)\left(\eta Y_{t,b} + \delta_{t,b}\right)\right] \tag{32}$$

$$= \eta^2 \underbrace{\mathbf{E}\left(Y_{t,a}, Y_{t,b}\right)}_{=\text{Cor}\left(Y_{t,a},Y_{t,b}\right)=\rho_{G,\text{Siblings}}} + \underbrace{\eta\mathbf{E}\left(Y_{t,a}, \delta_{t,b}\right) + \eta\mathbf{E}\left(Y_{t,b}, \delta_{t,a}\right)}_{=0 \text{ since error terms independent}} \tag{33}$$

$$= \eta^2 \rho^{Y}_{\text{Siblings}} \tag{34}$$

$$= \eta^2\left(\beta_1^2 + \beta_2^2 + 2\beta_1\beta_2\rho_1\right) \tag{35}$$

A similar process yields formulas for the cousin correlation under this process. Denoting the attainment of cousins A and B in generation t by $Y_{t,aa}$ and $Y_{t,bb}$ and using Eq. 19,

$$Y_{t,aa} = \beta_1 Y_{t-1,a} + \beta_2 Y_{t-2} + \epsilon_{t,a} \tag{36}$$

$$Y_{t,bb} = \beta_1 Y_{t-1,b} + \beta_2 Y_{t-2} + \epsilon_{t,b} \tag{37}$$

Multiplying Eq. 36 and 37 and taking the expected value,

$$
\begin{aligned}
\rho^{Y}_{\text{Cousins}} &= \mathbf{E}\left(Y_{t,aa}Y_{t,bb}\right) && (38)\\
&= \mathbf{E}\left[\left(\beta_1 Y_{t-1,a} + \beta_2 Y_{t-2} + \epsilon_{t,a}\right)\left(\beta_1 Y_{t-1,b} + \beta_2 Y_{t-2} + \epsilon_{t,b}\right)\right] && (39)\\
&= \beta_1^2 \underbrace{\mathbf{E}\left(Y_{t-1,a}Y_{t-1,b}\right)}_{=\text{Cor}\left(Y_{t-1,a}Y_{t-1,b}\right)=\rho_{G,\text{Siblings}}} + \beta_2^2 \underbrace{\mathbf{E}\left(Y_{t-2}^2\right)}_{=\text{V}(Y_{t-2})=1} && (40)\\
&\quad + \beta_1\beta_2 \underbrace{\mathbf{E}\left(Y_{t-1,a}Y_{t-2}\right)}_{=\text{Cor}(Y_{t-1,a},Y_{t-2})=\rho_1} + \beta_1\beta_2 \underbrace{\mathbf{E}\left(Y_{t-1,b}Y_{t-2}\right)}_{=\text{Cor}\left(Y_{t-1,b},Y_{t-2}\right)=\rho_1}\\
&\quad + \underbrace{\mathbf{E}\left[\left(\beta_1 Y_{t-1,a} + \beta_2 Y_{t-2}\right)\epsilon_{t,b}\right] + \mathbf{E}\left[\left(\beta_1 Y_{t-1,b} + \beta_2 Y_{t-2}\right)\epsilon_{t,a}\right]}_{=0 \text{ since error terms independent}}\\
&= \beta_1^2 \rho^{Y}_{\text{Siblings}} + \beta_2^2 + 2\beta_1\beta_2\rho_1 && (41)
\end{aligned}
$$

Denoting the observed cousin outcomes $\tilde{Y}_{t,aa}$ and $\tilde{Y}_{t,bb}$ and following similar steps, we derive the correlations in the observed values.

