Kinship Correlations and Intergenerational Mobility[*]

M. Dolores Collado[a] Ignacio Ortuño-Ortín[b] Jan Stuhler[b]

February 2019
– Preliminary draft –

Abstract

We propose a new methodology to estimate long-run intergenerational socioeconomic mobility. Our approach does not require information across multiple generations, but considers different degrees of kinship within the same generation. This "*horizontal*" approach has two key advantages. First, socioeconomic status can be measured in the same data source and at the same time. Second, the horizontal approach yields many more kinship moments than the vertical approach. This allows us to consider a more detailed intergenerational model, with *direct* and *indirect* (latent) transmission mechanisms; shared sibling influences and assortative mating in observable and unobservable characteristics; and gender asymmetries in all parameters. We fit that model for income and education using registry data from Sweden and census data from Spain. We find strong persistence in the intergenerational, sibling, and assortative processes. Striking is the high rate of assortative matching – to rationalize our kinship data, spouses must be far more similar to each other than they are in observables such as education. Our model nests the standard genetic model, as well as other models used in the recent multigenerational literature. The standard genetic model cannot fit the kinship pattern in educational attainment.

---

[*] We are grateful for very helpful comments and suggestions from Adrian Adermon, Gregory Clark, Mikael Lindahl, Martin Nybom and seminar participants at the Université du Québec à Montréal, Nuffield College at Oxford University, the Joint Research Center of the European Union in Ispra, Rotterdam University, Stockholm University, and the Selten Insitute in Cologne. (a) Universidad de Alicante; (b) Universidad Carlos III de Madrid

1 Introduction

Research on the degree of socioeconomic intergenerational has received renewed interest in recent years. Partly this interest is due to novel evidence on the variation of *parent-child* mobility across regions within countries (see for example Chetty et al. 2014), and its possible relation with the recent increase in income inequality (Corak 2013). This interest is also due to recent evidence on *multi*generational mobility, which suggests that mobility is perhaps much lower than what most economists used to think (Clark 2014, Lindahl et al. 2015, Barone and Mocetti 2016). By tracking socioeconomic success across multiple generations, this recent literature contradicts a common interpretation of the available parent-child evidence – that the correlation between individuals in one generation and their ancestors decreases fast as we go back in time, so that after, say, three or four generations the link is already very weak.[1]

Instead, recent empirical studies suggest a much higher persistence rate in socioeconomic status, and a significant link with grandparents or more distant ancestors. Economic historians have made important contributions. Using data from a series of countries and time periods, Gregory Clark and co-authors (e.g. Clark, 2014; Clark et al., 2015) show that the average socioeconomic status of *surnames* regresses at a rather slow rate. For example, in historical data from Florence the average status of surnames still correlates across generations that are six centuries apart (Barone and Mocetti, 2016). Other studies use individual-level data linked across multiple generations to estimate multigenerational persistence (Lindahl et al., 2015; Braun and Stuhler, 2018; Neidhöfer and Stockhausen, forthcoming; Adermon et al., 2018), and to study the role of grandparents in the transmission process (e.g. Breen, 2018; Anderson et al., 2018). These studies likewise find high intergenerational persistence, if not as high as studies on the surname level.

One problem with the multigenerational or *"vertical"* approach adopted in these works are the data requirements. It is difficult to obtain comparable socioeconomic information for more than two generations. For instance, in many countries there is very little variation in years of formal education for older generations if its majority had only basic education. Occupational classifications are useful, but going sufficiently back

---

[1] Such interpretations are based on the iteration of parent-child regressions (see Stuhler, 2012), i.e. the assumption that the correlation between grandparents and grandchildren outcomes is basically the square of the parent-offspring correlation. Since parent-offspring correlations in income, education or other socioeconomic outcomes are always moderate, ancestor correlations would decrease fast as we go back in time. As put by Becker and Tomes (1986), "*Almost all earnings advantages and disadvantages of ancestors are wiped out in three generations. Poverty would not [...] persist for several generations.*"

in time we invariably end up with a sample that consists mostly of farmers. Data that directly link individuals are often not available, so most studies use surnames as an imperfect proxy for actual ancestor relations – triggering a lively debate on how informative surname-level processes can be about individual-level mobility. Studies based on direct ancestor links avoid this problem, but only track the more immediate ancestors (i.e. the grandparent or great-grandparent generation). A limitation common to both surname- and individual-level studies is the small set of empirical moments that the vertical approach yields, limiting our ability to distinguish between competing models of intergenerational transmission.

