Multigenerational Inequality*

Jan Stuhler
Universidad Carlos III de Madrid

October 2022

Abstract

A growing literature provides evidence on *multi*generational inequality – the extent to which socio-economic advantages persist across three or more generations. This chapter reviews its main findings and implications. Most studies find that inequality is more persistent than a naive iteration of conventional parent-child correlations would suggest. We discuss three potential interpretations of this new "fact", related to (i) latent, (ii) non-Markovian and (iii) non-linear transmission processes, and empirical strategies to discriminate between them. Finally, we illustrate why strong multigenerational associations imply strong assortative mating.

Keywords: Multigenerational inequality, assortative mating, distant kins

1 Introduction

Studies of social mobility often focus on two generations, measuring how one's education, income or other outcomes are associated with that of one's parents. But how persistent are socio-economic inequalities in the long run, across multiple generations? A naive extrapolation from the available parent-child evidence would suggest that multigenerational inequality is low – that the influence of family background declines geometrically, and therefore washes out over three or four generations. However, a growing literature links families over multiple generations to provide direct evidence on *multigenerational inequality*. This chapter reviews this evidence and its implications.

We first illustrate why multigenerational regression coefficients deviate from the product of the corresponding parent-child coefficients (Section 2), even though the "iteration" of coefficients may appear natural in a regression framework. Indeed, this "*iterated regression fallacy*" has been a common source of misinterpretations in intergenerational research and other contexts. The intuition for why conventional parent-child correlations may not be very informative about multigenerational inequality is that those correlations measure only a descriptive rather than a structural relationship.

*We are grateful for helpful comments and suggestions from Adrian Adermon, Jo Blanden, Lola Collado, Maximilian Longmuir, Martin Nybom and Salvatore di Salvo. Financial support from the Ministerio de Ciencia e Innovación (RYC2019-027614-I) and Comunidad de Madrid (EPUC3M11 and H2019/HUM-5891) is gratefully acknowledged.

We then review recent evidence on multigenerational inequality (Section 3). A robust pattern across studies is that inequality is more persistent than a naive iteration of the parent-child correlations would suggest. Put differently, the coefficient on grandparents in a child-parent-grandparent regression tends to be positive: grandparent status predicts child status, even conditional on parent status. Part of this "excess persistence" appears to reflect the omission of important characteristics of the parent generation. A particularly vital omission is the second parent, and controlling for both maternal and paternal characteristics greatly decreases the coefficient on grandparent status.

Less clear is how this new "empirical fact" should be interpreted. We highlight three potential interpretations related to (i) latent, (ii) non-Markovian and (iii) non-linear transmission processes (Section 4). Some parental influences are inherently unobservable, and such latent processes could also rationalise high multigenerational persistence. An alternative interpretation is that "grandparents matter" in a causal sense, influencing the status of their grandchildren independently of the characteristics of the parent generation. A third possibility is that intergenerational processes are non-linear or heterogeneous across families. While such non-linearities have received much attention in other contexts, its implications for multigenerational inequality have received little consideration.

We also link this evidence to earlier work on sibling correlations, which suggests that parental characteristics such as schooling or income account only for a minor part of the family and community influences that siblings share (Björklund and Salvanes, 2011; Jäntti and Jenkins, 2014). Intuitively, as family background cannot be captured up by one single variable, intergenerational associations capture only the "the tip of the iceberg" of family background effects (Björklund and Jäntti, 2012). We argue that multigenerational associations may reflect this same insight, making those associations meaningful even if their absolute size remains limited.

One intriguing implication of strong multigenerational associations is that assortative mating must be strong, too (Section 5). Conventional measures of assortative mating, such as the spousal correlation in years of schooling, would imply rapid regression to the mean across generations – which appears at odds with multigenerational estimates. This observation illustrates that multigenerational associations are not only informative about inequality in the very long run, but also provide novel insights into fundamental aspects of *inter*generational transmission, and the degree of sorting in a population.

For brevity, this chapter omits many important questions. The stylised models we consider are "mechanical", abstracting from economic choices and behavioural patterns. We also ignore normative aspects, such as the question of whether multigenerational correlations are "too high", and abstain from a discussion of policy implications. Instead, this chapter focuses on basic empirical "facts" and their potential interpretations, which might inform future work. It is complementary to a review by Anderson, Sheppard and Monden (2018). While they provide a comprehensive review of earlier multigenerational estimates, this chapter focuses more on their interpretation and relation to alternative measures of family background. Other insightful discussions of multigenerational mobility include Pfeffer (2014) and Solon (2018). Moreover, Blanden, Doepke and Stuhler (2022) discuss how multigenerational mechanisms contribute to educational inequality, and Section 4 of this chapter draws on an earlier review by Stuhler (2012).

2 The Iterated Regression Fallacy

Our understanding of intergenerational processes has been shaped by theoretical and empirical research involving just two generations, parents and children (Mare 2011). It is therefore instructive to first consider why an extrapolation from the available parent-child evidence may not be very informative about the persistence of socio-economic status across multiple generations (in the "long run").

The degree of status persistence between *parents* and their children is often measured by the slope coefficient in a linear regression of outcome y in offspring generation t of family i on the parental outcome in generation $t-1$,

$$y_{it} = \alpha + \beta_{-1} y_{it-1} + \varepsilon_{it}. \tag{1}$$

For example, if y is the logarithm of income then β_{-1} captures the *intergenerational elasticity of income*; a high elasticity represents low mobility. For simplicity we assume below that β_{-1} remains constant across generations, but the arguments apply likewise in a non-stationary environment.

How does the coefficient from this Galtonian regression across two generations compare with the coefficient across three or more generations? The idea that the latter equals the square of the former, so that persistence declines geometrically, may appear as a natural consequence of regression: if β_{-1} captures to what degree deviations from the mean tend to be passed from parents to children then we might expect $(\beta_{-1})^2$ to represent their expected extent after being passed twice from parents to children, between grandparents and their grandchildren. Formally, we may use equation (1) to rewrite the grandparent-grandchild elasticity β_{-2} as

$$\beta_{-2} \equiv \frac{Cov(y_{it}, y_{it-2})}{Var(y_{it-2})} = \frac{Cov(\beta_{-1} y_{it-1} + \varepsilon_{it}, y_{it-2})}{Var(y_{it-2})} \stackrel{?}{=} (\beta_{-1})^2. \tag{2}$$

The fallacy is in the last step: while ε_{it} is by construction uncorrelated to y_{it-1}, it is not necessarily uncorrelated with grandparental status y_{it-2}. Put differently, the coefficient β_{-1} in equation (1) captures a statistical rather than a structural association.[1]

The belief that regression toward the mean between two observations implies iterated regression between multiple observations is a common fallacy. This "*iterated regression fallacy*" (Stuhler, 2012) has not only affected beliefs about long-run mobility, but has also been common in other fields; using the term "expectation fallacy", Nesselroade, Stigler and Baltes (1980) discuss its relevance in psychological research. It also has a long tradition: Francis Galton fell fault of it in his influential work on linear regression (Bulmer, 2003). As discussed in the next section, a naive iteration of parent-child correlations tends to instead understate the true extent of multigenerational inequality.

3 Multigenerational Inequality

How persistent are socio-economic inequalities? A string of recent studies provide a new perspective on this question by tracking such inequalities across multiple generations. Spurred by the increased availability of suitable datasets, this surge of research on multigenerational inequality has occurred nearly simultaneously in economics (e.g., Lindahl et al., 2015), sociology (e.g., Chan and Boliver

[1] Importantly, equation (2) may fail to hold even in a Markovian world, in which outcomes depend only on the previous generation (Mare, 2011). We illustrate this argument in Section 4.

2013), demography (e.g., Mare 2011), and economic history (e.g., Dribe and Helgertz 2016). However, different studies have emphasised different interpretations, a point to which we return in the next section. Anderson, Sheppard and Monden (2018) provide a systematic review of earlier studies in this recent literature, and is complementary to the more selective presentation here.

