Does Opportunity Skip Generations? Reassessing Evidence From Sibling and Cousin Correlations

Ian Lundberg [1]

Abstract

Sibling (cousin) correlations are empirically straightforward: they capture the degree to which siblings' (cousins') socioeconomic outcomes are similar. At face value, these quantities seem to summarize something about how families constrain opportunity. Their meaning, however, is complicated. One empirical set of sibling and cousin correlations can be generated from a multitude of distinct theoretical processes. I illustrate this problem in the context of multigenerational mobility: the relationship between the incomes of an ancestor and a descendant separated by several generations in a family. When cousins' outcomes are similar (an empirical fact), prior authors have favored the particular theoretical interpretation that extended kin affect life chances through pathways not involving the parents of the focal individual. I show that this evidence is consistent with alternative theories of latent transmission (measurement error) or dynamic transmission (a parent-to-child transmission process that changes over generations). Theoretical assumptions are required to lend meaning to a point estimate. Further, I show that point estimates alone may be misleading because they can be highly uncertain. To facilitate uncertainty estimation for the key test statistic, I develop a Bayesian procedure to estimate sibling and cousin correlations. I conclude by outlining how future research might use sibling and cousin correlations as effective descriptive quantities while remaining cognizant that these quantities could arise from a variety of distinct theoretical processes.

✉ Ian Lundberg
  ilundberg@princeton.edu

[1] Department of Sociology and Office of Population Research, 227 Wallace Hall, Princeton University, Princeton, NJ 08544, USA

Introduction

Sibling and cousin correlations are a standard tool in the study of stratification and mobility. Suppose we construct a scatterplot with each pair of siblings represented by a point, with one sibling's income on the x-axis and the other's income on the y-axis. The correlation between the two is often taken as an omnibus measure of the degree to which family of origin constrains life chances (Solon et al. 1991). Noting that cousins share extended kin (e.g., grandparents) but come from different nuclear families, recent research has used sibling and cousin correlations in combination to assess the degree to which life chances vary across nuclear and extended families (Hällsten 2014; Jæger 2012; Knigge 2016).

The problem with sibling and cousin correlations is that they are difficult to interpret. Siblings can be similar because parents strongly influence life chances, because they are raised in the same neighborhood, because they influence each other's outcomes, or for numerous other possible reasons. Many theoretical stories are consistent with a given set of empirical estimates. The same is true of cousin correlations. Although it is tempting to interpret a large cousin correlation as evidence that extended families shape life chances, cousins' outcomes would also be similar under a parent-to-child transmission process in which transmission is strong from grandparents to parents and from parents to offspring. Echoing Stuhler's (2012) argument that parent-offspring correlations can arise from a variety of causal processes, the present article shows that sibling and cousin correlations can also arise from a variety of causal processes. Similar to the way biologists interpret similarities of traits in light of path models that summarize the generative process (Otto et al. 1995), a sibling correlation in social science also takes its meaning only under a theory about how that correlation came to be.

As a reference point, I begin with the Becker and Tomes (1979) model of income mobility. This model predicts that the cousin correlation will equal the squared sibling correlation. When empirical evidence contradicts this model, I show that the evidence is consistent with at least three distinct theoretical interpretations. Estimates would deviate from the Becker-Tomes prediction if the process is dynamic (changing across generations), latent (outcomes measured with error), or second-order (involving a direct role for grandparents). Some prior authors have favored the third interpretation, but there is no particular reason to do so.

Second, I apply a Bayesian estimation procedure to produce new estimates of sibling and cousin correlations in permanent log family income in the United States. The advantage of this procedure is that it facilitates reporting of uncertainty about the key test statistic and implied parameters of theoretical models. When estimates seem to contradict a theoretical model, it becomes essential to assess our degree of confidence by incorporating uncertainty.

The article concludes with a discussion of the role that sibling and cousin correlations may continue to play in applied research. I highlight how these quantities remain useful in a descriptive sense and may serve as a springboard for new theories, even if they cannot provide definitive evidence for or against a particular theory under assumptions we can credibly defend.

Sibling and Cousin Correlations: A Variance Decomposition

Sibling correlations capture the expected degree to which an outcome tends to be similar for a pair of siblings. Although these quantities are standard in stratification and mobility (Jencks et al. 1972; Mazumder 2008), this section briefly introduces them for the general reader. Suppose an outcome μ_i for person i is decomposed into three components:

$$
\underset{\substack{\text{Socioeconomic}\\\text{outcome}}}{\underbrace{\mu_i}} = \underset{\substack{\text{Component unique}\\\text{to this person}}}{\underbrace{\delta_i}} + \underset{\substack{\text{Component shared}\\\text{within sibling set}}}{\underbrace{\eta_{S(i)}}} + \underset{\substack{\text{Component shared}\\\text{within cousin set}}}{\underbrace{\lambda_{C(i)}}} .
$$

(1)

In the notation in Eq. (1), $S(i)$ indexes the sibling set to which i belongs: all siblings of i have the same value of $S(i)$. Likewise, $C(i)$ indexes the cousin set to which i belongs: all siblings and cousins of i have the same value of $C(i)$. Denote the variance of each component by $\left\{ \sigma^2_{Person}, \sigma^2_{SiblingSet}, \sigma^2_{CousinSet} \right\}$.

The relationship of Eq. (1) to sibling and cousin correlations relies critically on a set of independence assumptions that components are independent across units and across levels. To make this concrete, sibling correlations assume that the person-specific component is independent across individuals ($\delta_i \perp \delta_j \;\; \forall \;\; i, j$). This assumption would be violated if siblings differentiate from each other, thereby making δ_i and δ_j negatively correlated for a pair of siblings i and j. In the extreme, this would occur in a dystopian world in which every amount that sibling 1 sacrifices helps sibling 2 to succeed (Conley 2008). This dystopian world may seem implausible, but in the actual world siblings do in fact differentiate from each other to some degree (Conley 2004). The use of variance decompositions to study sibling correlations requires us to set aside this possibility. Likewise, estimators for sibling and cousin correlations typically assume independence across levels—for instance, ruling out that the average attainment of one's sibling or cousin set is related to the variance of attainment for individuals within that set.

Armed with the independence assumption, the correlation between μ for a pair of siblings i and j is a function of the variance at each level (Mazumder 2008):

$$
\rho_{Sibling} = \mathrm{Cor}\left(\mu_i^{Person}, \mu_j^{Person} \right)
$$

(2)

$$
= \frac{\mathrm{Cov}\left(\mu_i^{Person}, \mu_j^{Person} \right)}{\left(\mathrm{V}(\mu_i^{Person}) \mathrm{V}\left(\mu_j^{Person} \right) \right)^{\frac{1}{2}}}
$$

(3)

by definition of correlation,

$$
= \frac{\mathrm{Cov}\left(\delta_i + \eta_{S(i)} + \lambda_{C(i)}, \delta_j + \eta_{S(j)} + \lambda_{C(j)} \right)}{\left(\mathrm{V}(\mu_i^{Person}) \mathrm{V}\left(\mu_j^{Person} \right) \right)^{\frac{1}{2}}} .
$$

(4)

by Eq. (1).