Thus, we propose a new approach to asses the degree of long-run intergenerational mobility that does not require information on previous generations. Instead we use "*horizontal*" information, that is, information about individuals of the same generation, or very close generations, who are relatives of a certain degree, for example siblings, cousins, second cousins, parent-child, uncle-nephew. The underlying idea is simple. Say that we would like to asses the link between grandparents and grandsons but we don't have data for grandparents to directly measure it. But if instead we have good data for cousins we can infer the grandparents-grandsons link from the cousins links. Thus, horizontal information can overcome the lack of vertical information. We believe that our method has an important advantage since it does not require information on individuals in previous generations, and therefore, it can be applied to study long-run intergenerational mobility in many countries in which there is no comparable data on individuals in several generations. A key advantage of the horizontal approach is that socio-economic outcomes can be measured within the same data source, and at approximately the same age and time.[2]

This horizontal approach can be considerably strengthened by considering *affine* kinships, such as siblings in-law. As *consanguine* ("blood") kinships, affine relations are a function of – and therefore informative about – intergenerational and assortative processes. However, the identification of distant consanguine kins requires first the identification of distant ancestors. For example, the identification of cousins requires observation of their shared grandparents, while the identification of second-degree cousins requires observation of their great-grandparents. In contrast, affine relationships are identified by spousal and parental links *irrespective* of their degree of separation. In a first step, we identify a person's sibling via their shared parent. In a second step, we identify the siblings' spouse via their shared descendant. Implemented once, these

---

[2] An important influence for our approach is the literature on sibling correlations, on which we comment below. In related work, Hällsten (2014) considers the correlation between cousins and second cousins, and discusses the advantage of using data from the same generation.

steps identify siblings in-law. Implemented twice, we identify the sibling-in-law of the sibling-in-law, and so on.

The approach scales particularly well in large-scale registry and census data, which cover a large share of the population. We consider two such data sources, from Sweden and Spain. The Swedish registers are very extensive, include family links over multiple generations, and education, income and other outcomes. Our sample covers more than one third of the population, allowing us to identify very distant in-laws. In this paper, we consider in-law relations up to five degrees of separations in this paper (i.e. individuals separated by five sibling plus five spousal links).[3] In contrast, the Spanish data are limited to a single cross-section from the region of Cantabria, which contains only educational outcomes, and names instead of direct family links. We exploit Spanish naming conventions to recover parent-child links (and therefore siblings, first-cousins, uncles-nephews and sibling-in-laws) for a substantial share of our sample. We can identify 141 distinct kinship moments in the Swedish registers, and 65 moments in the Spanish sources.

The horizontal perspective yields therefore *many more*, and more comparable and better measured kinship moments than the vertical approach. That opens the door for the estimation of more detailed intergenerational models, and a deeper understanding of transmission processes within the nuclear family, shared influences among siblings, and the assortative process. Our model builds on the key implication from the recent multigenerational literature: latent variables must play an important role in the transmission process, because parent-child correlations in observables are too low to rationalize the persistence of inequalities across multiple generations.[4] However, it deviates in three aspects from the prior literature. First, we allow for *both* direct and latent transmission mechanisms instead of considering only observable or latent factors (as in Clark, 2014, or other recent work). Second, we allow for assortative mating in both the observable and the latent dimension. Third, we allow each of those mechanisms to vary with the gender of the child and the parent.

We calibrate this model using the kinship moments from Sweden and Spain, using educational attainment as our baseline outcome. For Sweden, we also consider income

---

[3] While more distant family members can be identified in principle, we abstain from using them because selective fertility becomes an increasingly important issue that is not modeled in this paper.

[4] That part of the transmission process may be inherently unobservable has long been recognized in the theoretical literature (e.g. Duncan, 1969; Goldberger, 1972). The latent variable in our model is close in spirit to Becker and Tomes (1979), who assume that a person's *"endowment"* represent a great variety of cultural and genetic attributes. Latent transmission channels have received less attention in empirical work, but are central for the recent literature on *multigenerational* transmission.

and (for reasons that will become apparent later on) body height. In the first part of the paper, we study how intergenerational, sibling, and assortative processes contribute to the overall persistence of inequality across generations. Our objective here is to quantify the transmission of inequality, without trying to identify the precise channel of transmission. However, our model is general enough to nest specific models, such as the standard genetic model from behavioral genetics. We therefore test how well such genetic model could explain our kinship data in the final part of the paper.

We find strong persistence in the intergenerational, sibling, and assortative processes. The parent-child correlation of the latent factor is around 0.6 in our baseline for Sweden, and about 0.8 in the Spanish sources. Our results therefore imply that the educational levels of individuals in the current generation are still correlated in a non-negligible magnitude with the socioeconomic status of their ancestors as much as four or five generations back in time.[5] In contrast, the observed educational attainment of the parent has a positive but only limited independent association with the educational attainment of their child (consistent with evidence on the causal effect of parents' schooling, see e.g. Holmlund et al., 2011).

Striking is the high rate of assortative matching that our data imply. Kinship correlations decay only slowly across sibling in-laws, falling by only about 25% with each step of separation. To rationalize this slow decay, both spouses (and siblings) must be substantially more similar in those factors that determine the economic success of their children than they are in observable characteristics, such as educational outcomes. We estimate the spousal correlation in the latent factor to be about 0.75 in the Swedish and about 0.9 in our Spanish sources – far higher than the spousal correlation in educational attainment, which is below 0.5 in Sweden and only slightly higher in our Spanish sources.