$$
\begin{aligned}
\tilde{Y}_{t,aa} &= \eta Y_{t,aa} + \delta_{t,aa} && (42)\\
\tilde{Y}_{t,bb} &= \eta Y_{t,bb} + \delta_{t,bb} && (43)\\
\rho_{\text{Cousins}} &= \mathbf{E}\left(\tilde{Y}_{t,aa}\tilde{Y}_{t,bb}\right) && (44)\\
&= \mathbf{E}\left[\left(\eta Y_{t,aa} + \delta_{t,aa}\right)\left(\eta Y_{t,bb} + \delta_{t,bb}\right)\right] && (45)\\
&= \eta^2 \underbrace{\mathbf{E}\left(Y_{t,aa}, Y_{t,bb}\right)}_{=\text{Cor}\left(Y_{t,aa},Y_{t,bb}\right)=\rho^{Y}_{\text{Cousins}}} + \underbrace{\eta\mathbf{E}\left(Y_{t,aa}, \delta_{t,bb}\right) + \eta\mathbf{E}\left(Y_{t,bb}, \delta_{t,aa}\right)}_{=0 \text{ since error terms independent}} && (46)\\
&= \eta^2 \rho^{Y}_{\text{Cousins}} && (47)\\
&= \eta^2\left[\beta_1^2 \rho^{Y}_{\text{Siblings}} + \beta_2^2 + 2\beta_1\beta_2\rho_1\right] && (48)
\end{aligned}
$$

C Structural parameters implied by sibling and cousin correlations

Starting with the sibling and cousin correlation formulas derived in Appendix B, we can rearrange terms to produce formulas for the structural parameters β_1 and β_2 for a given set of sibling and cousin correlations $\{\rho_{Y,\text{Siblings}}, \rho_{Y,\text{Cousins}}\}$ and an assumed level of measurement reliability η.

$$\rho^Y_{\text{Cousins}} = \beta_1^2 \rho^Y_{\text{Siblings}} + \beta_2^2 + 2\beta_1\beta_2\rho_1 \tag{49}$$

$$\rho^Y_{\text{Siblings}} = \beta_1^2 + \beta_2^2 + 2\beta_1\beta_2\rho_1 \tag{50}$$

Subtract 49 from 50

$$\rho^Y_{\text{Siblings}} - \rho^Y_{\text{Cousins}} = \left(1 - \rho^Y_{\text{Siblings}}\right)\beta_1^2 \tag{51}$$

$$\beta_1 = \pm\sqrt{\frac{\rho^Y_{\text{Siblings}} - \rho^Y_{\text{Cousins}}}{1 - \rho^Y_{\text{Siblings}}}} \tag{52}$$

$$= \pm\sqrt{\frac{\frac{1}{\eta^2}\rho_{\text{Siblings}} - \frac{1}{\eta^2}\rho_{\text{Cousins}}}{1 - \frac{1}{\eta^2}\rho_{\text{Siblings}}}} \tag{53}$$

Solve 50 for β_2

$$\rho^Y_{\text{Siblings}} = \beta_1^2 + \beta_2^2 + 2\beta_1\beta_2\rho_1 \tag{54}$$

$$= \beta_1^2 + \beta_2^2 + 2\beta_1^2\frac{\beta_2}{1-\beta_2} \tag{55}$$

$$(1-\beta_2)\rho^Y_{\text{Siblings}} = \beta_1^2(1-\beta_2) + \beta_2^2(1-\beta_2) + 2\beta_1^2\beta_2 \tag{56}$$

$$0 = \beta_1^2(1-\beta_2) + \beta_2^2(1-\beta_2) + 2\beta_1^2\beta_2 - (1-\beta_2)\rho^Y_{\text{Siblings}} \tag{57}$$

$$0 = -\beta_2^3 + \beta_2^2 + \left(\beta_1^2 + \rho^Y_{\text{Siblings}}\right)\beta_2 + \beta_1^2 - \rho^Y_{\text{Siblings}} \tag{58}$$

$$0 = \beta_2^3 - \beta_2^2 - \left(\beta_1^2 + \rho^Y_{\text{Siblings}}\right)\beta_2 - \beta_1^2 + \rho^Y_{\text{Siblings}} \tag{59}$$

$$0 = \beta_2^3 - \beta_2^2 - \left(\beta_1^2 + \frac{1}{\eta^2}\rho_{\text{Siblings}}\right)\beta_2 + \left(\frac{1}{\eta^2}\rho_{\text{Siblings}} - \beta_1^2\right) \tag{60}$$