Expand, dropping cross-level terms by independence:

$$
\underbrace{\overbrace{\mathrm{Cov}\left(\delta_i,\ \delta_j\right)}^{\substack{= 0 \text{ by} \\ \text{independence}}} + \overbrace{\mathrm{Cov}\left(\eta_{S(i)},\ \eta_{S(j)}\right)}^{\substack{= \sigma^2_{SiblingSet} \\ \text{given that } S(i)=S(j) \\ \text{for siblings}}} + \overbrace{\mathrm{Cov}\left(\lambda_{C(i)},\ \lambda_{C(j)}\right)}^{\substack{= \sigma^2_{CousinSet} \\ \text{given that } C(i)=C(j) \\ \text{for siblings}}}}{\left(\underbrace{\overbrace{V\left(\mu_i^{Person}\right)}}_{= \sigma^2_{Person} + \sigma^2_{SiblingSet} + \sigma^2_{CousinSet}} \quad \underbrace{\overbrace{V\left(\mu_j^{Person}\right)}}_{= \sigma^2_{Person} + \sigma^2_{SiblingSet} + \sigma^2_{CousinSet}}\right)^{\frac{1}{2}}}
\tag{5}
$$

$$
= \frac{\overbrace{\sigma^2_{SiblingSet} + \sigma^2_{CousinSet}}^{\substack{\text{Variance between} \\ \text{sibling sets}}}}{\underbrace{\sigma^2_{Person} + \sigma^2_{SiblingSet} + \sigma^2_{CousinSet}}_{\text{Total variance}}}.
\tag{6}
$$

Under the independence assumption, the proportion of the variance that is between sibling sets equals the sibling correlation. Analogously, the proportion of the variance between cousin sets equals the cousin correlation. For this reason, even descriptive use of sibling and cousin correlations as measures of similarity within families relies on the independence assumption. This article's primary focus is the *interpretation* of sibling and cousin correlations in light of theoretical models of social mobility. The next section formalizes theoretical models of multigenerational mobility using parameters (e.g., β) linking outcomes across multiple generations. Sibling and cousin correlations provide a tool to infer these parameters given data on the outcome generation alone. As I will argue, inferring multigenerational parameters from sibling and cousin correlations requires conceptual arguments for a particular theoretical model. It is often difficult to make a compelling case for one theory over another, which makes inferences about multigenerational processes from sibling and cousin correlations difficult.

Deviation From Becker-Tomes: Three Distinct Interpretations

A deep literature has built on the formal model of Becker and Tomes (1979), in which the expected income (Y_g) of someone in generation g given the incomes of all ancestors ($\{Y_{g'}\}_{g'<g}$) is a linear function β of the parent income (Y_{g-1}):

$$
\mathrm{E}\left(Y_g \Big| \{Y_{g'}\}_{g'<g}\right) = \beta Y_{g-1} \qquad \forall g.
\tag{7}
$$

This model forms the basis of the famous claim that families can go "from shirtsleeves to shirtsleeves in three generations" (Becker and Tomes 1986:S28), rising from poverty to luxury and back again. A one-unit increase in generation g income produces a β increase in the expected income of generation $g + 1$, but produces only a β^k increase k generations in the future. If we conceptualize each income measure (Y_g) as standardized to a mean of 0 and a variance of 1, the coefficient β can be interpreted as the correlation between the incomes of parents and offspring within a family. Figure 1 plots the rapid decay across generations that results for the case when $\beta = .50$. Despite a moderately high .50 correlation between the

incomes of parents and offspring, a great-great-grandparent's income is correlated only .06 with offspring income. In short, families have abundant opportunity to move away from their origins over the course of a few generations.

The prediction of substantial mobility over multiple generations holds only insofar as the assumptions of the Becker and Tomes (1979) model are correct. Three assumptions are noteworthy (Fig. 2). First, the assumed process is *observed rather than latent*: observed parent income (Y_{g-1}) is sufficient to block all statistical dependence between earlier ancestor incomes ($\{Y_{g'}\}_{g'<g}$) and offspring income (Y_g). Even if this holds for true income, income reported with error $\widetilde{Y}_{g-1}$ would not be sufficient to block this dependence (Ferrie et al. 2016). This is the sense in which the process is observed. Second, the assumed process is *static rather than dynamic*: the process that generates incomes in generation g is the same as the process that generates incomes in generation $g - 1$. This would be violated if the process of mobility changes over time. If the ability to buy advantages for one's offspring differs across generations, a single β cannot represent this relationship in all generations. Third, the assumed process is *first-order rather than second-order*: once the incomes of parents are known, the incomes of grandparents provide no additional information about the life chances of offspring. This would not be true, for instance, if grandparents directly affect life chances net of parents (Mare 2011; Pfeffer 2014).

Statistical evidence that contradicts the Becker-Tomes model may mean that any one of these assumptions does not hold. Sibling and cousin correlations offer one source of evidence against the model. In particular, the model predicts a sibling correlation of β^2 and a cousin correlation of β^4 (online appendix, part 1). Without directly estimating β, one can assess evidence against this model by a test statistic τ: the difference between the cousin correlation and the squared sibling correlation,

$$\tau = \rho_{Cousin} - \rho^2_{Sibling}, \tag{8}$$

where τ = 0 under the model of Becker and Tomes (1979). The following subsections show how this prediction does not hold under violations of each of the assumptions that the process is static, observed, and first-order.
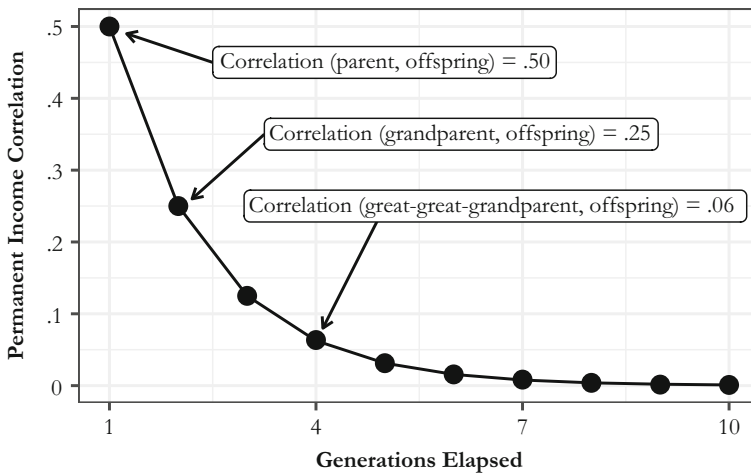

Fig. 1 Multigenerational permanent income correlations decay rapidly under the Becker-Tomes model. The attainment of a distant ancestor is only weakly associated with one's own attainment even if β is large (β = .5 depicted).