3.1 Measuring multigenerational inequality

Multigenerational evidence tends to be presented in one of two distinct forms. We may compare the relative size of inter- and multigenerational correlations, i.e. whether multigenerational correlations are larger or smaller than the naive iteration of parent-child correlations would suggest,

$$\beta_{-k} \lesseqqgtr (\beta_{-1})^k \tag{3}$$

where β_{-k} is the multigenerational correlation between generation t and generation $t-k$, for $k > 1$. Alternatively we may estimate a multivariate regression of the form

$$y_{it} = \alpha + \beta_p y_{it-1} + \beta_{gp} y_{it-2} + ... + \varepsilon_{it}. \tag{4}$$

and study the sign and magnitude of the slope coefficient on grandparents or earlier ancestors.[2] The distinction is presentational: using the Frisch-Waugh-Lovell theorem, the coefficient β_{gp} in the three-generation regression (4) can be re-expressed as (see Braun and Stuhler, 2018)

$$\beta_{gp} = \frac{\beta_{-2} - (\beta_{-1})^2}{1 - (\beta_{-1})^2}, \tag{5}$$

such that we have "excess persistence" in the sense of $\beta_{gp} > 0$ if and only if $\beta_{-2} > (\beta_{-1})^2$. Given this "duality", there is no substantive difference between studies providing bivariate estimates, as in (3), and those focusing on multivariate estimates, as in (4).

3.2 Multigenerational evidence

Until a recent surge, little multigenerational evidence has been available. Hodge (1966) notes that mobility may not be well described by a first-oder Markov process in which child outcomes depend only on the parent generation. Warren and Hauser (1997) study three generations of families in the Wisconsin Longitudinal Study, showing that the occupational status of grandparents is not very predictive of their grandchildren's education or occupational status once father's education, occupation and earnings, and mother's education, are all controlled for. However, their estimates are not very precise due to the limited size of their sample. Similarly, Erola and Moisio (2007) report that conditional on parents' class, the grandchildren's social class is almost conditionally independent from the grandparents' class in Finland, while Chan and Boliver (2013) find a more robust net effect of grandparents' class in British data.[3]

[2] One interesting observation is that the addition of grandparents or other ancestors often contributes little in an R^2 sense, even if the corresponding slope coefficients are large. We return to this observation in Section 4.5.

[3] See also Warren and Hauser (1997) and Hertel and Groh-Samberg (2013), who review other contributions from the earlier literature.

Table 1: Selected Multigenerational Studies

Study	Sample	Main outcomes	Excess persistence	Remarks
Warren and Hauser (1997)	US (Wisconsin)	Education, Occupation	No, cond. on parent characteristics	Conditioning test
Lindahl et al. (2015)	Sweden (Malmö region)	Education, Income	Yes, $\beta_{-k} > \beta^k$ and $\beta_{gp} > 0$	Link up to four generations for education
Adermon, Lindahl and Waldenström (2018)	Sweden	Wealth	Yes, $\beta_{gp} > 0$	Study role of bequests and wealth, earnings and education up to four generations
Braun and Stuhler (2018)	Germany	Education, Occupation	Yes, $\beta_{-k} > \beta^k$ and $\beta_{gp} > 0$	Estimate latent factor model, conditioning tests, up to four generations for education
Pfeffer and Killewald (2018)	US (PSID)	Wealth	Yes, $\beta_{-k} > \beta^k$ and $\beta_{gp} > 0$	Study role of bequests, educational attainment, marriage, homeownership, and business ownership
Sheppard and Monden (2018)	21 countries (SHARE)	Education	Yes, $\beta_{gp} > 0$	Test for grandparent, contact and interaction effects
Neidhöfer and Stockhausen (2019)	US, UK and Germany	Education	Yes, $\beta_{-k} > \beta^k$ and $\beta_{gp} > 0$ (UK and Germany)	Conditioning tests, multigenerational trends, latent factor model
Colagrossi, d'Hombres and Schnepf (2020)	28 EU countries (Eurobarometer)	Education, Occupation	Yes, $\beta_{-k} > \beta^k$ and $\beta_{gp} > 0$ (most countries)	Conditioning tests, latent factor model
Engzell, Mood and Jonsson (2020)	Sweden	Income	Unconditional: $\beta_{gp} > 0$ Conditional $\beta_{gp} \approx 0$	Detailed conditioning tests, heterogeneity of multigenerational associations, data quality
Hällsten and Kolk (2020)	Sweden (Skellefteå and Umeå regions)	Education, Occupation, Wealth	Yes, $\beta_{-k} > \beta^k$ and $\beta_{gp} > 0$	Up to seven generations, 5^{th}-order cousins
Modalsli (2021)	Norway	Occupation, Income	Yes, $\beta_{-k} > \beta^k$ and $\beta_{gp} > 0$	Test for contact effects, multigenerational trends

Notes: The table lists a subset of studies with multigenerational family links. Conditioning tests correspond to estimates of β_{gp} in equation (4) with increasingly large set of parental controls.

Lindahl et al. (2015) combine survey data from the Swedish *Malmö study* with administrative data to track earnings for three generations and educational attainment over four generations. They find that multigenerational persistence is much higher than would be predicted from the iteration of regression estimates for two generations, i.e. $\beta_{-k} > (\beta_{-1})^k$ and $\beta_{gp} > 0$ in the three-generation regression (4). The size of the coefficient on grandparents' (standardised) earnings is about one quarter of the corresponding coefficient on father's earnings, which is also the median ratio of β_{gp}/β_p across 40 research articles reviewed by Anderson, Sheppard and Monden (2018). The results by Lindahl et al. received much attention. Although restricted to one Swedish region, the quality of the data was high, which also contained a good measure of earnings. And interestingly, their findings are at odds with a well-known prediction by Becker and Tomes (1986) that "*Almost all earnings advantages and disadvantages of ancestors are wiped out in three generations. Poverty would not [...] persist for several generations.*"

Braun and Stuhler (2018) study multigenerational correlations in education and occupational prestige, using several German samples. They interpret those correlations through the lens of a latent factor model in which parents transmit advantages to their children that are not directly observed in the data (see Section 4.1). The implied parent–child correlation in "latent" advantages is about 0.6, nearly 50 percent larger than the parent-child correlation in years of schooling in their samples. Applying a similar approach on harmonised survey data, Neidhöfer and Stockhausen (2019) find slightly higher latent persistence around 0.7 for Germany and slightly lower persistence rates for the US and UK, while Colagrossi, d'Hombres and Schnepf (2020) estimate a mean rate of latent persistence of 0.66 in standardised educational outcomes across 28 European countries.

Similar patterns are found for other outcomes. Using Swedish data, Adermon, Lindahl and Waldenström (2018) show that grandparental wealth is predictive of grandchildren's wealth, above and beyond parent wealth.[4] This pattern is even more pronounced in studies by Boserup, Kopczuk and Kreiner (2013) for Denmark and Pfeffer and Killewald (2018) for the US. While it could be partially explained by direct bequests from grandparents to their grandchildren that "skip a generation" (Mare, 2011), the advantages associated with family wealth appear to arise at an earlier age that such direct bequests would imply. Pfeffer and Killewald (2018) find that grandparent wealth predicts the education and home ownership of their grandchildren, conditional on parental wealth, and Hällsten and Pfeffer (2017) document substantial associations with their grandchildren's school grades.

Some recent studies are able to link more than "just" three generations. In particular, Hällsten and Kolk, 2020, link administrative data and parish records from Northern Sweden to track up to seven generations. The observation of such long data coverage would also allow researchers to estimate *trends* in multigenerational persistence. For example, Modalsli (2021) links up to five generations of data in Norway, and finds substantial differences in the strength of multigenerational persistence over time. As yet there exists little evidence on multigenerational correlations in developing countries, with Razzu and Wambile (2020), Kundu and Sen (2021) and Gallegos and Celhay (2022) as recent exceptions.

In summary, most studies find that multigenerational correlations are larger than a naive iteration of parent-child estimates would suggest. This pattern appears to be robust across countries and holds for different socio-economic outcomes. However, a substantial share of this "excess persistence" can be explained by the omission of the second parent, and studies that control for both maternal and

[4] Hällsten and Pfeffer (2017) show that net of parental wealth, grandparental wealth also predicts their grandchildren's educational achievement. Clark and Cummins (2014) find much higher persistence of wealth on the surname level, which can be rationalised using the latent factor model presented in Section 4.1.

paternal characteristics (e.g., Engzell, Mood and Jonsson, 2020) tend to find a much smaller and sometimes insignificant coefficient β_{gp} in the child-parent-grandparent regression (4). Whether significant grandparent associations remain after controlling for both maternal and paternal characteristics differs across studies. Moreover, there exists only limited evidence on whether multigenerational persistence varies across countries or groups.[5] Most importantly, there is no consensus yet on how this multigenerational evidence should be interpreted. While some studies emphasise the potential role of "latent" transmission channels, others emphasise the potential causal influence of grandparents or the extended family. We discuss potential interpretations in Section 4.