D PSID sample selection

The original 1968 PSID sample contained two groups: the Survey of Economic Opportunity (SEO) sample and the Survey Research Center (SRC) sample. The SRC sample was a multistage probability sample of households in the contiguous 48 states in which every household had equal probability of selection. The SEO sample represents households residing in Standard Metropolitan Statistical Areas (SMSAs), with the exception of the South which included non-SMSAs, and was drawn from a sample used by the Census bureau to study economic opportunity among low-income households. A substantial share of the SEO sample was dropped in 1997 to reduce costs. In the SRC sample, over 98% of observations on grandchildren at ages 25-45 occur in 1997 or later, rendering the 1997 reduction in the SEO sample important for this particular application. Because the two samples are distinct and because it is much easier to assess the representativeness of the SRC sample, I treat the two samples separately in all analyses, reporting results for the SRC sample in the main text and the SEO sample in the appendix.

In addition to the samples that began in 1968, the PSID added refresher samples of Latino and immigrant families in 1990 and 1997, respectively. I exclude these samples from all analysis because they have not been followed long enough to produce sizable samples of grandchildren descended from the sampled families.

Not all families from the initial 1968 SRC sample have descendants who contribute to the analytic sample, reflecting a combination of fertility processes and survey attrition. As shown in Table A1, the analytic sample includes grandchildren descended from only 703 of the 3,000 households originally sampled. Some households initially sampled in 1968 were childless or had children who had already moved out of the home before the initial interview, and therefore were not followed. In the SRC sample, 998 households had no children in the household in 1968 despite the head and the wife (if applicable) being 40 or more years old. These families were likely either childless or had children who had already moved out of the household, leaving only 1,932 households that responded in 1968 and might plausibly have descendants who the PSID would attempt to follow. Given that 1,523 families have descendants who were actually interviewed, attrition appears to be comparably minor. The reduction in sample size between the 1,523 households with descendants and the 1,171 households with grandchildren likely reflects a combination of survey attrition and childlessness in the second generation. The reduction from 1,171 households with grandchildren to 703 households with grandchildren observed at ages 25-45 likely reflects a combination of survey attrition and delayed fertility such that some grandchildren were not yet 25 years old by the end of data collection. The overall reduction in the sample size from 3,000 sampled SRC families to an analytic sample of descendants from 703 families should give the reader pause about the representativeness of this sample because it suggests the possibility of non-random attrition. Given the arguments above, however, it is important to remember that factors other than attrition such as childlessness would produce a reduction even if survey response rates were perfect. Given that no other data exist to answer the research question, this paper proceeds cautiously with the understanding that the analytic sample may not accurately represent the population due to survey attrition.[4]

	SRC sample (main text)	SEO sample (appendix)
Sampled households	3,000	2,000
Interviewed in 1968	2,930	1,872
With descendants in the PSID	1,523	1,100
With grandchildren in the PSID	1,171	881
Analytic sample observed at ages 25-45:		
Extended families	703	365
Nuclear families	1,101	589
Persons	2,008	998
Observations	9,076	3,968

Table A1. PSID sample restrictions. The SRC sample represents a cross-section of American households in 1968, whereas the SEO sample includes low-income households selected through a more complex process. The main text reports estimates on the cross-sectional SRC sample. Estimates using the SEO sample are provided in the Appendix.