a Becker-Tomes

Grandparents   Parents   Offspring

b Dynamic

Grandparents   Parents   Offspring

c Latent

Grandparents   Parents   Offspring

d Second-order

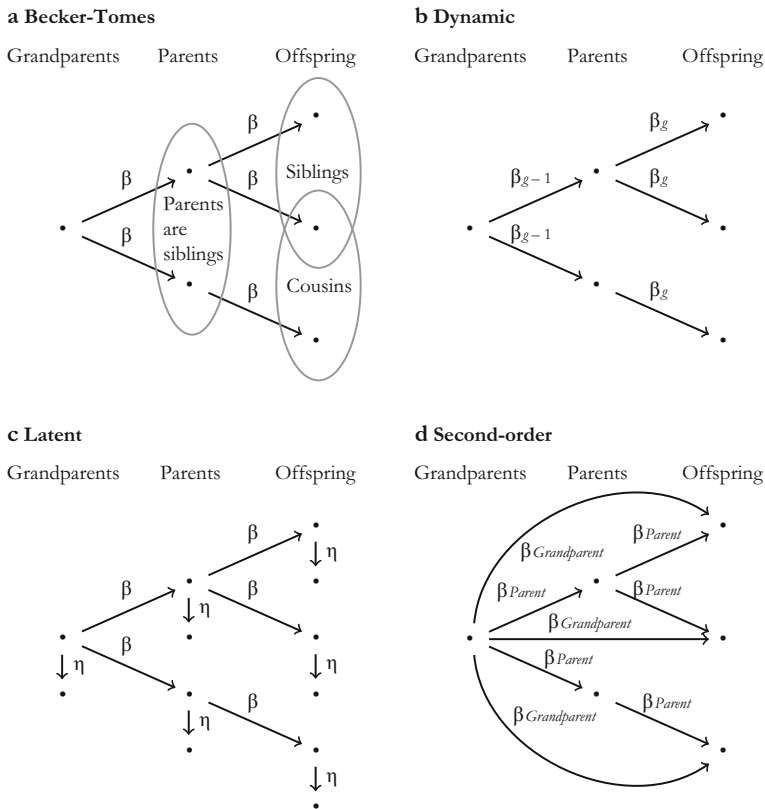Grandparents   Parents   Offspring

Fig. 2 Theoretical models of multigenerational mobility. Each model represents a variable in the transmission process. The figure depicts the outcomes of a pair of siblings and one of their cousins. Panel a is the Becker-Tomes model. Panel b allows the parent-offspring transmission coefficient to be dynamic across generations. Panel c maintains a static transmission parameter but adds a measurement step with reliability η. Panel d assumes a static process with perfect measurement but allows a direct grandparent-offspring link. Evidence against the Becker-Tomes model can be interpreted as evidence for any of the three alternatives or others not depicted.

Interpretation 1: From a Static to a Dynamic Process

Suppose the process is dynamic, so that the intergenerational income elasticity β_g is different in each generation g:

$$\mathrm{E}\left(Y_g \middle| \{Y_{g'}\}_{g'<g}\right) = \underbrace{\beta_g}_{\substack{\text{Dynamic across} \\ \text{generations } g}} Y_{g-1} \qquad \forall g. \tag{9}$$

In this case, a one-unit increase in parent income increases offspring income in generation g by β_g. A one-unit increase in grandparent income increases parent income in generation $g - 1$ by a different value of β_{g-1}. Combining these, a one-unit increase in grandparent income increases offspring income by $\beta_{g-1}\beta_g$. This could be greater or less than β_g^2 (the prediction of the static model) depending on

whether the grandparent-parent transmission, β_{g-1}, is greater or less than the parent-offspring transmission, β_g.

In the analogous intragenerational case, each sibling's income is correlated β_g with parent income, producing a correlation of β_g^2 between siblings (online appendix, part 1). Cousin's incomes are correlated $\beta_{g-1}\beta_g$ with the income of a shared grandparent, producing a cousin correlation of $\beta_{g-1}^2\beta_g^2$. Notably, this cousin correlation can be larger or smaller than the squared sibling correlation. Given an estimated sibling correlation, the Becker-Tomes model is consistent with only one cousin correlation, whereas a dynamic model is consistent with any cousin correlation.

$$\rho_{Sibling} = \beta_g^2. \tag{10}$$

$$\rho_{Cousin} = \beta_g^2\beta_{g-1}^2. \tag{11}$$

$$\rho_{Sibling}^2 < \rho_{Cousin} \text{ if } \beta_g < \beta_{g-1}. \tag{12}$$

$$\rho_{Sibling}^2 = \rho_{Cousin} \text{ if } \beta_g = \beta_{g-1}. \tag{13}$$

$$\rho_{Sibling}^2 > \rho_{Cousin} \text{ if } \beta_g > \beta_{g-1}. \tag{14}$$

Interpretation 2: From an Observed to a Latent Process

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 1968:144). For instance, whereas Blau and Duncan (1967) found substantial returns to education net of family background variables, Bowles (1972:S222) called this conclusion "seriously misleading" because of reporting error in family background variables that downwardly bias their association with the outcome and produce an overstatement of the returns to education. In studies of mobility over three generations, it is well known that credible estimates require measurement models for socioeconomic outcomes (Warren and Hauser 1997). Central conclusions in the field of mobility often hinge on assumptions about measurement.

Sibling and cousin correlations are no exception to this rule. As first noted in the context of sibling correlations (Solon et al. 1991), intragenerational correlations in measured attainment $(\widetilde{Y})$ will be less than correlations in true attainment (Y) to the extent that attainment is reported with classical measurement error. This occurs because measurement error makes siblings (and cousins) look more different than they actually

are.[1] Formally, suppose that reported income, $\widetilde{Y}$ (standardized to a mean of 0 and variance of 1), is a noisy indicator of true income, Y (also standardized to mean 0 and variance 1), with reliability factor η linking the two. Suppose that reported income is uncorrelated across individuals given actual income.

$$E\left(\widetilde{Y}_g \big| \{Y_{g'}\}_{g'<g}\right) = \underbrace{\eta Y_g}_{\substack{Y_g \text{ is latent} \\ \text{and measured} \\ \text{with error}}}. \tag{15}$$

$$\mathrm{Cov}\left(\widetilde{Y}_{gi}, \widetilde{Y}_{gj} \big| Y_{gi}, Y_{gj}\right) = 0. \tag{16}$$

Pairing this measurement model with the Becker-Tomes transmission model (Eq. (7)) yields the following sibling and cousin correlations.

$$\rho_{Sibling} = \eta^2 \beta^2. \tag{17}$$

$$\rho_{Cousin} = \eta^2 \beta^4. \tag{18}$$

$$\rho_{Sibling}^2 \leq \rho_{Cousin}. \tag{19}$$

In short, measurement error depresses the sibling and cousin correlations by a factor of $\eta^2 \leq 1$. Counterintuitively, the typical practice of squaring the sibling correlation also squares the contribution of measurement error, thereby driving the squared sibling correlation down to a greater degree than the cousin correlation, which is not squared. Because the central test statistic for the Becker-Tomes model involves a comparison of the squared sibling correlation and the cousin correlation, conclusions can be especially sensitive to the presence of measurement error. Even a shift from annual to permanent income could be sufficient to alter conclusions about multigenerational persistence.

Interpretation 3: From a First-Order to a Second-Order Process

Finally, evidence against the Becker-Tomes model can be interpreted in light of a second-order process under which the attainment of grandparents and offspring are directly linked through a path that does not involve the attainment of parents:[2]

$$E\left(Y_g \big| \{Y_{g'}\}_{g'<g}\right) = \beta_{Parent} Y_{g-1} + \underbrace{\beta_{Grandparent} Y_{g-2}}_{\substack{\text{Second−order process adds this} \\ \text{direct grandparent contribution}}} \quad \forall g. \tag{20}$$

---

[1] The measurement error that biases sibling and cousin correlations is error in the outcome variable. This makes the problem distinct from econometric work on the use of proxies for latent predictor variables (e.g., Lubotsky and Wittenberg 2006).

[2] I focus on a theoretical model that assumes only one income variable in each generation, even though a given person has two sets of grandparents, four sets of great-grandparents, and so on. This simplification is similar to the assumption in many demographic studies of a one-sex population. In the setting of multigenerational mobility, Becker and Tomes (1979:1183) motivated it as an assumption of perfect assortative mating on social origins; under this assumption, all the grandparents here have the same family incomes and can be represented by a single variable. Extending the Becker-Tomes model to allow some independence between social origins is beyond the scope of this article but may be an important topic for future work.