The strong transmission of the latent factor from parents to their children also contribute to the slow decay of kinship correlations in the horizontal dimension. We demonstrate that siblings share important influences that are only partially reflected in their educational attainment. Similar as spouses, siblings must be far more similar to each other in characteristics that matter for the socioeconomic success of their descendants than what is reflected in their observable characteristics. The implied sibling correlations are about 0.7 in both Spanish and Swedish sources, while the sibling correlation in observables such as years of schooling is below 0.5.

---

[5] Collado et al. (2014) analyze long-run mobility in the same Spanish region using census data from the XIX and the XX century. They find a higher level of mobility than the one in this paper. This discrepancy might be explained because they only consider two socioeconomic levels whereas here individuals are classified according to 10 possible levels of education (years of schooling)

These results are based on educational attainment, which is seen as the key mediator for the transmission of socio-economic advantages (Goldthorpe 2014). To study if our results generalize to other socioeconomic outcomes, we calibrate our model with income data obtained from the Swedish registers. We do not model labor supply decisions, but our model is flexible enough to account for the fact that kinship correlations are systematically lower for mixed or female than for male pairs. The results are qualitatively similar to our benchmark calibration based on years of schooling, but differ in magnitude. The latent advantages are more strongly transmitted than income itself, across all three dimensions of our model. However, our results suggest that those latent factors that determine educational attainment are more strongly transmitted from one generation to the next than those factors that influence earnings.

Because we observe a greater set of empirical moments than previous studies, we can study the robustness of these results in more detail. In particular, we show that our results are robust to calibration from different subset of our empirical moments, and test the out-of-sample performance of our model. We show that our baseline results remain similar when dropping two thirds of the empirical moments from the Swedish registers. Moreover, the model calibrated from this restricted set of moments successfully predicts a diverse set of kinship correlations not included in the calibration, including vertical, horizontal and distant moments. Our model appears identifiable from information that is far more limited than what we observe in the Swedish data, or even our Spanish sources.

Our results contribute to several strands of the literature on inequality and intergenerational mobility. First, we corroborate the main finding from the recent literature on multigenerational persistence, that inequality is far more persistent than previously thought (e.g. Clark, 2014; Barone and Mocetti, 2016). Importantly, our approach is very different from those studies. We use direct family links instead of group-level estimators on the surname level, and our approach does not require observation of socioeconomic outcomes in distant ancestor generations. Our results therefore suggest that the high multigenerational persistence is not just an artifact of the use of surname group-level estimators, as has been speculated in the literature.[6]

Second, our results have implications for the relation between assortative mating and inequality. A recent literature studies if an increased educational assortative mating may explain the rise in economic inequality that many developed countries have experienced since the early 1980s (for example, Greenwood et al. 2014, Eika et al. 2014). But

---

[6] See for example Chetty et al. (2014), Torche and Corvalan (2015), Vosters and Nybom (2017), Braun and Stuhler (2018), Güell et al. (2018), Solon (2018), Adermon et al. (2016), Clark (2018).

while contributing to inequalities between households, trends in educational assortative mating may not affect inequality transmission across generations – unless they reflect assortative trends in the latent factor. Because our results suggest that educational sorting explains only a limited share of the sorting in latent factors, even large shifts in educational sorting may have little implication for assortative mating in those factors that affect child outcomes.

Third, our approach relates to the literature on siblings correlations, which estimates the impact of family background on observable outcomes such as income or education (see Solon 1999, Levine and Mazumder 2007, Björklund et al. 2009). The family background captured by sibling correlations is a latent component accounting for all factors shared by siblings, including factors that are orthogonal to the observed socioeconomic status of parents. Sibling correlations are therefore a more comprehensive measure of family background than parent-child correlations (Jäntti and Jenkins, 2014). However, they quantify the similarity of siblings in *observables*, which may not reflect how similar siblings are in unobservable advantages. By considering more distant relatives in the horizontal dimension, our approach allows us to account for such latent factors. We find that siblings must be far more similar in latent advantages than they are in education or income. Most of the shared influences of siblings are not reflected in observables.

By considering relatives of different degrees of kinship we are able to disentangle the "non-inheritable" part of family background that is only shared by siblings from the "inheritable" part that is also partially shared by more distant relatives, such as cousins or sibling in-lars. Our decomposition of the family background into the inheritable and non-inheritable components is related to the nature and nurture decomposition. Many papers in this literature try to estimate the relative importance of nature and nurture by looking at the correlations in observed outcomes for different type of siblings like MZ twins, DZ twins, siblings, half siblings or adoptees (see Sacerdote 2011 for a literature review). The standard approach in this literature decomposes the total variance of the output of interest into three additive terms, representing genetic factors, environmental factors shared by the siblings, and a third factor that is idiosyncratic to the individuals. Among these papers, the most related to our work are Behrman and Taubman (1989) and Björklund et al. (2005). Both papers make use of correlations across several sibling types and find the values of the parameters that best fit the empirical correlations, in a similar way as we do here. Behrman and Taubman focus on years of schooling and assume that family environment and genes are uncorrelated (see Goldberger (1979) for a critique of this approach). Björklund et al. (2005) focus on earnings and consider different possible models, and allow for the possibility that environment and genes are correlated.