To assess the robustness of the results to the choice of sample, I estimated an alternative

[4] If one is willing to assume that attrition is ignorable given observed values, the PSID weights may address attrition and produce representative estimates. Unfortunately, the use of weights in Bayesian models is an area of active research with no clear answers (i.e. Si et al. 2015). The use of these models is critical in this application, because non-Bayesian models do not yield reliable uncertainty estimates (see Section 3.2). I therefore rely on unweighted models with the equal-probability SRC sample, producing estimates that may be biased only to the extent that attrition is non-random.

set of models using the SEO sample. To the extent to which results differ in the SEO vs. the SRC sample, it would suggest that multigenerational processes may differ among low-income urban families as opposed to American families in general. In agreement with the SRC sample results from the main text, Fig. A4 shows that cousin correlations in annual age-adjusted log family income in the SEO sample are slightly higher than expected given the sibling correlations, but that the difference disappears if one assumes measurement reliability of 0.84. The estimated sibling and cousin correlations are almost exactly the same as those from the SRC sample (Fig. 5), and the implied structural parameters (Fig. A5) and long-run pattern of mobility (Fig. A6) are likewise similar to those presented in the main text (Fig. 6 and 7). Similarly, estimates of sibling and cousin correlations in permanent income are very similar in the SEO sample (Fig. A7) to those reported for the SRC sample (Fig. 8). Overall, the choice of sample seems to be relatively unimportant.