Numerous theories suggest that $\beta_{Grandparent} \neq 0$. Many children today see their grandparents frequently (Dunifon et al. 2018), sometimes living in doubled-up multigenerational households (Mykyta and Macartney 2011; Pilkauskas and Martinson 2014; Pilkauskas et al. 2014; Taylor et al. 2010). Prior research has documented substantial multigenerational associations in education (Sheppard and Monden 2018; Zeng and Xie 2014), occupation (Chan and Boliver 2013), and wealth (Adermon et al. 2018; Pfeffer and Killewald 2017). Extended kin may compensate for the shortcomings of parents, providing resources that parents cannot (Erola et al. 2018; Prix and Pfeffer 2017), or may serve as a cultural reference point that propels offspring toward academic achievement and attainment (Hällsten and Pfeffer 2017; Hertel and Groh-Samberg 2014). Although some evidence casts doubt on a direct role for grandparents in the attainment process (Warren and Hauser 1997), much of the literature suggests that one exists in at least some contexts. For a variety of reasons, we might expect grandparent income to be informative about offspring income net of parent income.

Correlations in a second-order process are mathematically complicated. To build intuition for this complexity, consider the case in which both parents and grandparents directly cause the incomes of offspring. The parent-offspring correlation then includes a causal component as well as components that arise through confounding (e.g., both are influenced by the grandparents and previous ancestors). The following equations show the sibling and cousin correlations generated by a second-order process (derivations in the online appendix, part 1):

$$
\rho_{Sibling} = \underbrace{\beta_{Parent}^2}_{\substack{\text{Direct path} \\ \text{through shared} \\ \text{parent}}} + \underbrace{\beta_{Grandparent}^2}_{\substack{\text{Direct path} \\ \text{through shared} \\ \text{grandparent}}} + \underbrace{\frac{2\beta_{Parent}^2 \beta_{Grandparent}}{1 - \beta_{Grandparent}}}_{\substack{\text{Other paths} \\ \text{through several} \\ \text{ancestors}}}. \tag{21}
$$

$$
\rho_{Cousin} = \underbrace{\beta_{Parent}^2 \rho_{Sibling}}_{\substack{\text{Path through} \\ \text{parents who} \\ \text{are siblings}}} + \underbrace{\beta_{Grandparent}^2}_{\substack{\text{Direct path} \\ \text{through shared} \\ \text{grandparent}}} + \underbrace{\frac{2\beta_{Parent}^2 \beta_{Grandparent}}{1 - \beta_{Grandparent}}}_{\substack{\text{Other paths} \\ \text{through several} \\ \text{ancestors}}}. \tag{22}
$$

For the present purpose, it is sufficient to note that the cousin correlation need not equal the squared sibling correlation in Eqs. (21) and (22).

Three Interpretations: What We Learn and Do Not Learn From Evidence

The sections above introduced the Becker-Tomes model as well as three theoretical extensions to processes that are dynamic, latent, or second-order. The Becker-Tomes model has only one unknown (β), but sibling and cousin correlations provide two sources of empirical information. This makes the Becker-Tomes model overidentified: the two

sources of empirical evidence can point toward different estimates of β and thus cast doubt on the validity of the model.[3] Because the extensions have more unknown parameters, they are consistent with a wider range of possible sibling and cousin correlations. Thus, in general, the data cannot cast doubt on these alternatives.[4] Conclusions that favor one interpretation over another require some other reason, such as additional data or a conceptual argument, to justify the preferred interpretation over the other candidates.

Reinterpreting Published Evidence

Table 1 presents a few published sibling and cousin correlations. In the text, I highlight one estimate from each study. Every author provided compelling evidence that casts doubt on the Becker-Tomes model, but the language they used to discuss this evidence often favors a second-order process in which grandparents directly influence offspring attainment. Yet, the evidence is often also consistent with a dynamic or latent process.

Analyzing educational attainment in the Wisconsin Longitudinal Study, Jæger (2012: table 2) estimated that 14% of the variance in years of education is between extended families, and 26% of the variance is between nuclear families within extended families. These estimates indicate a sibling correlation of $.14 + .26 = .41$ (with the sum reflecting rounding), which squares to predict a cousin correlation under the Becker-Tomes model of .17. The estimated cousin correlation is .14. This evidence is consistent with several theories. The process could be dynamic, with grandparent-parent income transmission of $\beta_{g-1} = .58$ and parent-offspring transmission of $\beta_g = .64$. It could also arise from a second-order model in which parent income raises offspring income by $\beta_{Parent} = .68$ and grandparent income directly reduces offspring income by a small amount, $\beta_{Grandparent} = -.06$. Jæger (2012:913) concluded, "a two-generation Markov process [parent-to-child transmission] does not represent the total effect of family background on educational success." The evidence does indeed cast doubt on a two-generation Markov process defined as the Becker-Tomes model, but it does not imply that the extended family plays any direct role in the attainment of offspring.

Hällsten (2014) referenced an estimate by Björklund and Jäntti (2012) that sibling's years of education are correlated .39 in Sweden. Squaring this yields a Becker-Tomes prediction that the cousin correlation would be .15. This is almost exactly equal within rounding to the .15 cousin correlation that Hällsten (2014: table 2) estimated with Swedish register data. This estimate is consistent with a static, first-order Markov process with intergenerational transmission by $\beta = .62$. Nevertheless, on the basis of the inability to explain away the cousin correlation with measured covariates, Hällsten (2014:31) concluded, "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 [parent-to-child transmission]" (emphasis in original). The robustness of the cousin correlation to controls for parental characteristics is an intriguing result; this represents one example where new external information may inform multigenerational theories beyond

---

[3] The Becker-Tomes model is overidentified if we assume that β > 0; because β enters formulas only in squared form, the data would be equally consistent with β < 0 of equal magnitude.
[4] Empirical evidence can sometimes cast doubt on the latent transmission model because one subspace of parameters (cousin correlations less than the squared sibling correlation) is impossible in this model.

Table 1 Published estimates of cousin correlations providing evidence against the Becker-Tomes model if the cousin correlation deviates from the squared sibling correlation