One common issue in multigenerational studies is that the marginal distributions tend to be very different for distant ancestors. Educational attainment is often low, income rarely observed, and even occupational classifications can be problematic as the share of farmers tends to be high in older generations. Moreover, very different mechanisms can generate similar "vertical" transmission patterns (Cavalli-Sforza and Feldman, 1981), making it difficult to distinguish between competing models. One alternative explored in recent studies is to also consider distant relatives in the "horizontal" dimension. Adermon, Lindahl and Palme (2021) measure "dynastic human capital" based on a broad set of kinships in the parent generation, including uncles and aunts, and show that it has a much stronger association with child education than conventional parental measures. And Collado, Ortuño-Ortín and Stuhler (2022) show that a single transmission model can fit both vertical and horizontal kinship patterns.

While this chapter focuses on studies that use *direct* family links across generations, these studies also link to recent *name*-based evidence by Clark (2014), Barone and Mocetti (2020) and others. Names are informative about multigenerational persistence for two distinct reasons. The more obvious one is that using names, we can link very distant generations. For example, Barone and Mocetti (2020) find that the average status of surnames correlates across five centuries, which suggests that some components of the transmission process must exhibit high persistence. The second, more subtle reason is that the regression of surname averages between two generations might tell us something about aspects of the *individual* transmission process that is not visible from individual-level regressions Clark (2014), Clark and Cummins (2014). We return to this observation in the next section.

4 Interpretations and Mechanisms

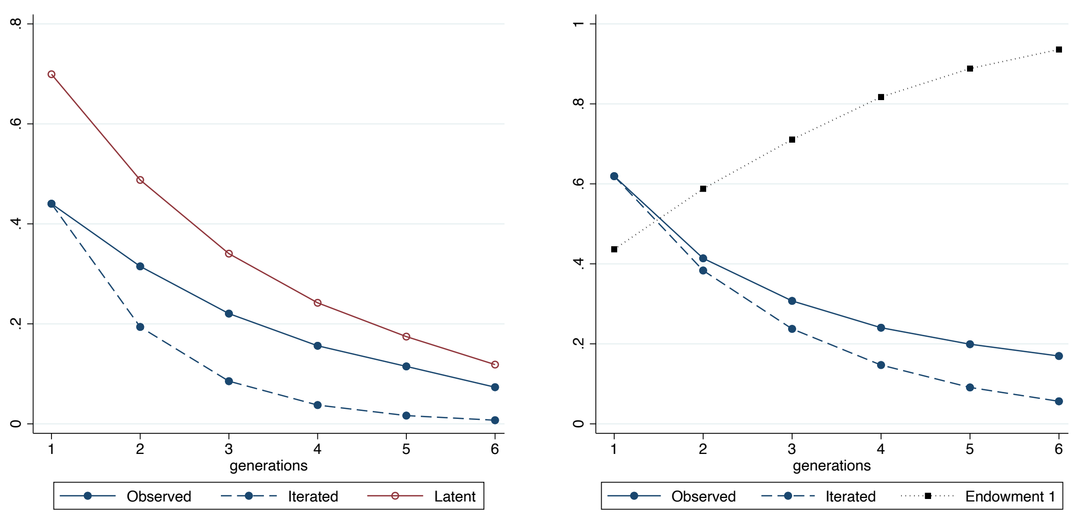
How should this multigenerational evidence be interpreted? While informative about the extent of multigenerational inequality, it is not directly informative about the underlying mechanisms: different theoretical models could generate similar multigenerational pattern. In this section, I review three potential interpretations related to (i) latent, (ii) non-Markovian and (iii) non-linear transmission processes.[6]

[5] Cross-country comparisons by Neidhöfer and Stockhausen (2019), Colagrossi, d'Hombres and Schnepf (2020) or GDIM (2018) do find such variation. However, due to limited sample sizes, some of the cross-country differences cannot be precisely estimated. Open questions for future research include whether cross-country differences in parent-child correlations (e.g., Blanden, 2013) are also reliable indicators of differences in multigenerational inequality.

[6] The presentation in this section draws and extends on Stuhler (2012). For brevity we focus on purely "mechanical" transmission models, which can however be viewed as the reduced form of behavioural or economic models (Goldberger, 1989). Moreover, the interpretations proposed here are of course not exhaustive; see for example Hertel and Groh-Samberg (2013), who discuss the potential role of war-related destruction and expulsions or racial and ethnic segregation.

Figure 1: Multigenerational pattern in different models

(a) A latent factor model (b) A model with multiplicity



Notes: Simulated data with n=10,000 observations. Panel (a) corresponds to the latent factor model with returns $\rho = 0.8$ and transferability $\lambda = 0.7$. The solid blue (red) line corresponds to the rate of persistence in outcome y (latent endowment e). The dashed line corresponds to the predicted persistence in y based on the iteration of the parent-child correlation. Panel (b) corresponds to the multiplicity model with $\rho_1^2 = 0.3$, $\rho_2^2 = 0.7$, $\lambda_1 = 0.9$ and $\lambda_2 = 0.5$. The black dashed line corresponds to the share of the first endowment in overall persistence.

4.1 Latent transmission processes

Some of the advantages that parents transmit to their children may be inherently unobservable, as has long been recognised in both the social sciences (e.g., Duncan 1969, Goldberger 1972, Becker and Tomes, 1979) and population genetics (Rice, Cloninger and Reich 1978, Cavalli-Sforza and Feldman 1981). As noted by Clark (2014), Clark and Cummins (2014) or Stuhler (2012), such "latent" transmission has also interesting implications for the pattern of multigenerational transmission

To understand the basic argument, consider a simplified one-parent one-offspring family structure, in which the transmission in generation t of family i is governed by

$$y_{it} = \rho e_{it} + u_{it} \tag{6}$$

$$e_{it} = \lambda e_{it-1} + v_{it}, \tag{7}$$

in which an observed outcome y depends on latent endowments e (according to *returns* ρ), which are partially transmitted within families (according to *transferability* λ), and where the white-noise error terms u and v represent market and endowment luck, uncorrelated with each other and past values. To simplify the presentation we drop the i subscript and assume that e and y are standardised with mean zero and variance one, such that slopes ρ and λ can be interpreted as correlations.

To make matters concrete assume that our outcome of interest is income, and that e measures an individual's human capital, although the argument can be applied in other contexts. The parameter ρ then measures the fraction of income that is explained by an individual's own human capital, as opposed to factors or events outside of individual control, such as market luck or market-level shocks; and $\rho = 1$

would imply that income differences are fully explained by an individuals' own characteristics.

Given equations (6) and (7), and the assumption that all variances are one, the intergenerational elasticity equals

$$\begin{aligned}\beta_{-1} &= Cov(y_t, y_{t-1}) \\ &= \rho^2\lambda, \end{aligned} \tag{8}$$

and the elasticity across three generations instead equals

$$\begin{aligned}\beta_{-2} &= Cov(y_t, y_{t-2}) \\ &= \rho^2\lambda^2. \end{aligned} \tag{9}$$

The extrapolation error from the iteration of the parent-child elasticity equals

$$\begin{aligned}\Delta &= (\beta_{-1})^2 - \beta_{-2} \\ &= \left(\rho^2 - 1\right)\rho^2\lambda^2, \end{aligned} \tag{10}$$

which is negative if $0 < \rho < 1$, that is as long as income is not perfectly determined by human capital.

Figure 1a provides a numerical example based on simulated data from this model, with $\rho = 0.8$ and $\lambda = 0.7$, implying an intergenerational correlation of $\beta_{-1} = \rho^2\lambda \approx 0.45$.[7] A naive iteration of this parent-child correlation across multiple generations would imply rapid regression to the mean (dashed line); after only three generations, the iterated correlation falls below 0.1, such that the distribution of income among ancestors explains less than one percent of the variance in their descendants' income. But the *actual* correlation in income in this model decays much more slowly (solid blue line), falling below 0.1 after only six generations. The reason is the strong transmission of human capital (red line), which serves as the actual state variable in this model.