Our approach differs from these works in several fundamental aspects. First, we do not make use of twins, and our analysis is based on correlations for all type of relatives. Second, and more importantly, we focus on intergenerational mobility and ”persistence”, not on the nature-nurture debate. As noticed, we decompose the family background into an *inheritable* and a *non-inheritable* component rather than into genetic and environmental components. The obvious disadvantage of our approach is that we remain largely agnostic about what mechanisms the latent factor and the pathways of our model represent. The principal advantage however is that it provides a more comprehensive account of intergenerational transmission. For example, our latent factor is a more comprehensive object than the “genotype” considered in behavioral genetics, as it captures also non-genetic advantages that matter for the socioeconomic success of the next generation. That is, by avoiding any a-priori stand on the causal channels via which transmission occurs, we capture those channels more completely.[7]

However, one may ask if genes are an important component of the advantages encapsulated in the latent factor of our model. We can investigate this question because the standard model in quantitative genetics is nested by our more general model. We can therefore study to what degree a genetic model can fit the wide array of kinship correlations that we observe in this study. We first consider *body height*, which we observe from military enlistment tests in the Swedish registers. Height is an interesting reference point, because it is less affected by assortative processes (Galton, 1886), and variation in body height is known to be primarily due to the influence of genes (Silventoinen 2003). Indeed, we find that the genetic model fits the kinship correlations in body height nearly as well as our more general model. These results suggest that our model is flexible enough to capture genetic processes, without specifically targeting those processes.

In contrast, the genetic model cannot fit the kinship correlations in educational attainment. Even the in-sample fit is extremely bad, with kinship correlations being heavily overstated for close kins, and heavily understated for more distant kins. Our findings are therefore inconsistent with a purely genetic interpretation. The limitation is the assortative process – spouses must be far more similar to each other than they are in “*phenotype*” educational attainment. The observation of a wide range of kinship moments is key for this finding. We show that the genetic model can fit a small number

---

[7] We also do not have to deal with the complicated problem of the relationship between genes and environment. In our case, by definition, the non-inherited component captures all the effects that siblings share and are not correlated with the inheritable components.

of close kinship moments; only when asked to fit a large set of kinship moments the inadequacy of the genetic model becomes apparent. Of course, the observation that a purely genetic model cannot fit the data is not necessarily inconsistent with the view that genes are an important contributor to the latent factor in our model. To quantify their role more throughly we extend our model by incorporating both a genetic and a non-genetic latent factor. The results from this extended model appear consistent with recent evidence from molecular genetic data, and genome-wide association study (GWAS) for years of schooling (e.g. et al, 2016; Lee et al., 2018).

Our approach is subject to two important limitations. A fundamental problem for the estimation of distributional models is that the moments that we wish to explain may vary over time (Atkinson and Jenkins, 1984). Our calibration relies on the assumption that these moments are in a steady-state equilibrium, as is standard in this literature. Because the outcome for horizontal kins can be measured at approximately the same time, our approach is arguably less sensitive to this assumption than the multigenerational approach based on information across three or even more generations. Still, changes in kinship correlations over time could affect our results. We address those concerns by dropping those moments from our analysis that appear unstable over our analysis period. Another important limitation is that we do not model the fertility process. Again this is standard in the literature, but it might be a concern here because the matching of distant sibling in-laws depends on the existence of a brother or sister, skewing the sample towards families with more children and lower socioeconomic status. We show that this skewness is modest and has only a small effect on the kinship correlations. Still, to limit the potential role of selective fertility we consider sibling in-laws only up to five degrees of separation. Future work could improve on our approach by explicitly modeling selective fertility and the off-steady-state dynamics.

The paper proceeds as follows. Section 2 sets out the basic model and develops our method. Section 3 describes the data. Section 4 presents our baseline empirical findings and robustness checks for Sweden, while Section 5 presents our results for Spain. Section 6 calibrates our model for other outcomes. In Section 7 we discuss the interpretation of our findings, consider alternative models, and study how transmission pattern vary with gender. Section 8 concludes. We include some additional information about the models in the Appendixes.

2 Theory

Our model deviates in three important aspects from the prior literature. First, we allow for *direct* and *indirect* transmission mechanisms. Second, we allow for assortative mating along two distinct dimensions, and account for both parents explicitly. Third, we consider how the strength of the transmission mechanisms vary with the gender of the child and the parent.

2.1 General Model

Suppose that y is a socioeconomic outcome of interest in our economy, such as income or education. We henceforth identify y with years of schooling, the baseline outcome in our empirical exercise. All theoretical implications remain valid when studying other outcomes. We want to study the link of such variable y between individuals and their ancestors.

Specifically, assume that the outcome y for an individual from generation t is given by

$$y_t^k = \beta^k \widetilde{y}_{t-1}^k + z_t^k + x_t^k + u_t^k \tag{1}$$

where the superscript k stands for male ($k = m$) and female ($k = f$). The first component $\widetilde{y}_{t-1}^k$ is the weighted average socioeconomic status of parents,

$$\widetilde{y}_{t-1}^k = \alpha_y^k y_{t-1}^m + (1 - \alpha_y^k) y_{t-1}^f,$$

where $\alpha_y^k \in [0, 1]$. The parameters β^k and α_y^k capture therefore the *direct transmission* of parental on child outcomes. As the importance of this channel may vary with the gender of the child ($\beta^f \neq \beta^m$) and the gender of the parent ($\alpha_y^f \neq \alpha_y^m$), we effectively allow it to be distinct for each of the four parent-child gender combinations.