Author	Context	Outcome	Sibling Correlation	Squared Sibling Correlation	Cousin Correlation	Parameters Consistent With Estimates Under Three Alternative Models		
						Dynamic	Latent	Second-Order
Jæger (2012)	Wisconsin	Years of education	.41	.17	.14	$\{\beta_g = .64; \beta_{g-1} = .58\}$	None	$\{\beta_{Parent} = .68; \beta_{Grandparent} = -.06\}$
		Log odds of college completion	.56	.31	.21	$\{\beta_g = .75; \beta_{g-1} = .61\}$	None	$\{\beta_{Parent} = .89; \beta_{Grandparent} = -.22\}$
	NLSY-CYA[a]	PIAT math score	.38	.14	.23	$\{\beta_g = .62; \beta_{g-1} = .78\}$	$\{\beta = .78; \eta = .79\}$	$\{\beta_{Parent} = .49; \beta_{Grandparent} = .18\}$
		PIAT reading score	.37	.14	.19	$\{\beta_g = .61; \beta_{g-1} = .72\}$	$\{\beta = .72; \eta = .85\}$	$\{\beta_{Parent} = .53; \beta_{Grandparent} = .11\}$
		Log odds of high school completion	.37	.14	.14	$\{\beta_g = .61; \beta_{g-1} = .62\}$	$\{\beta = .62; \eta = .99\}$	$\{\beta_{Parent} = .60; \beta_{Grandparent} = .01\}$
		Years of education	.37	.14	.26	$\{\beta_g = .61; \beta_{g-1} = .84\}$	$\{\beta = .84; \eta = .73\}$	$\{\beta_{Parent} = .42; \beta_{Grandparent} = .26\}$
		Log odds of college completion	.45	.20	.29	$\{\beta_g = .67; \beta_{g-1} = .80\}$	$\{\beta = .80; \eta = .84\}$	$\{\beta_{Parent} = .54; \beta_{Grandparent} = .18\}$
Hällsten (2014)	Sweden	Years of education	.39[b]	.15	.15	$\{\beta_g = .62; \beta_{g-1} = .62\}$	None	$\{\beta_{Parent} = .63; \beta_{Grandparent} = .00\}$
		Occupational prestige	.29	.08	.11	$\{\beta_g = .54; \beta_{g-1} = .62\}$	$\{\beta = .62; \eta = .87\}$	$\{\beta_{Parent} = .50; \beta_{Grandparent} = .06\}$
		GPA (9th grade)	.51[b]	.26	.19	$\{\beta_g = .71; \beta_{g-1} = .61\}$	None	$\{\beta_{Parent} = .81; \beta_{Grandparent} = -.14\}$
		Cognitive ability	.47[b]	.22	.16	$\{\beta_g = .69; \beta_{g-1} = .58\}$	None	$\{\beta_{Parent} = .76; \beta_{Grandparent} = -.13\}$
Knigge (2016)	Netherlands	Occupation status	.50	.25	.32	$\{\beta_g = .71; \beta_{g-1} = .80\}$	$\{\beta = .80; \eta = .88\}$	$\{\beta_{Parent} = .60; \beta_{Grandparent} = .14\}$
Pfeffer et al. (2016)	United States	Wealth	.34	.12	.19	$\{\beta_g = .58; \beta_{g-1} = .75\}$	$\{\beta = .75; \eta = .78\}$	$\{\beta_{Parent} = .48; \beta_{Grandparent} = .16\}$

Note: Deviations can arise under a dynamic model (parameter values change across generations), a latent model (outcome is measured with error), or a second-order model (direct grandparent effects).

[a] Jæger (2012) noted that the NLSY-CYA sample was young and highly selected at the time of analysis, so results should be interpreted with caution.

[b] With the exception of occupational prestige, Hällsten (2014) did not directly estimate sibling correlations but cited those reported by Björklund and Jäntti (2012).

what we can learn from a sibling and cousin correlation alone, in a way not formalized in this article. The purpose of including this estimate here is simply to highlight that the sibling and cousin correlations alone are insufficient to provide evidence against a process of parent-to-child transmission.

Knigge (2016) studied occupational status attainment in the Netherlands in the late nineteenth and early twentieth century, estimating a sibling correlation of .50. Squaring the sibling correlation yields a predicted cousin correlation of .25, which is slightly smaller than the estimated cousin correlation of .32, as the author noted (Knigge 2016:1233). The conclusion that "these results are not congruent with the Markovian model" (Knigge 2016:1233) is thus true only if "the Markovian model" strictly means the static, observed, first-order Becker-Tomes model. The evidence is in fact consistent with a dynamic process (coefficients of .80 from grandparents to parents and .71 from parents to offspring) or a latent process (β = .80 and measurement reliability η = .88). Finally, it is also consistent with a view that the Markovian model is wrong and that the parent effect (β_{Parent} = .60) combines with a direct grandparent effect ($\beta_{Grandparent}$ = .14) to produce offspring income. It is not clear a priori which of these theories is most conceptually plausible.

Finally, using PSID data from the United States, Pfeffer et al. (2016) estimated a sibling correlation in log net worth of .34, which implies a cousin correlation of $.34^2$ = .12. The authors estimated a cousin correlation of .19, which is slightly higher than expected and warrants their 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" (p. 12). The authors made no claims about whether this multigenerational persistence violates parent-to-child transmission; the study is included here only because readers may suspect that wealth transmission in the United States is a context in which multigenerational transmission is especially strong. The estimate is consistent with any of the three models presented here: dynamic transmission (β_{g-1} = .75, β_g = .58), latent transmission (β = .75, η = .78), or a second-order process (β_{Parent} = .48, $\beta_{Grandparent}$ = .16).

Each of the aforementioned studies found additional evidence based on intergenerational correlations to support its 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: cousins' outcomes are in many settings more similar than would be expected under the Becker-Tomes model. However, the estimates do not necessarily imply the particular alternative model in which cousin similarities arise because of direct multigenerational effects. These estimates are instead consistent with a variety of alternative models, including processes that operate from parent to child but are dynamic or latent.

Empirical Tools to Estimate Uncertainty and Incorporate Repeated Measures

The preceding review comments solely on whether published *point* estimates diverge from the prediction of the Becker-Tomes model. This omits an entirely separate reason why estimates may be different than expected: estimation uncertainty. In this section, I develop a Bayesian estimation procedure for sibling and cousin correlations that

quantifies uncertainty about the test statistic (the cousin correlation minus the squared sibling correlation). By incorporating repeated measures, the estimator also addresses some concerns that arise from classical measurement error and life cycle bias, each of which could produce a situation of latent transmission if left unaddressed.

Proposed Model

Suppose we obtain data on the log family income, Y_{it}, experienced by each person i at various ages t from 25 to 45. I standardize income to a mean of 0 and standard deviation of .5 for analysis,[5] calculated by pooling all persons and years. Denote the population-average income at age t by β_t, and assume that each person's income is a normal draw around a person-specific mean $\mu_i^{Person} + \beta_t$ with variance $\sigma_{Observation}^2$:

$$Y_{it} \sim \mathcal{N} \left(\underbrace{\mu_i^{Person}}_{\substack{\text{Permanent} \\ \text{income}}} + \underbrace{\beta_t}_{\substack{\text{Population} \\ \text{mean at age } t}} , \underbrace{\sigma_{Observation}^2}_{\substack{\text{Observation-level} \\ \text{variance}}} \right). \quad (23)$$

The person-specific intercept μ_i^{Person} allows that each person's income may be shifted up or down by a fixed amount from the population average. The population mean β_t at each age t adjusts for the life cycle bias problem that some people are observed at ages when incomes tend to be lower or higher. The observation-level variance, $\sigma_{Observation}^2$, subsumes actual income volatility and classical measurement error. To examine sibling and cousin correlations in permanent income, μ_i^{Person}, I decompose the variance of this quantity into three components as formalized in Eq. (1): a person-specific component (δ_i), a sibling set component ($\eta_{S(i)}$), and a cousin set component ($\lambda_{C(i)}$). I assume a normal prior on the values of each component:

$$\delta_1, \cdots, \delta_{n_{Person}} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma_{Person}^2\right) \quad (24)$$

$$\eta_1, \cdots, \eta_{n_{SiblingSet}} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma_{SiblingSet}^2\right) \quad (25)$$

$$\lambda_1, \cdots, \lambda_{n_{CousinSet}} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \sigma_{CousinSet}^2\right). \quad (26)$$