E Supplemental figures

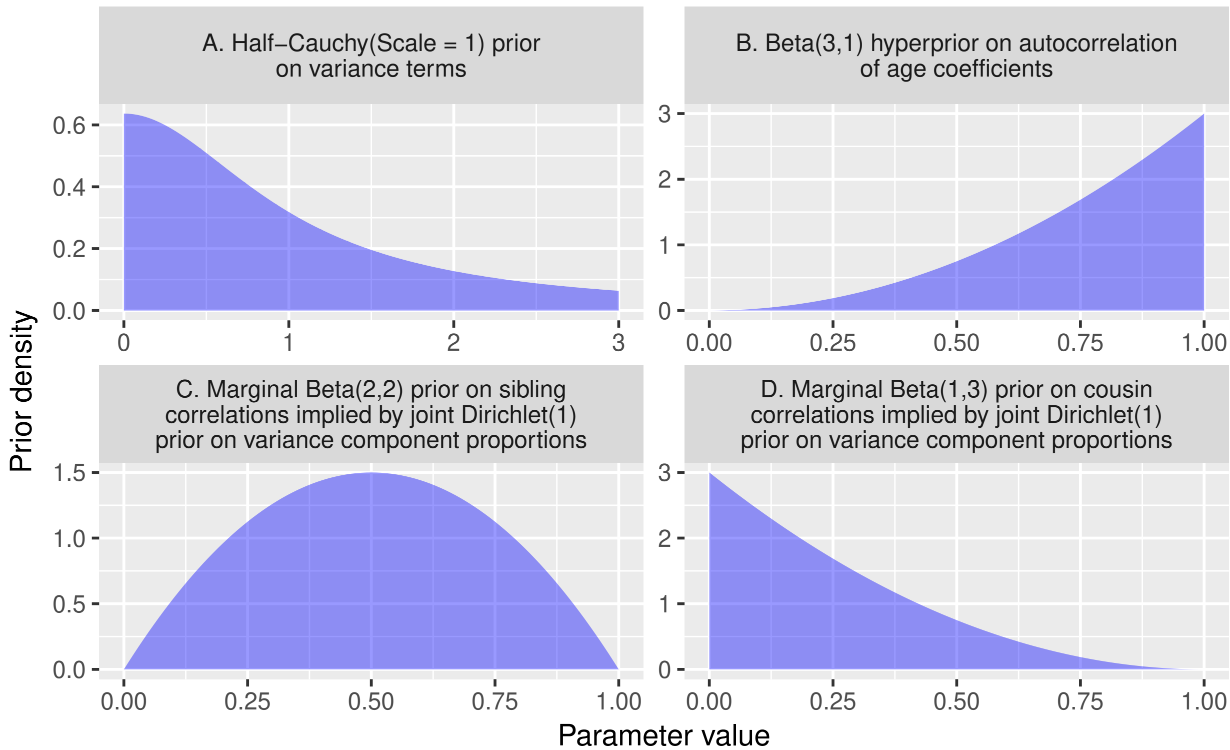

Fig. A1. Graphical depiction of prior distributions

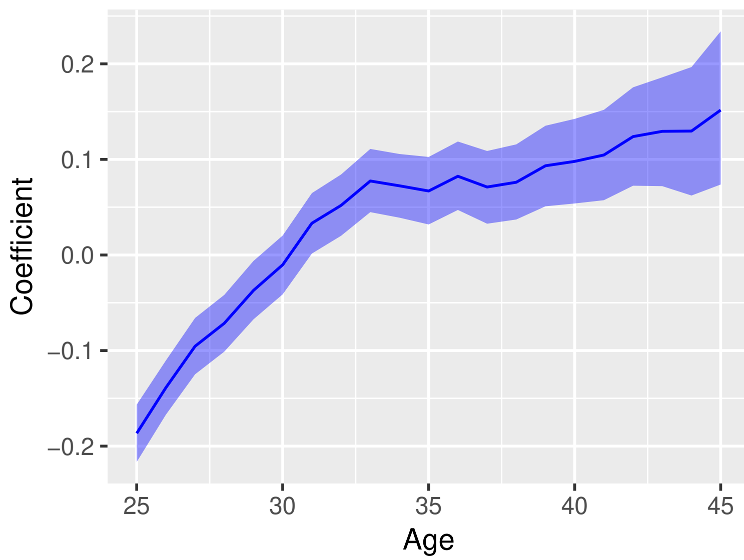

Fig. A2. Estimated association between age and log family income

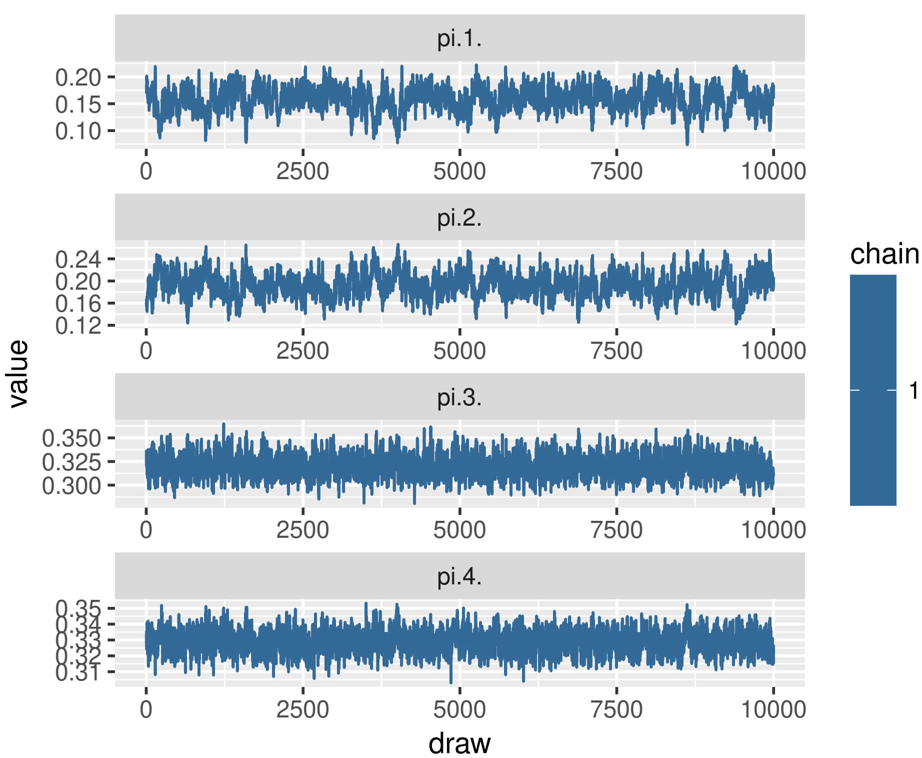

Fig. A3. Trace plots showing Hamiltonian Monte Carlo sampling for the proportion of the variance at each level for age-adjusted log family income.