The *latent factor* z_t^k captures the importance of unobservable determinants of child outcomes that are passed from parents to children (see Section 2.2). As the observable determinant, it depends on the weighted average latent status of the parents $\widetilde{z}_{t-1}^k$,

$$z_t^k = \gamma^k \widetilde{z}_{t-1}^k + e_t^k + v_t^k, \tag{2}$$

where

$$\widetilde{z}_{t-1}^k = \alpha_z^k z_{t-1}^m + (1 - \alpha_z^k) z_{t-1}^f \tag{3}$$

and $\alpha_z^k \in [0,1]$. The parameters γ^k and α_z^k capture the strength of indirect transmission channels, i.e. factors that impact observable outcomes but that are not directly observed themselves. We allow for distinct transmission pattern across all four parent-child gender combinations. Equations (2) and (3) do not necessarily map into one particular (e.g. genetic or behavioral) mechanism, but may represent a great number of underlying mechanisms. Such "reduced-form" representations have been common in theoretical work, including Becker and Tomes (1986), and we discuss its advantages and interpretation below (see Section 2.5).

Finally, the model includes three types of shocks. The individual component u_t^k in the observed outcome is a white-noise error term. The *sibling component* x_t^k is shared by siblings of the same gender, can be correlated across siblings of different gender, and is uncorrelated with the other variables (in particular with z_t and y_{t-1}). Similarly, the error term v_t^k in the latent factor is a white-noise error term, and the sibling component e_t^k is shared by all siblings of the same gender and can be correlated across siblings of different genders. Allowing for shared influences among siblings over and above the parental influence, and accounting for them in a flexible way, allows us to extend our analysis in the horizontal dimension (see Section 2.4).

We allow for assortative mating both in the observable and latent socioeconomic status (see Section 2.3).[8] In particular, we consider the linear projections of z_{t-1}^f and y_{t-1}^f on z_{t-1}^m and y_{t-1}^m:

$$\begin{pmatrix} z_{t-1}^f \\ y_{t-1}^f \end{pmatrix} = \begin{pmatrix} r_{zz}^m & r_{zy}^m \\ r_{yz}^m & r_{yy}^m \end{pmatrix} \begin{pmatrix} z_{t-1}^m \\ y_{t-1}^m \end{pmatrix} + \begin{pmatrix} w_{t-1}^m \\ \varepsilon_{t-1}^m \end{pmatrix} \tag{4}$$

where w_{t-1}^m and ε_{t-1}^m might be correlated but are uncorrelated with z_{t-1}^m and y_{t-1}^m, and the r_{sd}^m $(s,d=y,z)$ coefficients are functions of the following correlations and standard deviations $\rho_{z^m y^m}, \rho_{z^m z^f}, \rho_{z^m y^f}, \rho_{y^m z^f}, \rho_{y^m y^f}, \sigma_{z^m}, \sigma_{z^f}, \sigma_{y^m}$ and σ_{y^f}. Alternatively, we can consider the linear projections of z_{t-1}^m and y_{t-1}^m on z_{t-1}^f and y_{t-1}^f:

$$\begin{pmatrix} z_{t-1}^m \\ y_{t-1}^m \end{pmatrix} = \begin{pmatrix} r_{zz}^f & r_{zy}^f \\ r_{yz}^f & r_{yy}^f \end{pmatrix} \begin{pmatrix} z_{t-1}^f \\ y_{t-1}^f \end{pmatrix} + \begin{pmatrix} w_{t-1}^f \\ \varepsilon_{t-1}^f \end{pmatrix} \tag{5}$$

where w_{t-1}^f and ε_{t-1}^f might be correlated but are uncorrelated with z_{t-1}^f and y_{t-1}^f, and the r_{sd}^f $(s,d=y,z)$ coefficients are functions of the following correlations and standard deviations $\rho_{z^f y^f}, \rho_{z^m z^f}, \rho_{z^m y^f}, \rho_{y^m z^f}, \rho_{y^m y^f}, \sigma_{z^m}, \sigma_{z^f}, \sigma_{y^m}$ and σ_{y^f}, where we drop the subscript t based on the steady-state assumption that all second moments are constant

---

[8] See Behrman and Rosenzweig (2002) for a related model with assortative mating in two dimensions.

across generations. In Appendix A we provide the formulas for all these coefficients and we show that $\rho_{z^m y^m}$ and $\rho_{z^f y^f}$ are functions of the other parameters through two steady state equations.

2.2 Direct and Latent Transmission Channels

The model incorporates both direct (via observables) and indirect (via latent variables) transmission channels. That part of the transmission process may be inherently unobservable has long been recognized in the literature. Goldberger (1972) describes how latent factors such as "ambition" played a central role in the early sociological research, such as Duncan (1969). The latent variable in our model is closer in spirit to Becker and Tomes (1979), who assume that a person's *"endowment"* represent a great variety of cultural and genetic attributes.