The central task for estimation of sibling and cousin correlations is to estimate the values of the variance components $\left\{\sigma_{Person}^2, \sigma_{SiblingSet}^2, \sigma_{CousinSet}^2\right\}$ because these imply particular values of sibling and cousin correlations (see Eqs. (2)–(6)). I place priors on the variance components and the other unknown parameters to target the search for a solution to the range of reasonable values and provide the basis for fully Bayesian inference:

$$\vec{\beta} \sim \mathcal{N}\left(\vec{0}, \mathbf{I}\right) \quad (27)$$

---

[5] I use .5 to follow the recommendations of Gelman et al. (2008) to place continuous variables on a scale that works well with the default priors they proposed.

$$\left\{ \sigma_{Observation}, \sigma_{Person}, \sigma_{SiblingSet}, \sigma_{CousinSet} \right\} \overset{\text{iid}}{\sim} \text{half–Cauchy}\,(0, 1). \qquad (28)$$

The normal prior on the age trajectory shrinks all coefficients toward 0, similar to ridge regression; in this setting, though, the prior is vastly overwhelmed by the data because the population-mean income at each age can be estimated with considerable precision (see Fig. A4 in the online appendix). The half-Cauchy prior on the variance components is the absolute value of a Cauchy distribution, with scale parameter corresponding to the median of the distribution. It is very diffuse (see Fig. A3 in the online appendix) and is a standard choice of a weakly informative prior for variance parameters (Gelman et al. 2008).

Data: Panel Study of Income Dynamics

I fit the model to U.S. data from the Panel Study of Income Dynamics (PSID). The PSID selected a probability 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 United States in 1968 had an equal probability of selection (for a discussion of other samples, see the online appendix, part 2). The SRC sample initially selected 3,000 households, of whom 755 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 the online appendix, part 2). The analytic sample includes 11,730 observations on 2,363 respondents nested in 1,260 sibling sets in 755 cousin sets. A total of 1,854 respondents have at least one cousin in the sample, and 1,431 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 $5,000.

Posterior Sampling

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 for each of four chains executed sequentially from independent starting points, which produce a pooled set of 40,000 posterior draws.[6] Point estimates reported are posterior means. Credible intervals report the 5th and 95th percentiles of posterior draws. Trace plots for key parameters of interest are provided in Fig. A5 in the online appendix.

Results

Fig. 3 presents results. The estimated sibling correlation is .38 (95% confidence interval (CI): .32–.44). Squaring all posterior draws to predict the cousin correlation under the

---

[6] This procedure takes several hours on a Windows cluster computer with 512 GB of RAM and an Intel Xeon CPU E7-4850 v3 @2.20 GHz processor.
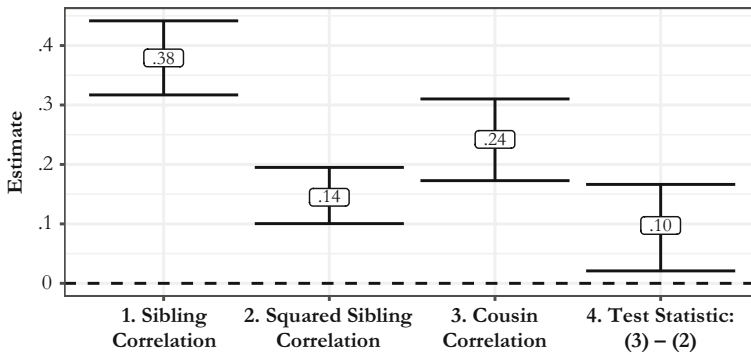
Fig. 3. The cousin correlation in permanent log family income is slightly greater than expected under the Becker-Tomes model. This empirical result does not point to any particular preferred alternative interpretation without additional assumptions. Error bars denote 95% posterior credible intervals.

Becker-Tomes model yields a posterior mean estimate of .14 (95% CI: .10–.20). The observed cousin correlation is slightly larger, at .24 (95% CI: .17–.31). The key test statistic subtracts the squared sibling correlation from the cousin correlation. For this test statistic, the estimate is .10 (95% CI: .02–.17). The cousin correlation therefore provides some evidence against the Becker-Tomes model for transmission of permanent income. Estimation uncertainty clarifies the strength of this evidence.

Rejection of the Becker-Tomes model (Fig. 3) leaves open the question of which alternative is most theoretically plausible. The results are consistent with at least three alternatives. The transmission process may be dynamic with grandparent-parent transmission $\beta_{g-1} = .80$ (95% CI: .67–.91) and parent-offspring transmission $\beta_g = .62$ (95% CI: .56–.66).[7] The use of permanent income removes the possibility that this result could be explained by classical measurement error at the observation level, but two alternative latent processes are possible. One is that individuals consistently over- or underreport their incomes in all waves of the survey, thereby producing classical measurement error at the person level with reliability η. A second is that the variable transmitted across generations is actually a latent variable distinct from income, and this variable is correlated with permanent income at only η. The observed data are consistent with either scenario, with $\beta = .80$ (95% CI: .67–.91) and measurement reliability $\eta = .78$ (95% CI: .64–.94). Finally, results are consistent with the interpretation favored by some prior researchers: a second-order process. In this case, parents affect offspring incomes by $\beta_{Parent} = .47$ (95% CI: .30–.61), and grandparents directly affect offspring incomes by $\beta_{Grandparent} = .21$ (95% CI: .04–.38). Assessing the relative merits of these three theories would require conceptual argument; the data do not point toward any one theory over the others.

---

[7] The slightly higher coefficient required for grandparent-parent transmission is consistent with a supplemental result. Following a procedure analogous to the main model, I estimate the parent-generation sibling correlation to be .51 (95% CI: .46–.55), slightly higher than the estimate of .38 (95% CI: .32–.44) in the offspring generation.

Comparison With Prior Approaches

This section briefly reviews what has been gained by the Bayesian statistical model of this article relative to other approaches. Under some assumptions, the model solves three key problems: uncertainty, life cycle bias, and measurement error.

Quantification of uncertainty is the primary advantage of the Bayesian approach. The key test statistic is a complex function of the variance components of a model, involving division, squaring, and subtraction. Posterior samples make uncertainty straightforward even in this setting: pass each posterior sample through the transformation, and then summarize the distribution of posterior samples that result. It is possible to arrive at similar uncertainty estimates using frequentist methods, but doing so is mathematically complicated. Standard software packages (e.g., *nlme*; Pinheiro et al. 2019) estimate the frequentist variance of a transformation of the variance of random intercepts at each level of the model, but mathematical derivations are required to convert these to estimates for the variance of sibling and cousin correlations (online appendix, part 3). It is therefore understandable that prior research has generally omitted uncertainty intervals on key quantities. A Bayesian model provides one solution to this problem.

The model also conducts an adjustment for age that addresses some concerns about life cycle bias: the tendency of substantive conclusions in mobility research to be sensitive to the age at which income is measured (Haider and Solon 2006). The problem arises because incomes are not stable over the life course and because not all individuals report incomes at exactly the same ages. Sibling correlations might appear large simply because the age at which incomes are reported is more similar within than between sibling sets. This source of similarity is a nuisance; sibling similarities that arise because of similarities in the age of measurement are not typically of interest for substantive theories about mobility. The proposed model corrects for this problem to the degree that every individual's income trajectory is an intercept shift μ_i^{Person} away from the population-average income trajectory $\overrightarrow{\beta}$. If we observe incomes at ages 25–30 for one pair of siblings and ages 35–40 for another pair of siblings, the population-average trajectory upwardly adjusts the estimate of μ_{Person}^i in the former case and downwardly adjusts μ_{Person}^i in the latter case.