The key idea underlying this "latent factor model" is that the true transmission mechanisms are distinct from the status y observed by the researcher.[8] But the representation of this idea in equations (6) and (7) is sufficiently generic to nest several distinct interpretations, with diverging implications. One possible interpretation is that y corresponds in fact to the "true" socio-economic status of an individual, but status is transmitted not directly but indirectly via other pathways. For example, income y may be a good proxy for status, but parents transmit not income but human capital e to their children. In this interpretation, β_{-1} is in fact a truthful measure of status persistence between one generation and the next – it just happens to be not a good measure of persistence in the long run.

Alternatively, we may assume that y is only a coarse proxy of socio-economic status, while e is the "true" or "generalised socio-economic status" of a person (as in Clark, 2014). For example, y may be short-run income, while e may represent a broader measure of socio-economic success. In this interpretation, β_{-1} is not only an inappropriate measure of multigenerational persistence; it is not even a good measure of the *inter*generational persistence of status differences from parents to their

[7]The distributions u and v are chosen such that e and y are normally distributed with mean zero and variance one.

[8]This simple model suggests not only that naive iterations understate persistence ($\Delta < 0$), but also that this extrapolation error Δ will be particularly large when ρ is small, i.e. when the observable outcome y is not a good proxy for the latent endowments e. The gap between inter- and multigenerational correlations should therefore be particularly large when considering outcomes that are hard to measure, such as cognitive or non-cognitive skills (e.g., Anger and Schnitzlein, 2017).

children (which would instead be captured by the red line in Figure 1a). One important special case of this interpretation is that the outcome y contains measurement error (Solon 2014, Ferrie, Massey and Rothbaum 2020).

The basic proposition underlying the latent factor model is intuitive, and also consistent with earlier insights from the literature. In particular, it is consistent with the argument that sibling correlations are a more comprehensive measure of family background effects than intergenerational correlations, as they capture the influence of all the advantages that siblings share, not only the advantages encapsulated by parental income or education (Björklund and Salvanes, 2011). More generally, it is consistent with the argument that conventional intergenerational correlations just measure the "tip of the iceberg" of family background effects (Björklund and Jäntti, 2012). However, not all models with latent transmission mechanisms generate high multigenerational persistence. In particular, the Becker-Tomes model in which child outcomes y also depend on parental income can produce either high or low multigenerational persistence (in the sense of $\beta_{-k} \lesseqgtr (\beta_{-1})^k$), depending on parameter values (Stuhler, 2012).[9]

4.2 Non-Markovian and extended-family processes

A second possibility is that intergenerational transmission deviates from a Markovian process, in the sense that earlier ancestors have an independent influence over and above the influence of parents. In particular, grandparents might have a direct causal influence on their grandchildren. This possibility has already been considered by Warren and Hauser (1997), but the role of the wider family has received renewed attention after Mare (2011) published his eloquent criticism of the "two-generation paradigm" in intergenerational research (e.g., Chan and Boliver 2013, Hertel and Groh-Samberg 2013, Ferrie, Massey and Rothbaum 2020).

For illustration, assume that offspring human capital depends on both parents and grandparents, such that equation (7) becomes

$$y_t = \gamma_p y_{t-1} + \gamma_{gp} y_{t-2} + v_t, \tag{11}$$

where v_t is a white-noise error term assumed to be uncorrelated to the outcome y of parents or earlier ancestors. In comparison to the descriptive associations captured by equation (4), we chose different symbols for the slope coefficients and the error term to indicate that equation (11) has a *structural* interpretation. Of course, if this "grandparent-effects" model is indeed the right model then the coefficients would coincide ($\beta_p = \gamma_p$ and $\beta_{gp} = \gamma_{gp}$), and we would observe "excess persistence" ($\Delta < 0$) iff $\gamma_{gp} > 0$.

Why might grandparents have an independent influence on their grandchildren? Mare (2011), Chan and Boliver (2013), Hertel and Groh-Samberg (2013) and Solon (2014) discuss various potential channels. Some of these mechanisms require overlapping lifespans or direct contact between grandparents and their children, such as when grandparents help in the upbringing of their grandchildren (Been et al. 2022 provide causal evidence), encourage or pay for educational investments, or transfer wealth. But other mechanisms do not: grandchildren may still benefit from the former contacts or reputation

[9] The model will tend to produce low multigenerational correlations (i.e., $\Delta < 0$) if parental income has a strong direct effect on child outcomes (Stuhler, 2012); moreover, $\Delta < 0$ holds with certainty in simplified versions of the Becker-Tomes model that abstract from the stochastic nature of the relation between income y and human capital e (Solon, 2014).

of their deceased grandparents, or may consider them as reference points for their own choices (Hertel and Groh-Samberg, 2013).

More generally, other members of the extended family may affect child outcomes. For example, Erola et al. (2018) find that in US and Finnish data, aunts and uncles are better predictors of child educational attainment earnings than grandparents, conditional on parent status. They note that this observation could be consistent with a model in which extended family members help with educational investments, provide access to educational opportunities and jobs, or serve as role models. And while equation (7) assumes linear effects, a more realistic model might account for interactions between different family members (in particular, the extended family might compensate for a lack of resources in the nuclear family; see Jæger 2012, Erola and Kilpi-Jakonen, 2017), the overlap in the lifespan of grandparents and their grandchildren, or the size of the extended family network (Lehti, Erola and Tanskanen, 2018).

4.3 Non-linear transmission processes and multiplicity

A third fundamental reason why multigenerational correlations may decay less rapidly than iterations of the parent-child correlation would suggest are non-linearity and other forms of heterogeneity in the transmission processes. To see this, consider a transmission process with "multiplicity" of pathways. Specifically, we introduce a second factor or "endowment" into the latent factor model,

$$y_t = \rho_1 e_{1t} + \rho_2 e_{2t} + u_t \tag{12}$$

$$e_{1t} = \lambda_1 e_{1,t-1} + v_{1t} \tag{13}$$

$$e_{2t} = \lambda_2 e_{2,t-1} + v_{2t}, \tag{14}$$

assuming that two endowments are inherited from parents according to transferability λ_1 and λ_2. For simplicity, assume that the noise terms v_{1t} and v_{2t} are uncorrelated, such that $Cov(e_{1t}, e_{2t}) = 0 \ \forall \ t$, and that both endowments affect incomes ($0 < \rho_1 < 1$ and $0 < \rho_2 < 1$). The parent-child and grandparent-grandchild correlations then equal

$$\begin{aligned}\beta_{-1} &= \rho_1^2\lambda_1 + \rho_2^2\lambda_2\\ \beta_{-2} &= \rho_1^2\lambda_1^2 + \rho_2^2\lambda_2^2\\ \dots\ &\ \dots\end{aligned}$$

and so on. In the special case in which incomes are *perfectly* determined by individual endowments, such that $\rho_1^2 + \rho_2^2 = 1$ and $Var(u_t) = 0$, the extrapolation error can be written as (Stuhler, 2012)

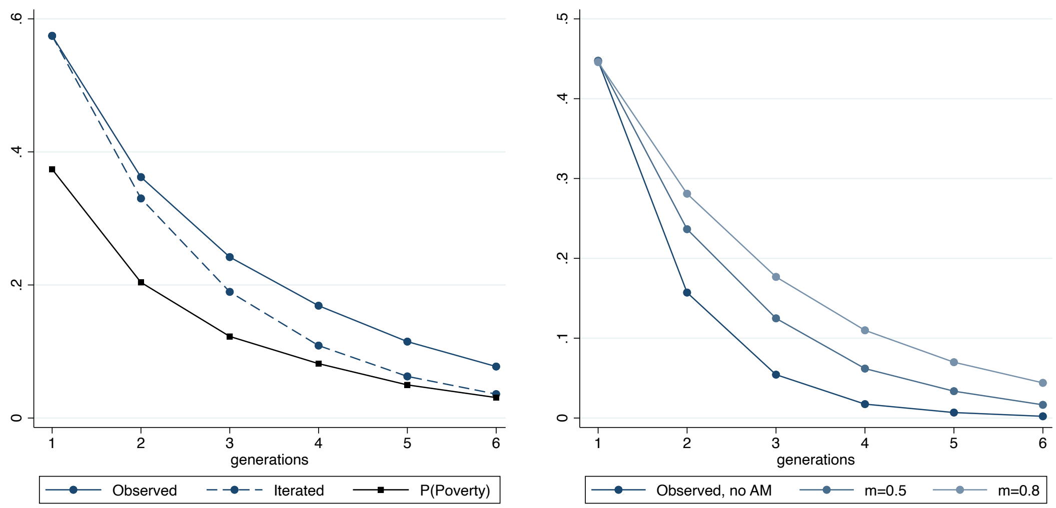
$$\Delta = \rho_1^2(\rho_1^2 - 1)(\lambda_1 - \lambda_2)^2, \tag{15}$$

which is negative for $\lambda_1 \neq \lambda_2$. This result can be understood as the application of Jensen's inequality: the square of the average transferability across endowments is smaller than the average of their square. Inequalities therefore decay more slowly if intergenerational income persistence stems from multiple causal pathways with differing rates of persistence.