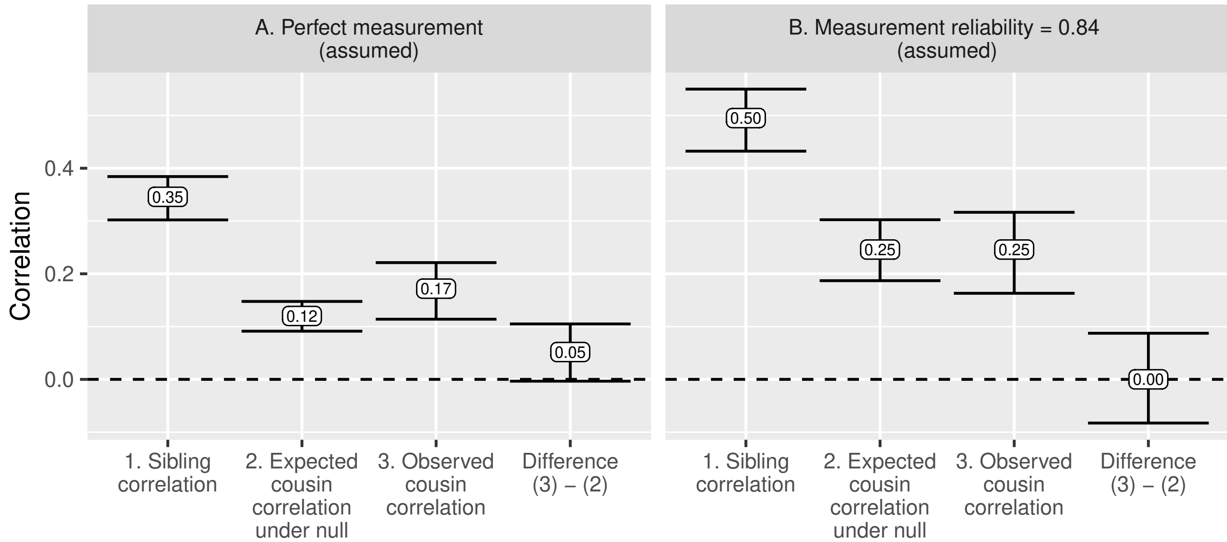

Fig. A4. SEO sample estimates of sibling and cousin correlations in log family income. Error bars represent 95% credible intervals. Analogous SRC sample estimates are provided in the main text, Fig. 5.