The model does not correct for life cycle bias if income trajectories follow different patterns across individuals. If, for instance, individuals with higher long-term earnings also experience steeper earnings profiles, as in the example of Haider and Solon (2006), then a homogeneous population-average trajectory will not correct for this problem. Nonetheless, the problem is not likely to be severe in this example because the reporting ages do not vary dramatically within the analytic sample: across individuals, the interquartile range of the mean age of reporting is 27.0 to 31.4. Similarity in reporting ages across individuals means that, even if the age adjustment is incorrect, results are likely to approximate the sibling and cousin correlation around the ages at which incomes are reported in the sample. The model is especially promising if we benchmark against the most standard approach of averaging income over whatever ages are observed for each individual within a specified age window (e.g., Björklund and Jäntti 2012). Incorporating the age trajectory directly into the model is an

improvement over this benchmark because it explicitly adjusts for age differences across individuals in the sample.

Net of life cycle bias, the model addresses one type of measurement error by estimating the observation-level residual variance, $\sigma^2_{Observation}$, and omitting this component from calculation of sibling and cousin correlations. The information to estimate $\sigma^2_{Observation}$ comes from the fact that each person is observed an average of five times, with an interquartile range from two to seven. If only one observation per person were available, then sibling and cousin correlations would be unable to distinguish the person-specific permanent variance component, σ^2_{Person}, from the observation-level variance component, $\sigma^2_{Observation}$. Pulling these apart avoids a problem in which random observation-level reporting error would drive down sibling and cousin correlations (Solon et al. 1991). The model does not resolve problems of nonclassical measurement error. If individuals tend to over- or underreport their incomes in the same direction over time in a way independent of all other individuals, this would inflate σ^2_{Person} and drive down sibling and cousin correlations. This might happen, for instance, if respondents seek to make themselves look good on the survey by overreporting income whenever they are asked. The tendency to overreport may be correlated among siblings, for instance, if their parents instilled an especially high value on presenting oneself as materially successful. This would inflate $\sigma^2_{SiblingSet}$, thereby increasing sibling correlations and reducing cousin correlations. Finally, cousin correlations would be higher if a tendency to overreport is correlated among cousins, thereby inflating $\sigma^2_{CousinSet}$. Addressing these types of nonclassical measurement error would likely require new forms of data about how individuals respond to surveys and whether response errors are correlated within families. I leave these tasks to future research. For the present article, a focus on permanent income resolves measurement error only if this error is classical at the level of observations.

Discussion

Sibling and cousin correlations in socioeconomic outcomes serve a useful descriptive purpose, summarizing the degree to which outcomes are similar for those born into the same family. Knowledge of these quantities and their variation across outcomes, over time, and under various institutional arrangements can spur new theories about how these correlations come to be. Sibling and cousin correlations serve as descriptive facts that force us to look at the situation in a new light and develop hypotheses about the relationship between family origins and socioeconomic opportunity.

Distinguishing between various theories, however, is difficult because the range of conceptual processes that could produce sibling and cousin correlations involves far more parameters than the amount of data available. This is in some sense unsurprising: data in one generation can tell us about a process occurring in past generations only under incredibly strong assumptions. To interpret a variance decomposition in terms of sibling and cousin correlations requires assumptions about independence across units. To further interpret those correlations as evidence of a dynamic, latent, or second-order process requires a conceptual argument against the other interpretations equally consistent with the observed data. To conclude that extended kin shape life chances requires assumptions that outcomes have been measured precisely and that the

transmission process has unfolded with the same parameter values over the past two generations. A credible defense of these assumptions is difficult to build. For this reason, caution is warranted when using sibling and cousin correlations to draw multigenerational inferences.

Beyond sibling and cousin correlations, the claims of this article relate to a growing body of evidence that various indicators of socioeconomic attainment persist over three or more generations to a greater degree than the Becker-Tomes model would predict (Anderson et al. 2018; Chan and Boliver 2013; Colagrossi et al. 2019; Hällsten and Pfeffer 2017; Hertel and Groh-Samberg 2014; Huang et al. 2015; Jæger 2012; Knigge 2016; Lindahl et al. 2015; Liu 2018; Olivetti et al. 2016; Pfeffer 2014; Pfeffer and Killewald 2015, 2017; Wightman and Danziger 2014; Ziefle 2016). This body of work clearly demonstrates that decades of past research have overstated societal openness by leaving complexities out of the model. Already, others have pointed out that evidence of grandparent-offspring associations conditional on parent characteristics can arise from various processes (Braun and Stuhler 2016; Breen 2018; Ferrie et al. 2016; Solon 2018; Stuhler 2012). I join these authors in arguing that evidence from multigenerational and intragenerational correlations remains unable to pinpoint the precise complexity that should be added to the model. The interpretation favored by some authors (a direct role for grandparents) is only one candidate among many possibilities that could make society more rigid than Becker and Tomes (1979) predicted. One model that is wrong opens the door to numerous new theories that might be more consistent with the observed data.

Proper interpretation of evidence is especially important in the domain of multigenerational mobility because of the theoretical positions at stake. As an extreme example, Clark (2014) noted that in many societies, certain surnames tend to have high average attainment across many generations and concluded that genetic factors must be responsible for this high degree of persistence. Torche and Corvalan (2018) thoroughly critiqued this argument by clarifying that the persistence of group-level attainment (averaged over all those with a given surname) is distinct from the persistence of individual-level attainment, so that the former does not imply the latter. This level of clarity about evidence is obviously essential when a researcher claims that genetics underly mobility processes. I argue that researchers must take a similarly critical view when assessing the evidence on any theory of multigenerational transmission. Sociologists and demographers may be drawn to theories that predict strong socioeconomic persistence because individuals are embedded in complex family networks that constrain opportunity socially from many directions. Supporting such a strong and complex theory, however, requires an accurate statement of the evidence in support of this theory. Researchers must devote at least as much thought to more mundane alternatives, such as measurement error, which could produce the empirical finding.

To apply this lens critically to the work presented here, I cannot rule out direct effects of extended kin on life chances. The actual process by which income is transmitted over generations is likely to be far more complicated than the three simple models considered here. All I can credibly conclude is that the estimated sibling and cousin correlations could have been generated by a dynamic, latent, or second-order process, or something else that I have not considered. With confidence, I can conclude that the evidence does not point cleanly toward any one of these as the correct interpretation. Sibling and cousin correlations provide a useful point of reference, but

in the case of multigenerational mobility, the answers they provide are far from definitive.

Compliance With Ethical Standards

Ethics and Consent   This study conducts secondary analysis of existing data collected for research purposes and distributed under the oversight of the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board.

Conflict of Interest   The author declares no conflict of interest.

References

Adermon, A., Lindahl, M., & Waldenström, D. (2018). Intergenerational wealth mobility and the role of inheritance: Evidence from multiple generations. *Economic Journal, 128,* F482–F513.

Anderson, L. R., Sheppard, P., & Monden, C. W. (2018). Grandparent effects on educational outcomes: A systematic review. *Sociological Science, 5,* 114–142. https://doi.org/10.15195/v5.a6

Becker, G. S., & Tomes, N. (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy, 87,* 1153–1189.