Figure 1b provides a numerical example based on simulated data from this model. As in the latent

Figure 2: Multigenerational pattern in different models (continued)

(a) A non-linear model with "poverty traps" (b) A latent factor model with assortative mating



Notes: Simulated data. Panel (c) corresponds to the model in equation (16) with $\gamma_1 = 0.9$, $\gamma_2 = 0.2$ and $\underline{y} = -0.3$. The solid black line corresponds to the persistence in poverty ($y < \underline{y}$). Panel (d) corresponds to the latent factor model with assortative mating assuming $\rho = 0.8$, $\lambda = 0.7$ and $m = \{0, 0.5, 0.8\}$.

factor model, multigenerational correlations (solid blue line) decay more slowly than an iteration of parent-child correlations (dashed blue line) would suggest. Moreover, the more transferable trait explains an increasing share of the long-run persistence in income (dashed line). One interesting implication is that multigenerational persistence might be primarily due to factors that are not very important in explaining intergenerational persistence; in our illustration, endowment e_1 contributes less to parent-child transmission than endowment e_2, but causes most of the multigenerational persistence. One relevant example could be racial segregation and discrimination, which could generate strong multigenerational persistence (Hertel and Groh-Samberg 2013, Margo 2016).

Similar implications arise with other forms of heterogeneity. For example, the strength of the transmission processes may vary across the distribution of socio-economic status y, with higher persistence of status in some parts of the distribution than others. For illustration, consider an example with "poverty traps" in which there is modest transmission of status in the upper parts of the distribution but high persistence in the bottom below some threshold $\underline{y}$, i.e.

$$y_t = \gamma_1 I(y_{t-1} < \underline{y})\left(y_{t-1} - \underline{y}\right) + \gamma_2 I(y_{t-1} \geq \underline{y})\left(y_{t-1} - \underline{y}\right) + u_t \tag{16}$$

where $\gamma_1 > \gamma_2$. Figure 2a provides a numerical example; we again find excess persistence in the form of $\Delta < 0$ in multigenerational correlations (blue solid line) compared to the prediction based on iterated parent-child correlations (dashed line). The probability to remain poor, in the sense of $y < \underline{y}$, decays slowly as well (black line).

These examples illustrate that the observation of high multigenerational correlations may simply be a consequence of our tendency to ignore non-linearities and heterogeneity in the parent-child transmission process. But while such non-linearities have received much attention in other contexts, its implications for multigenerational transmission have received little attention so far. Important recent exceptions are Bingley and Cappellari (2019), who note that the strength of transmission may vary systematically across families or different groups, Colagrossi et al. (2020), who note that such heterogeneity would affect the relative size of intergenerational and sibling correlations, and Benhabib, Bisin and Fernholz (2022), who show that a model with permanent heterogeneity in wealth returns could match the wealth distribution in both the short- and long-run.

4.4 Testing theories of multigenerational transmission

Very different mechanisms could therefore explain the multigenerational patterns documented in recent studies. Indeed, different fields have emphasised different interpretations. The latent-factor interpretation has been popular in economics (e.g., Clark 2014, Braun and Stuhler, 2018, Neidhöfer and Stockhausen 2019, Colagrossi, d'Hombres and Schnepf 2020). The grandparent-effect interpretation has received particular interest in demography, after Mare (2011) and Mare (2014) called attention to the role of the wider family. Both perspectives are found in sociology (Chan and Boliver, 2013; Engzell, Mood and Jonsson, 2020). Finally, the implications of non-linear transmission on multigenerational patterns has received little attention in any field.

How can we then distinguish between these distinct interpretations? First, the observation of multigenerational persistence as such does not point to any particular theory. In particular, the finding that $\beta_{gp} > 0$ when estimating equation (4) should not be interpreted as favouring the grandparent-effect interpretation. As follows from (5), *any* process that generates high multigenerational persistence in the sense of $\Delta < 0$ will also generate a positive grandparent coefficient, and vice versa. The observations that $\beta_{gp} > 0$ or $\beta_{-2} > (\beta_{-1})^2$ do therefore not imply that intergenerational transmission has a memory of more than one generation, and researchers need to consider more specific evidence to distinguish between different candidate models.

One testable implication of the latent factor model is that the coefficient β_{gp} should be sensitive to which parental characteristics are controlled for.[10] Warren and Hauser (1997) show that after conditioning on a more comprehensive set of parental characteristics, the occupational status of grandparents is not predictive of child status. Many recent studies report similar "conditioning tests" (see Table 1), and typically find that β_{gp} shrinks but remains positive even after controlling for a wide set of parental controls (Sheppard and Monden, 2018). Such residual associations can still be consistent with a latent factor model, as some parental characteristics might be systematically missing in the data.[11] Using rich data from Sweden, Engzell, Mood and Jonsson (2020) show that even models that control for both parents' education, earnings, occupation, and wealth may still suffer from omitted parental variables.[12]

One obvious source of omitted variable bias is the omission of one of the parents: the coefficient β_{gp} on grandparent's status tends to be smaller when controlling for *both* paternal and maternal characteristics. For example, Braun and Stuhler (2018) find that in German samples, the coefficient on

[10] See also a recent strand of the literature that combines multiple proxy measures of parental status to explain child outcomes (Vosters and Nybom 2017, Blundell and Risa 2018, Hsu 2021 or Eshaghnia et al. 2021), or the literature on inequality of opportunity that often considers a wide set of "circumstances" (Brunori, Ferreira and Peragine 2013, Brunori, Hufe and Mahler 2022).

[11] To distinguish latent factor and non-Markovian models, the more relevant question may therefore be *by how much* β_{gp} declines when controlling for additional characteristics of the parents, rather than whether β_{gp} remains positive.

[12] Related, Modalsli and Vosters (2022) highlight the potential influence of measurement error in income, showing that it may generate a spurious grandparent coefficient even if a long-term average of parental income is controlled for.

grandfather's schooling or occupational prestige declines strongly once we condition on the corresponding status of the mother. Neidhöfer and Stockhausen (2019) and Engzell, Mood and Jonsson (2020) find that the coefficient on grandparents can become negligible once both parents are accounted for. Detailed conditioning tests are also provided by Chiang and Park (2015), Fiel (2018) or Sheppard and Monden (2018).

Another interesting test is whether the association between grandparent and child status varies with the extent of contact between parents and children, pointing to "grandparent effects". However, Anderson, Sheppard and Monden (2018) find in their review of the literature that the coefficient β_{gp} does not appear to vary systematically with the likelihood of contact between grandparent and grandchild (as do Helgertz and Dribe 2021), with Zeng and Xie (2014) and Song and Mare (2019) as notable exceptions. However, a direct effect of grandparents on their grandchildren is also possible through means that do not require contact, such as financial transfers (Hällsten and Pfeffer, 2017). And while conditioning tests appear to suggest that "grandparent effects" largely reflect omitted variables in the parent generation, those tests are subject to interpretative issues, too (Breen, 2018).[13]

A third strategy is to estimate persistence at an aggregate rather than individual level. For example, the persistence rate λ in the simple latent factor model could be identified in a two-step process, averaging y across multiple relatives to then estimate the regression to the mean on the surname level. With this motivation in mind, Clark (2014), Clark and Cummins (2014) and related studies document strikingly strong persistence of socio-economic status at the *surname* level, across many different countries and time periods. This finding has led to much debate about the potential pitfalls of name-based estimators (see Santavirta and Stuhler 2021 for a review). While their precise interpretation is contested, those estimators might be useful to discriminate between competing models of multigenerational inequality: for example, a latent factor model generates higher persistence at the group than the individual level, while a simple "grandparent effects" model would not.