Despite their role in the theoretical literature, latent transmission channels have received less attention in empirical work. Solon (1999) and Black and Devereux (2011) summarize the many dimensions in which empirical work has progressed in recent decades, but interest has centered on mobility in observable characteristics (as captured by the intergenerational elasticity of lifetime earnings), or the strength of particular causal channels (such as the causal effect of parental education on child education).

An exception is the recent literature on *multigenerational* transmission, which describes how economic inequalities persist across multiple generations. A key finding is that inequalities appear more persistent than indicated by traditional parent-child correlations. Studies based on historical records demonstrate that inequalities *between surnames* persist across centuries (Clark and Cummins 2012, Clark 2014, Barone and Mocetti 2016). And while name-level estimators have been criticized, the evidence from individual-level estimators points likewise to high persistence (see Lindahl et al. 2015, Dribe and Helgertz 2016, Braun and Stuhler 2018, Adermon et al. 2018, Long and Ferrie 2018, Neidhöfer and Stockhausen, forthcoming). Purely direct transmission mechanisms cannot explain this pattern (Clark and Cummins, 2012).

The existence of Latent transmission mechanisms would rationalize these findings (Clark and Cummins, 2012).[9] If the latent variable is comparatively persistent across generations ($\gamma^k > \beta^k$), but explains only part of the inequality in socioeconomic outcomes

---

[9] Other potential rationalizations include the idea that intergenerational transmission occurs via multiple channels of varying rates of persistence (see Stuhler, 2012), or that it does not follow a Markov process. In particular, a recent literature on "grandparent effects" aims to quantify the causal contribution of grandparents on their children (see Anderson et al., 2018). Intuitively, this literature agrees with

($\sigma^2_{z^k} < \sigma^2_{y^k}$), then the parent-child correlation may greatly understate the actual transmission of advantages or how this transmission varies across groups, areas and time.[10] This observation is related to the insight that *sibling correlations* are a more comprehensive measure of the importance of family background than parent-child correlations (see Jäntti and Jenkins 2014). [11]

Moreover, data across more than two generations can be used to identify the parameters of a transmission model, including the inherently unobserved component. This potential has been noted already by Becker and Tomes (1979) and Goldberger (1989), and has been exploited in the recent multigenerational literature. Clark and Cummins (2012) use surname averages across two or more generations to identify the underlying rate of persistence in a simple latent factor model. Braun and Stuhler (2018) illustrate that this model can be also identified from direct family linkages across three or more generations.

However, these multigenerational studies have been based on a stylized model, in which intergenerational transmission occurs exclusively via the latent factor, without any role for direct transmission mechanisms (i.e. $\beta^m = \beta^f = 0$ in our model) or shared influences among siblings (i.e. $\sigma^2_x = \sigma^2_e = 0$). We aim to account for latent transmission mechanisms without imposing such restrictions, and tp quantify the importance of both direct and indirect transmission channels.

2.3 Assortative Mating

Most intergenerational studies consider a simplified one-parent family structure, as often the assortative mating process is not directly relevant. But the assortative process is fundamental to understand the recent multigenerational evidence, as high rates of persistence across generations requires strong assortative mating (Diaz-Vidal and Clark 2015, Clark 2017, Braun and Stuhler 2018). The intuition for this argument follows from equations (2) and (3), which suggest that the father's and mother's latent status z^m_{t-1} and z^f_{t-1} can only be both strongly correlated with their child's status z^k_t if they are also strongly correlated with each other ($0 \ll \rho_{z^m z^f}$).

Spousal correlations in socioeconomic outcomes such as income or education are typically between 0.4 and 0.6 (e.g. Fernández and Rogerson 2001, Ermisch et al. 2006,

---

the diagnosis (a missing component in the transmission process) but considers a more specific solution (considering a missing person, i.e. grandparent) than the approach that we follow here (considering a *latent factor*).

[10] Intuitively, the observable socioeconomic status is only an imperfect proxy for status, a type of measurement error that attenuates the traditional parent-child correlations.

[11]

Greenwood et al. 2014). But as we show below, assortative mating needs to be higher to rationalize the socioeconomic similarity between distant kins. Spousal correlations in observable characteristics appear too low to explain multigenerational dependence, or the pattern of dependence across horizontal kins that we present in this paper.

We rationalize this discrepancy by allowing spouses to be similar not only in observable but also *unobservable* characteristics that determine the socioeconomic status of their offspring. In particular, spouses may be more similar to each other in unobservables (z_{t-1}^k in our model) than observable outcomes (y_{t-1}^k) – spousal correlations in socioeconomic outcomes may reflect only a superficial similarity, and not the effective degree of assortative mating in more fundamental characteristics that determine the success of future generations. Our data provide an opportunity to test this hypothesis.[12]

The assortative process is therefore a key component to rationalize the pattern of dependence in our kinship data, but the latter will also be informative about formerly unknown aspects of the former. We therefore model the assortative process in more detail than the previous literature. We allow for assortative mating along two dimensions, with the spousal correlation in the socioeconomic outcome ($\rho_{y^m y^f}$) potentially differing from the spousal correlation in the latent status ($\rho_{z^m z^f}$). Moreover, we do not impose assumptions on the relative importance of each parent, and allow for all transmission processes to vary with parent and child gender.