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

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

Blau, P. M., & Duncan, O. D. (1967). *The American occupational structure.* New York, NY: Wiley.

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

Braun, S. T., & Stuhler, J. (2016). The transmission of inequality across multiple generations: Testing recent theories with evidence from Germany. *Economic Journal, 128,* 576–611.

Breen, R. (2018). Some methodological problems in the study of multigenerational mobility. *European Sociological Review, 34,* 603–611.

Chan, T. W., & Boliver, V. (2013). The grandparents effect in social mobility: Evidence from British birth cohort studies. *American Sociological Review, 78,* 662–678.

Clark, G. (2014). *The son also rises: Surnames and the history of social mobility.* Princeton, NJ: Princeton University Press.

Colagrossi, M., d'Hombres, B., & Schnepf, S. V. (2019). *Like (grand)parent, like child? Multigenerational mobility across the EU* (IZA Discussion Paper No. 12302). Bonn, Germany: Institute for the Study of Labor.

Conley, D. (2004). *The pecking order: Which siblings succeed and why.* New York, NY: Pantheon.

Conley, D. (2008). Bringing sibling differences in: Enlarging our understanding of the transmission of advantage in families. In A. Lareau & D. Conley (Eds.), *Social class: How does it work?* (pp. 179–200). New York, NY: Russell Sage Foundation.

Dunifon, R. E., Near, C. E., & Ziol-Guest, K. M. (2018). Backup parents, playmates, friends: Grandparents' time with grandchildren. *Journal of Marriage and Family, 80,* 752–767.

Erola, J., Kilpi-Jakonen, E., Prix, I., & Lehti, H. (2018). Resource compensation from the extended family: Grandparents, aunts, and uncles in Finland and the United States. *European Sociological Review, 34,* 348–364.

Ferrie, J., Massey, C., & Rothbaum, J. (2016). *Do grandparents and great-grandparents matter? Multigenerational mobility in the US, 1910–2013* (NBER Working Paper No. 22635). Cambridge, MA: National Bureau of Economic Research.

Gelman, A., Jakulin, A., Pittau, M. G., & Su, Y.-S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics, 2,* 1360–1383.

Haider, S., & Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings. *American Economic Review, 96,* 1308–1320.

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., & Pfeffer, F. T. (2017). Grand advantage: Family wealth and grandchildren's educational achievement in Sweden. *American Sociological Review, 82,* 328–360.

Hertel, F. R., & Groh-Samberg, O. (2014). Class mobility across three generations in the U.S. and Germany. *Research in Social Stratification and Mobility, 35,* 35–52.

Huang, J. Y., Gavin, A. R., Richardson, T. S., Rowhani-Rahbar, A., Siscovick, D. S., & Enquobahrie, D. A. (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,* 568–578.

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

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., . . . Michelson, S. (1972). *Inequality: A reassessment of the effect of family and schooling in America.* New York, NY: Basic Books.

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

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

Liu, H. (2018). Social and genetic pathways in multigenerational transmission of educational attainment. *American Sociological Review, 83,* 278–304.

Lubotsky, D., & Wittenberg, M. (2006). Interpretation of regressions with multiple proxies. *Review of Economics and Statistics, 88,* 549–562.

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

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

Mykyta, L., & Macartney, S. (2011). *The effects of recession on household composition: "Doubling up" and economic well-being* (SEHSD Working Paper No. 2011-4). Washington, DC: U.S. Census Bureau, Social and Economic Housing Statistics Division.

Olivetti, C., Paserman, M. D., & Salisbury, L. (2016). *Three-generation mobility in the United States, 1850–1940: The role of maternal and paternal grandparents* (NBER Working Paper No. 22094). Cambridge, MA: National Bureau of Economic Research.

Otto, S. P., Christiansen, F. B., & Feldman, M. W. (1995). *Genetic and cultural inheritance of continuous traits* (Paper No. 64). Stanford, CA: Stanford University, Morrison Institute for Population and Resource Studies.

Pfeffer, F. T. (2014). Multigenerational approaches to social mobility: A multifaceted research agenda. *Research in Social Stratification and Mobility, 35,* 1–12.

Pfeffer, F. T., & Killewald, A. (2015). *How rigid is the wealth structure and why? Inter- and multigenerational associations in family wealth* (PSC Report 15-845). Ann Arbor, MI: Population Studies Center.

Pfeffer, F. T., & Killewald, A. (2017). Generations of advantage: Multigenerational correlations in family wealth. *Social Forces, 96,* 1411–1442.

Pfeffer, F. T., Killewald, A., & Siliunas, A. (2016, June). *The concentration of wealth within family lineages and intergenerational transfers.* Paper presented at the New Directions in the Study of Intergenerational Transfers and Time Use in Later Life: A Panel Study of Income Dynamics Conference, Ann Arbor, MI. Retrieved from https://psidonline.isr.umich.edu/Publications/Workshops/IntergenTransfers/Papers.aspx

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

Pilkauskas, N. V., & Martinson, M. L. (2014). Three-generation family households in early childhood: Comparisons between the United States, the United Kingdom, and Australia. *Demographic Research, 30,* 1639–1652. https://doi.org/10.4054/DemRes.2014.30.60

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2019). *nlme: Linear and nonlinear mixed effects models. R package version 3.1-141* [Code base]. https://cran.r-project.org/web/packages/nlme/index.html

Prix, I., & Pfeffer, F. T. (2017). Does Donald need Uncle Scrooge? Extended family wealth and children's educational attainment in the United States. In J. Erola & E. Kilpi-Jakonen (Eds.), *Social inequality across the generations: The role of compensation and multiplication in resource accumulation* (pp. 112–135). Cheltenham, UK: Edward Elgar.

Sheppard, P., & Monden, C. (2018). The additive advantage of having educated grandfathers for children's education: Evidence from a cross-national sample in Europe. *European Sociological Review, 34,* 365–380.

Siegel, P. M., & Hodge, R. W. (1968). A causal approach to the study of measurement error. In H. M. Blalock Jr. & A. B. Blalock (Eds.), *Methodology in Social Research* (pp. 28–59). New York, NY: McGraw-Hill.

Solon, G. (2018). What do we know so far about multigenerational mobility? *Economic Journal, 128,* F340–F352.

Solon, G., Corcoran, M., Gordon, R., & Laren, D. (1991). A longitudinal analysis of sibling correlations in economic status. *Journal of Human Resources, 26,* 509–534.

Stan Development Team. (2017). RStan: The R interface to Stan (R package version 2.16.2). Retrieved from https://mc-stan.org/

Stuhler, J. (2012). *Mobility across multiple generations: The iterated regression fallacy* (IZA Discussion Paper No. 7072). Bonn, Germany: Institute for the Study of Labor.

Taylor, P., Passel, J., Fry, R., Morin, R., Wang, W., Velasco, G., & Dockterman, D. (2010). *The return of the multi-generational family household* (Social & Demographic Trends Report). Washington, DC: Pew Research Center.

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

Warren, J. R., & Hauser, R. M. (1997). Social stratification across three generations: New evidence from the Wisconsin Longitudinal Study. *American Sociological Review, 62,* 561–572.

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

Zeng, Z., & Xie, Y. (2014). The effects of grandparents on children's schooling: Evidence from rural China. *Demography, 51,* 599–617.

Ziefle, A. (2016). Persistent educational advantage across three generations: Empirical evidence for Germany. *Sociological Science, 3,* 1077–1102. https://doi.org/10.15195/v3.a47