A fourth strategy is to confront competing transmission models with a wider set of kinship correlations. The "right" model of intergenerational transmission would explain not only the relation between multi- and intergenerational correlations, but also their relation to sibling and many other type of kinship correlations. For example, Collado, Ortuño-Ortín and Stuhler (2022), show that a generalised latent factor model can provide a good fit to a wide set of 141 distinct kinship moments in Swedish data. An interesting question is whether a non-linear or non-Markovian model could provide a similarly good fit to extended-kin data, and whether they could motivate why status appears to still correlate between very distant ancestors (Barone and Mocetti, 2020).

These obstacles in the interpretation of multigenerational correlations are of course the same obstacles that limit our understanding of social mobility more generally. Statistical associations are difficult to map to mechanisms, and while robust causal evidence exists for certain pathways, most of the statistical associations remain unaccounted for (Björklund and Jäntti, 2020). Moreover, distinguishing between the models considered in Section 4 is only a first challenge, as those stylised models are not precise about the *specific* mechanisms and behavioural patterns that matter, or how policies and institutions would affect them.

Still, the recent multigenerational evidence is not "toothless", as it reduces the range of permissible

[13] Breen (2018) shows that *partial* conditioning for some but not all relevant characteristics of the parent generation may not always reduce bias. Using the language of causal graphs, the issue is that parental status might not only be a "mediator" for the effect of grandparent on child status, but also a "collider" that is affected by other causal factors influencing both parent and child status (such as neighbourhood effects).

Table 2: On the explanatory power of multigenerational associations

	Dependent variable: Child status y					
	(1)	(2)	(3)	(4)	(5)	(6)
Parent's y	0.450***	0.389***	–	0.392***	–	0.310***
	(0.004)	(0.004)		(0.004)		(0.004)
Grandparent's y	–	0.138***	–	–	–	–
		(0.004)				
Sibling's y	–	–	0.306***	0.131***	0.458***	0.318***
			(0.004)	(0.004)	(0.004)	(0.004)
R^2	0.204	0.219	0.095	0.218	0.210	0.286

Notes: Simulated data from the latent factor model in equations (6) and (7) with returns $\rho = 0.8$ and transferability $\lambda = 0.7$ over three generations (n=50,000). For columns (3) and (4), the noise term u is uncorrelated between siblings. For columns (5) and (6), u is drawn from a joint normal distribution with correlation 0.4 between siblings.

models. For example, the standard implication of the Becker-Tomes model for β_{gp} to be negative (see Section 4.1) is rejected by most papers. Related, Collado, Ortuño-Ortín and Stuhler (2022) argue that a purely genetic model with phenotypic assortment could not fit the kinship pattern in educational advantages. And while there is no consensus yet on the underlying mechanisms, multigenerational estimates are of course directly informative about status persistence in the long run, thereby providing novel insights about an important dimension of inequality.

4.5 Does this matter? The R^2 controversy

Do multigenerational associations "matter", in that they tell us something meaningful about social mobility? One frequent observation is that conditional on parents, other ancestors do not add much explanatory power – even if the corresponding slope coefficients appear sizeable, the regression R^2 does not increase much. For example, Pfeffer and Killewald (2018) report that switching from a two- to a three-generations regression, the R^2 increases only mildly (from 0.146 and 0.160), even though the coefficient on the grandparents' wealth is highly significant and nearly half as large as the coefficient on the parents' wealth.[14] This low contribution in a R^2 sense is also a key point in an interesting recent debate between Björklund, Hedros and Jäntti (2022) and Adermon, Lindahl and Palme (2022) regarding the "dynastic" estimates presented by Adermon, Lindahl and Palme (2021). Are deviations from the "iterated" parent-child regression (Section 2) then just not very important?

To provide an illustration, we generate simulated data for three generations, based on the latent factor model given by (6) and (7) and the same parameters as in Figure 1 ($\lambda = 0.7$ and $\rho = 0.8$). Table 2 reports regression estimates from this simulated data. Column (1) reports estimates of the intergenerational regression coefficient β_{-1}, which according to equation (8) equals $\beta_{-1} = \rho^2\lambda$. In column (2) we add the grandparent status to the model. Given our choice of parameter values, the coefficient on grandparents is sizeable, at about one third of the coefficient on parent status. However, the regression R^2 hardly increases. Hence the conundrum: are deviations from the parent-child model sizeable, as indicated by the regression slopes, or negligible, as seemingly implied by the $R2$?

The answer depends on *why* we observe multigenerational associations. While the R^2 differs little

[14] Similarly, Erola and Moisio (2007) note that considering more than two consecutive generations adds little explanatory power in an analysis of social mobility in Finnish data.

between columns (1) and (2), the significant coefficient on grandparents signals that the transmission process deviates from the simple parent-child regression. And in our chosen example, that deviation turns out to be important: advantages are transmitted at a *much* higher rate than the parent-child correlation would suggest ($\lambda = 0.7$ vs. $\beta \approx 0.45$) and the multigenerational implications differ substantially (Figure 1a). The observation of independent multigenerational associations can therefore be meaningful, even if the regression R^2 moves little.

It is instructive to consider a similar simulation for *sibling correlations*, which are an alternative measure of family influences. We simulate two children per parent, initially assuming that the errors u and v in equations (6) and (7) are uncorrelated between siblings. As shown in column (3) of Table 2, the implied sibling correlation (0.31) is much larger than the square of the parent-child correlation (0.31 vs. $0.45^2 = 0.20$), in line with the pattern in actual applications (Björklund and Salvanes, 2011, Jäntti and Jenkins 2014).[15] What if we condition on siblings *and* parents, as in column (4)? While the coefficient on siblings is sizeable, the R^2 hardly increases compared to the bivariate parent-child regression in column (1) – mirroring the pattern for multigenerational correlations. The example illustrates that the "excess" persistence found in multigenerational studies (as in column 2) and the wedge between (the square of) intergenerational and sibling correlations (columns 1 and 3) could reflect similar underlying mechanisms.

However, the measures are not interchangeable: sibling correlations are a broad measure of family background that also capture influences that siblings share over and above their parental influences (such as neighbourhood or peer effects), while multigenerational correlations capture the effect of ancestry in a more narrow sense. To illustrate this point, columns (5) and (6) replicate the regressions from columns (3) and (4) but allow for shared influences between siblings. Specifically, we draw the error u in equation (6) from a joint normal distribution with correlation 0.4 between siblings. This increases the sibling correlation further (cf. columns 3 and 5), and the R^2 in a joint regression with parents *and* siblings is now substantially larger as compared to the bivariate regression conditioning only on parents (cf. columns 1 and 6). This is the pattern we observe in actual data, reflecting that in contrast to multigenerational correlations, sibling correlations capture more than just the effect of ancestors.

5 Assortative Matching

The observation of high multigenerational correlations has important implications for the extent of assortative mating between spouses. To see this, extend the latent factor model as given by equations (6) and (7) to a two-parent setting, assuming that an offspring's endowment depends on the average of the father's and mother's endowment,

$$e_t = \tilde{\lambda}\frac{e^m_{it-1} + e^p_{it-1}}{2} + v_t, \tag{17}$$

where the m and p superscripts denote maternal and paternal variables, respectively.[16] Normalising all y and e to one, the parent-child correlation in y between a child and *one* of his or her parents is

[15] We measure the sibling correlation as the slope in a regression of a child's outcome on her sibling's outcome.

[16] The presentation here follows Braun and Stuhler (2018), which contains further details.

now given by

$$\beta_{-1} = \rho^2 \tilde{\lambda} \frac{1+m}{2}, \tag{18}$$

where $m = Corr\left(e^m_{i,t-1}, e^p_{i,t-1}\right)$ measures the extent of assortative matching among parents. The multigenerational correlations with earlier ancestors are similarly given by

$$\beta_{-k} = \rho^2 \tilde{\lambda}^k \left(\frac{1+m}{2}\right)^k \tag{19}$$

for $k > 1$. Because β_{-k} depends on $\left(\frac{1+m}{2}\right)^k$, multigenerational correlations will decay quickly if assortative matching is weak – even if the average parental endowments are transmitted perfectly from parents to children.