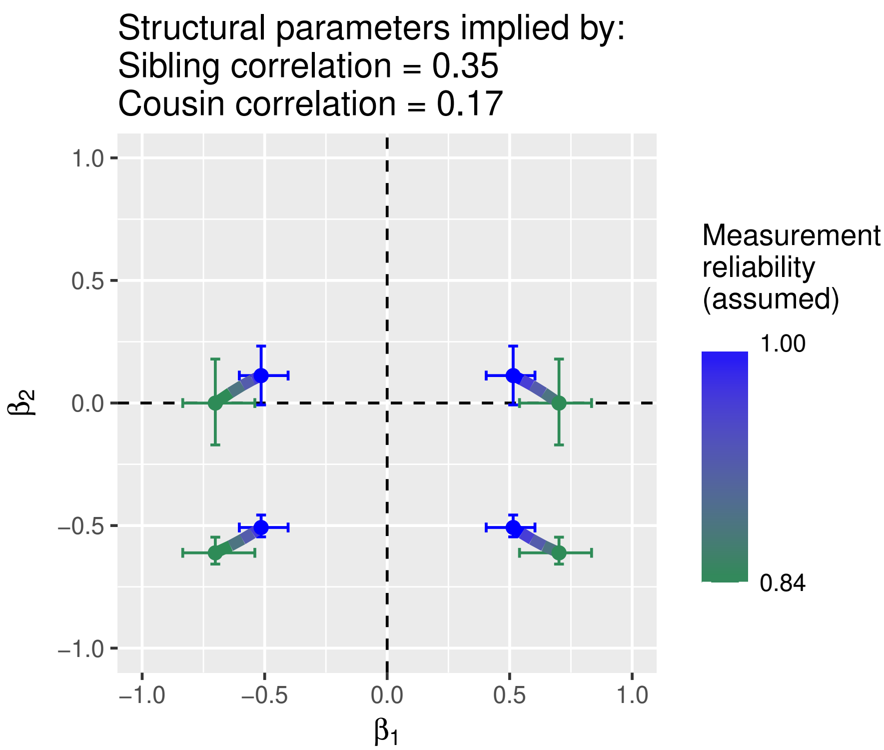

Fig. A5. SEO sample estimates of structural parameters for log family income transmission implied by sibling and cousin correlations under a second-order Markov model assuming the most positive β_1 and β_2 consistent with the observed data. Error bars represent 95% credible intervals. Analogous SRC sample estimates are provided in the main text, Fig. 6.

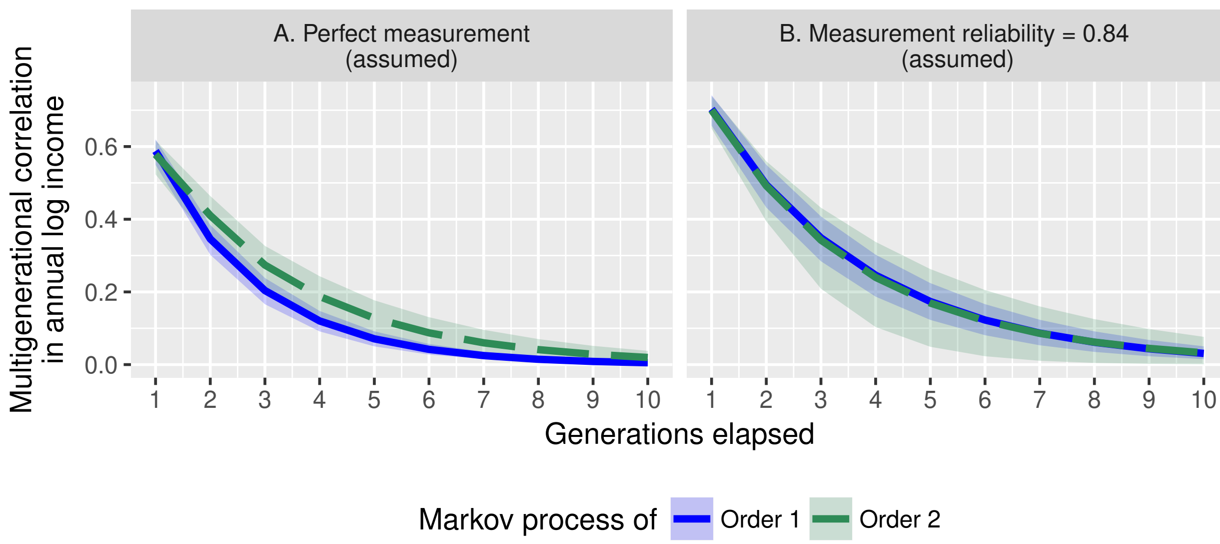

Fig. A6. SEO sample estimates of multigenerational correlations in log family income implied by sibling and cousin correlations under a second-order Markov model assuming the most positive β_1 and β_2 consistent with the observed data. Error bars represent 95% credible intervals. Analogous SRC sample estimates are provided in the main text, Fig. 7.