2.4 Horizontal Kinship

The model described in Section 2.1 is comparatively general, and therefore challenging to estimate. It allows for (i) both direct and latent transmission mechanisms, (ii) a two-parent structure and assortative mating in observable and unobservable characteristics, and (iii) asymmetries in the strength of these mechanisms that vary with both parent and child gender. We can directly estimate σ_{y^m} and σ_{y^f} from the data, and therefore have 20 unknown parameter. With at least 20 empirical moments required, this model is too complex to be identified from inter- or multigenerational moments alone, which has been the traditional approach (see e.g. Atkinson and Jenkins 1984).

We therefore switch from from this "*vertical*" to a "*horizontal*" approach, considering information about relatives in the same generation (such as siblings or cousins). Because

---

[12] Ermisch et al. (2006) shows the potential of this approach, illustrating how the degree of assortative mating in one-dimensional matching model with a latent human capital variable can be identified from data on children, their parents, and their parents-in-law.

siblings may share influences over and above the parental influence, we model these influences in a flexible way. We allow for shared sibling components in both observable and unobservable characteristics, and for the distribution of those components to vary across gender combinations. Importantly, by modeling both assortative and sibling processes we can also consider sibling in-laws in our analysis. The horizontal approach offers important advantages in terms of data quality, feasibility and scope:

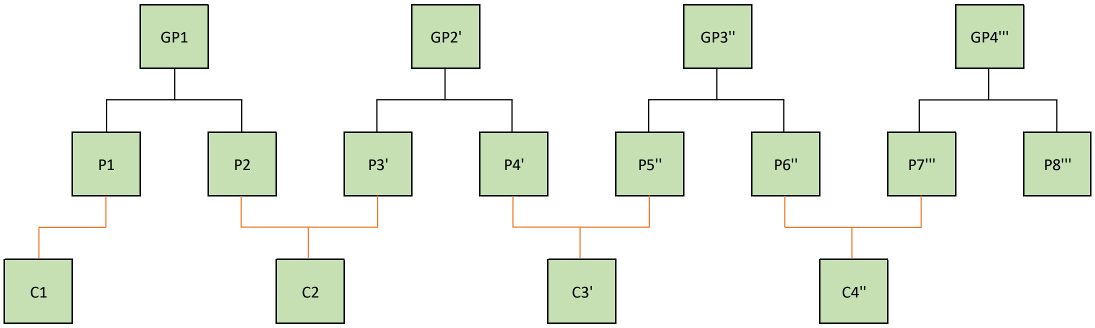
First, studies of multigenerational persistence have to compare socio-economic outcomes from distant generations. As income is rarely observed in historical sources, this problem typically boils down to making occupational outcomes more comparable over time – a difficult task, which has received considerable debate (see e.g. Long and Ferrie, 2013b; Modalsli, 2017). This problem is much diminished when considering horizontal kins, for whom socio-economic outcomes can be measured at approximately the same age and time. We also avoid problems related to recall bias, or the comparison of data from different sources.

Second, the horizontal approach remains feasible in settings in which there is only limited information on ancestors available. This benefit has long been recognized in the extensive literature on sibling correlations (Solon et al. 1991, Levine and Mazumder 2007, Björklund et al. 2009, Jäntti and Jenkins 2014). Sibling correlations can be estimated even when parental outcomes are not or not well observed, and therefore require less data than intergenerational measures. The horizontal perspective therefore allows us to estimate our model in settings in which the intergenerational dimension of the data is restrictive, such as the Spanish sources discussed below.

Third, and most importantly, the incorporation of horizontal kins yields a much greater set of empirical moments than a purely vertical approach. With better data becoming available, the literature has started to document the pattern of socioeconomic inequalities beyond the nuclear family. For example, Hällsten (2014) studies kinship correlations between first and second cousins in Sweden, while Adermon et al. (2016) provide evidence across a broad range of kinship relations, including horizontal kins. We aim to demonstrate that such evidence is valuable not only from a descriptive perspective, but that it also opens the door for the identification of more detailed intergenerational models – providing a deeper understanding of transmission processes within the nuclear family and the assortative process.

We can distinguish several dozen *consanguine* ("blood") kinships in our data, but extend our analysis also to *affinity* ("in-law") kinships such as siblings in-law. While they may not descend from a common ancestor, affine kins have similar informational value as

Figure 1: The Identification of Sibling In-Laws



consanguine kins – in particular, both are a function of and therefore informative about intergenerational and assortative processes (see Appendix A). The key advantage of affine relations is however that they can be traced over exceptionally long "distances".