Figure 2b provides a numerical example with $\rho = 0.8$ and $\tilde{\lambda} = 0.7$. We plot the implied multigenerational correlations in y for three different assortative correlations, with $m = 0$ (no assortative matching), $m = 0.5$ (a typical value for the spousal years of schooling, Fernandez, Guner and Knowles, 2005) and $m = 0.8$. The example illustrates that high multigenerational correlations are possible only if assortative matching is very strong, a point that is also discussed by Clark (2014). While we derived this argument in the context of the latent factor model, it applies similarly to alternative transmission models.

Moreover, inverting this argument suggests that multigenerational estimates are informative about assortative pattern. Indeed, conventional assortative measures may understate the sorting in a population for the same reasons that intergenerational correlations understate intergenerational transmission: the similarity of spouses in observable characteristics, such as years of schooling, may not capture their resemblance in unobserved characteristics that influence child outcomes. Allowing for latent sorting, Collado, Ortuño-Ortín and Stuhler (2022) estimate that in Sweden, the spousal correlation in latent endowments must be around 0.75 to explain the correlation of educational outcomes between distant in-laws – more than 50 percent larger than the spousal correlation in years of schooling. Such strong degree of sorting is surprising in so far as Sweden is thought to be comparatively egalitarian; one open question is whether spousal sorting might be even stronger in more unequal countries.

6 Conclusions

This chapter provided an overview of the fast-growing literature on multigenerational inequality. Using data across three or more generations, recent studies show that socio-economic inequality tends to be more persistent than a naive extrapolation from conventional parent-child measures would suggest. However, very distinct interpretations of that new empirical "fact" are possible. Its significance might lie less in the magnitude of multigenerational associations per se, as adding grandparents or extended kins often contributes little to the overall explanatory power of intergenerational models. Instead, multigenerational associations tell us something novel about the nature of intergenerational processes.

One leading interpretation is that parent-child transmission, and assortative mating, must be much stronger than what is captured by conventional measures. But the chapter also reviews alternative interpretations, related to "grandparent effects" or non-linearities in the transmission process, and more work is needed to distinguish between competing models of multigenerational transmission. One exciting aspect is that this work is happening simultaneously in multiple fields of the social sciences, enriching the discussion by connecting different strands of research on social mobility that otherwise progress in isolation.

We have ignored some issues that warrant attention in applications. On the empirical side, one important concern is sample selection. The data requirements for multigenerational studies are high, and authors might run into selection issues regarding the coverage of different generations and the age at which the outcomes y can be observed. A more explicit incorporation of fertility choices, as in Song (2021), could also be a useful avenue for future research. Some multigenerational studies relyb on data with limited regional coverage, and are therefore subject to sample selection related to migration. On the theoretical side, many studies rely on *steady-state assumptions*, but the assumption that distributional moments are fixed is particularly questionable in data spanning across multiple generations, and when comparing socio-economic outcomes today to the "same" outcomes in the distant past (Nybom and Stuhler, 2019). A more explicit consideration of transitional dynamics of intergenerational processes could be fruitful for future work.

References

Adermon, Adrian, Mikael Lindahl, and Daniel Waldenström. 2018. "Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations." *The Economic Journal*, 128(612): F482–F513.

Adermon, Adrian, Mikael Lindahl, and Mårten Palme. 2021. "Dynastic Human Capital, Inequality, and Intergenerational Mobility." *American Economic Review*, 111(5): 1523–48.

Adermon, Adrian, Mikael Lindahl, and Mårten Palme. 2022. "Svar på Anders Björklund, Karin Hederos och Markus Jäntti" Social rörlighet och jämlikhet i möjligheter–en kommentar"." *Ekonomisk Debatt*, 50(6): 62–64.

Anderson, Lewis R., Paula Sheppard, and Christiaan W. S. Monden. 2018. "Grandparent Effects on Educational Outcomes: A Systematic Review." *Sociological Science*.

Anger, Silke, and Daniel D Schnitzlein. 2017. "Cognitive Skills, Non-cognitive Skills, and Family Background: Evidence From Sibling Correlations." *Journal of Population Economics*, 30(2): 591–620.

Barone, Guglielmo, and Sauro Mocetti. 2020. "Intergenerational Mobility in the Very Long Run: Florence 1427-2011." *Review of Economic Studies*.

Becker, Gary S., and Nigel Tomes. 1979. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*, 87(6): 1153–1189.

Becker, Gary S., and Nigel Tomes. 1986. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4(3): 1–39.

Been, Jim, Anne C Gielen, Marike Knoef, Gloria Moroni, et al. 2022. "Prolonged Worklife Among Grandfathers: Spillover effects on Grandchildren's Educational Outcomes." Tinbergen Institute.

Benhabib, Jess, Alberto Bisin, and Ricardo T Fernholz. 2022. "Heterogeneous Dynasties and Long-Run Mobility." *The Economic Journal*, 132(643): 906–925.

Bingley, Paul, and Lorenzo Cappellari. 2019. "Correlation of Brothers' Earnings and Intergenerational Transmission." *Review of Economics and Statistics*, 101(2): 370–383.

Björklund, Anders, and Kjell G. Salvanes. 2011. "Education and Family Background: Mechanisms and Policies." In *Handbook of the Economics of Education.* Vol. 3, , ed. Eric A. Hanushek, Stephen Machin and Ludger Woessmann, 201–247. Elsevier.

Björklund, Anders, and Markus Jäntti. 2012. "How Important is Family Background for Labor-economic Outcomes?" *Labour Economics*, 19(4): 465–474.

Björklund, Anders, and Markus Jäntti. 2020. "Intergenerational Mobility, Intergenerational Effects, Sibling Correlations, and Equality of Opportunity: A Comparison of Four Approaches." *Research in Social Stratification and Mobility*, 70: 100455.

Björklund, Anders, Karin Hedros, and Markus Jäntti. 2022. "Social rörlighet och jämlikhet i möjligheter–en kommentar." *Ekonomisk Debatt*, 50(5): 59–64.

Blanden, Jo. 2013. "Cross-Country Rankings in Intergenerational Mobility: A Comparison of Approaches from Economics and Sociology." *Journal of Economic Surveys*, 27(1).

Blanden, Jo, Matthias Doepke, and Jan Stuhler. 2022. "Educational Inequality." In *Handbook of the Economics of Education.* Elsevier.

Blundell, Jack, and Erling Risa. 2018. "Intergenerational Mobility as a Prediction Problem." mimeo.

Boserup, Simon Halphen, Wojciech Kopczuk, and Claus Thustrup Kreiner. 2013. "Intergenerational Wealth Mobility: Evidence from Danish Wealth Records of Three Generations." University of Copenhagen.

Braun, Sebastian Till, and Jan Stuhler. 2018. "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany." *The Economic Journal*, 128(609): 576–611.

Breen, Richard. 2018. "Some Methodological Problems in the Study of Multigenerational Mobility." *European Sociological Review*, jcy037.

Brunori, Paolo, Francisco HG Ferreira, and Vito Peragine. 2013. "Inequality of Opportunity, Income inequality, and Economic Mobility: Some International Comparisons." In *Getting Development Right.* 85–115. Springer.

Brunori, Paolo, Paul Hufe, and Daniel Mahler. 2022. "The Roots of Inequality: Estimating Inequality of Opportunity from Regression Trees and Forests." *Scandinavian Journal of Economics.*

Bulmer, Michael. 2003. *Francis Galton: Pioneer of Heredity and Biometry.* Johns Hopkins University Press.

Cavalli-Sforza, Luigi Luca, and Marcus W Feldman. 1981. *Cultural Transmission and Evolution: A Quantitative Approach.* Princeton University Press.

Chan, Tak Wing, and Vikki Boliver. 2013. "The Grandparents Effect in Social Mobility: Evidence from British Birth Cohort Studies." *American Sociological Review*, 78(4): 662–678.