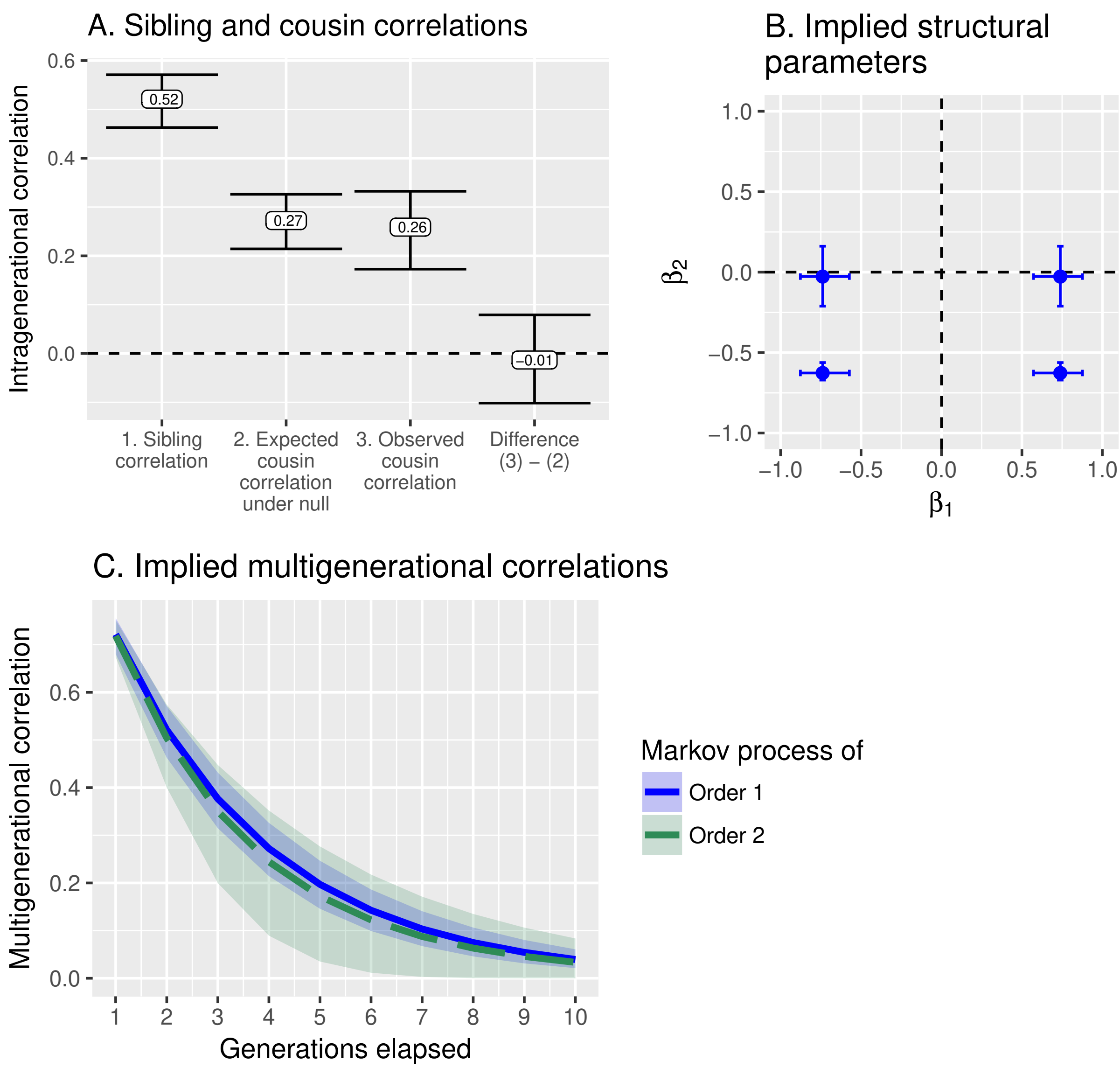

Fig. A7. SEO sample estimates of sibling and cousin correlations in permanent age-adjusted log family incomes and the implied mobility regimes with which they are compatible. Error bars and shaded bands represent 95% credible intervals. Analogous SRC sample estimates are provided in the main text, Fig. 8.