Chiang, Yi-Lin, and Hyunjoon Park. 2015. "Do grandparents matter? A multigenerational perspective on educational attainment in Taiwan." *Social science research*, 51: 163–173.

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

Clark, Gregory, and Neil Cummins. 2014. "Intergenerational Wealth Mobility in England, 1858–2012: Surnames and Social Mobility." *The Economic Journal*, 125(582): 61–85.

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

Colagrossi, Marco, Claudio Deiana, Andrea Geraci, Ludovica Giua, and Gianluca Mazzarella. 2020. "Two Generations of Siblings and Cousins Correlations Across the Ancestors' Wealth Distribution." mimeo.

Collado, M. Dolores, Ignacio Ortuño-Ortín, and Jan Stuhler. 2022. "Estimating Intergenerational and Assortative Processes in Extended Family Data." *Review of Economic Studies (forthcoming).*

Dribe, Martin, and Jonas Helgertz. 2016. "The Lasting Impact of Grandfathers: Class, Occupational Status, and Earnings over Three Generations in Sweden 1815–2011." *The Journal of Economic History*, 76(4): 969–1000.

Duncan, Otis Dudley. 1969. "Contingencies in Constructing Causal Models." *Sociological Methodology*, 1: 74–112.

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

Erola, Jani, and Elina Kilpi-Jakonen. 2017. *Social Inequality Across the Generations: The Role of Compensation and Multiplication in Resource Accumulation.* Edward Elgar Publishing.

Erola, Jani, and Pasi Moisio. 2007. "Social Mobility over Three Generations in Finland, 1950-2000." *European Sociological Review*, 23(2): 169–183.

Erola, Jani, Elina Kilpi-Jakonen, Irene Prix, and Hannu Lehti. 2018. "Resource Compensation from the Extended Family: Grandparents, Aunts, and Uncles in Finland and the United States." *European Sociological Review*, 34(4): 348–364.

Eshaghnia, Sadegh, James Heckman, Rasmus Landersø, Rafeh Qureshi, and Victor

Ronda. 2021. "The Intergenerational Transmission of Lifetime Wellbeing." mimeo.

Fernandez, Raquel, Nezih Guner, and John Knowles. 2005. "Love and money: A theoretical and empirical analysis of household sorting and inequality." *The Quarterly Journal of Economics*, 120(1): 273–344.

Ferrie, Joseph, Catherine Massey, and Jonathan Rothbaum. 2020. "Do Grandparents Matter? Multigenerational Mobility in the US, 1940-2015." *Journal of Labor Economics.*

Fiel, Jeremy E. 2018. "The Transmission of Multigenerational Educational Inequality." *Social Forces*, soy090.

Gallegos, Sebastian, and Pablo A. Celhay. 2022. "Patterns of Educational Mobility Across Three Generations in Latin America." CAF Working Paper Series 2022/09.

GDIM. 2018. "Global Database on Intergenerational Mobility." *World Bank Group Washington, DC.*

Goldberger, Arthur S. 1972. "Structural Equation Methods in the Social Sciences." *Econometrica*, 40(6): 979–1001.

Goldberger, Arthur S. 1989. "Economic and Mechanical Models of Intergenerational Transmission." *The American Economic Review*, 79(3): pp. 504–513.

Hällsten, Martin, and Fabian T Pfeffer. 2017. "Grand Advantage: Family Wealth and Grandchildren's Educational Achievement in Sweden." *American Sociological Review*, 82(2): 328–360.

Hällsten, Martin, and Martin Kolk. 2020. "The Shadow of Peasant Past: Seven Generations of Inequality Persistence in Northern Sweden." Stockholm University.

Helgertz, Jonas, and Martin Dribe. 2021. "Do Grandfathers Matter for Occupational and Earnings Attainment? Evidence from Swedish Register Data." *European Sociological Review.* jcab029.

Hertel, Florian R., and Olaf Groh-Samberg. 2013. "Class Mobility Across Three Generations in the U.S. and Germany." *Research in Social Stratification and Mobility*, 35: 35–52.

Hodge, Robert. 1966. "Occupational Mobility as a Probability Process." *Demography*, 3(1): 19–34.

Hsu, Hsiu-Fen. 2021. "Intergenerational Persistence in Latent Socioeconomic Status: Evidence From Taiwan." *The Journal of Economic Inequality*, 19(4): 753–772.

Jæger, Mads Meier. 2012. "The Extended Family and Children's Educational Success." *American Sociological Review*, 77(6): 903–922.

Jäntti, Markus, and Stephen Jenkins. 2014. "Income Mobility." In *Handbook of Income Distribution.* Vol. 2, , ed. Anthony Atkinson and François Bourguignon. Elsevier.

Kundu, Anustup, and Kunal Sen. 2021. "Multigenerational Mobility in India." *WIDER Working Paper Series wp-2021-32, World Institute for Development Economic Research (UNU-WIDER).*

Lehti, Hannu, Jani Erola, and Antti O Tanskanen. 2018. "Tying the Extended Family Knot—Grandparents' Influence on Educational Achievement." *European Sociological Review*, 35(1): 29–48.

Lindahl, Mikael, Mårten Palme, Sofia Sandgren Massih, and Anna Sjögren. 2015. "Long-term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations." *Journal of Human Resources*, 50(1): 1–33.

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

Mare, Robert D. 2014. "Multigenerational aspects of social stratification: Issues for further research." *Research in Social Stratification and Mobility*, , (35): 121–128.

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

Modalsli, Jørgen. 2021. "Multigenerational persistence: Evidence from 146 years of administrative data." *Journal of Human Resources*, 1018–9825R2.

Modalsli, Jørgen, and Kelly Vosters. 2022. "Spillover Bias in Multigenerational Income Regressions." *Journal of Human Resources*, 0520–10919R2.

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

Nesselroade, John R., Stephen M. Stigler, and Paul B. Baltes. 1980. "Regression Toward the Mean and the Study of Change." *Psychological Bulletin*, 88(3).

Nybom, Martin, and Jan Stuhler. 2019. "Steady-state Assumptions in Intergenerational Mobility Research." *The Journal of Economic Inequality.*

Pfeffer, Fabian T. 2014. "Multigenerational Approaches to Social Mobility. A Multifaceted Research Agenda." *Research in Social Stratification and Mobility*, 35: 1–12.

Pfeffer, Fabian T, and Alexandra Killewald. 2018. "Generations of Advantage. Multigenerational Correlations in Family Wealth." *Social Forces*, 96(4): 1411–1442.

Razzu, Giovanni, and Ayago Wambile. 2020. "Three-Generation Educational Mobility in Six African Countries." University of Reading Discussion Paper No. 2020-23.

Rice, John, Robert C. Cloninger, and Theodore Reich. 1978. "Multifactorial Inheritance with Cultural Transmission and Assortative Mating. I. Description and Basic Properties of the Unitary Models." *American Journal of Human Genetics*, 30(6): 618.

Santavirta, Torsten, and Jan Stuhler. 2021. "Name-Based Estimators of Intergenerational Mobility." mimeo.

Sheppard, Paula, and Christiaan Monden. 2018. "The Additive Advantage of Having Educated Grandfathers for Children's Education: Evidence from a Cross-National Sample in Europe." *European Sociological Review*, 34(4): 365–380.

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

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

Song, Xi. 2021. "Multigenerational Social Mobility: A Demographic Approach." *Sociological Methodology*, 51(1): 1–43.

Song, Xi, and Robert D Mare. 2019. "Shared Lifetimes, Multigenerational Exposure, and Educational Mobility." *Demography*, 56(3): 891–916.

Stuhler, Jan. 2012. "Mobility Across Multiple Generations: The Iterated Regression Fallacy." Institute for the Study of Labor (IZA) IZA Discussion Papers 7072.

Vosters, Kelly, and Martin Nybom. 2017. "Intergenerational Persistence in Latent Socioeconomic Status: Evidence from Sweden and the United States." *Journal of Labor Economics*, 35(3): 869–901.

Warren, John Robert, and Robert M. Hauser. 1997. "Social Stratification across Three Generations: New Evidence from the Wisconsin Longitudinal Study." *American Sociological Review*, 62(4): 561–572.

Zeng, Zhen, and Yu Xie. 2014. "The Effects of Grandparents on Children's Schooling: Evidence From Rural China." *Demography*, 51(2): 599–617.