Figure 1 illustrates the logic in a hypothetical family tree across a child (C), parent (P) and grandparent (GP) generation. The identification of more distant consanguine kins necessitates the observation of more distant ancestors. For example, the identification of cousins (e.g. C1-C2) requires the observation of their shared grandparents (GP1), while the identification of second-degree cousins would require the identification of their great-grandparents. The consequence of this hard limitation is that very distant kins are never directly observed. Instead, the available evidence on long-run mobility is based on historical sources and the probabilistic approximation of family links via *surnames* (e.g. Clark 2014, Barone and Mocetti, 2016).

In contrast, affine relationships are defined only via spousal and parental links – irrespective of their degree of separation. In a first step, we identify a person's sibling (e.g. P1-P2) via their shared parent (GP1). In a second step, we identify the siblings' spouse via their shared descendant (C2). Implemented once, these steps identify a pair of *siblings in-law* (P1-P3'). Implemented twice, we identify the *sibling-in-law of the sibling-in-law* (P1-P5"), and so on. In population-wide data in which *every* spouse and sibling is observed, one can reiterate these linkages – and therefore the number of empirical moments – ad infinitum.

In this study we observe more than one third of the Swedish population, and sizable samples for more than 160 different types of kinships up to the fifth-order affinity relations. We systematically consider all kinship types, such as *siblings in-law* (e.g. P1-P3'), the *sibling of the siblings in-law* (P1-P4'), or *spouses of the sibling of the siblings in-laws* (P2-P5"), as well as vertical relations such as the uncles in-law (C1-P3') and cousins-law (C1-C3'). The observation of distant affine relations will play a key role, as in distant kins it becomes directly discernible if a model can fit the pattern of economic inequalities across generations.

The horizontal perspective yields therefore many more and much more distant family relations than the vertical perspective. As a result, the model from Section 2.1 is heavily over-identified. Over-identification helps us to pin down the parameters of the model and, more importantly, test the fit of our model more systematically than what has been possible in the prior literature. Specifically, we test if our results remain robust to considering different types of kins, if alternative models can provide a similarly good fit to the data, and if our model can fit the pattern of socioeconomic inequality out-of-sample.

2.5 Interpretation

Our model is causal in the sense of capturing the causal effect of family background. However, we do not impose any specific – such as genetic or cultural – interpretation of its parameters. At the outset, our objective is instead a statistical one – to fit the pattern of economic inequalities across a wide range of kinship types by formulating a model that is sufficiently flexible along a number of central dimensions.

In contrast, most existing work either focuses on simpler descriptive measures (such as intergenerational correlations) or specific causal channels.[13] For example, a large literature in both behavioral genetics and economics assesses the role of genes, using targeted research designs that exploit variation in the presence of genetic mechanisms (as in twin or adoption studies, e.g. Björklund et al., 2006). Other strands of the literature study the intergenerational causal effect of parental education, income or wealth (e.g. Holmlund et al., 2011), neighborhoods and peers (e.g. Chetty and Hendren, 2018), or institutions (e.g. Havnes and Mogstad, 2015).

Our approach has distinct advantages and disadvantages compared to more targeted approaches. The obvious disadvantage is that we have to remain largely agnostic about what mechanisms the pathways of our model represent. This indeterminacy becomes most obvious with respect to the latent factor z_t^k, a catch-all object that may represent

---

[13] Standard theoretical models in economics often represent a mixture, featuring similar indeterminacy as our model in some of its aspects but targeting specific mechanisms in others. For example, Becker and Tomes (1979) focus on the effect of parental investments on child outcomes, but also consider the role of "endowments" that have a similarly broad interpretation as the latent factor in our model.

a great variety of cultural, behavioral and genetic attributes. The principal advantage of our approach however is that it provides a more *comprehensive* account of intergenerational transmission. For example, our latent factor z_t^k is a more comprehensive object than the "genotype" considered in behavioral genetics literature, as it captures also nonhereditary advantages that matter for the socioeconomic success of the next generation. That is, by avoiding any a-priori stand on the causal channels via which transmission occurs, we capture those channels more completely.

Our approach shares aspects with the literature on siblings correlations, which estimates the impact of family background on income, education, or other observable outcomes (Solon et al. 1991, Levine and Mazumder 2007, Björklund et al. 2009). Because they account for all factors shared by siblings, including latent factors that are orthogonal to the observed status of parents, sibling correlations are a more comprehensive measure of family background than intergenerational correlations (Jäntti and Jenkins, 2014). However, sibling correlations still capture only those advantages that are directly observable in the child generation. In contrast, our approach captures also advantages that are *not* visible in the child generation, but that nevertheless affect a family's prospects in future generation (i.e. the sibling correlation in *latent* advantages).[14]

A central question in the literature is the extent to which advantages are transmitted from one generation to the next. Different schools of thought offer very different views (Piketty 2000), and the empirical evidence on this question has changed drastically over recent decades. In comparison to earlier studies (see Becker and Tomes 1986), a more careful treatment of measurement issues nearly tripled estimates of income persistence in the United States (Solon 1999). However, Clark (2014) and other multigenerational studies imply that such estimates still drastically understate the transmission of advantages. This controversial hypothesis can be studied separately from the question which particular pathways contribute to the overall rate of persistence.

While our primary objectives are statistical ones, indirectly our approach does become informative about the underlying mechanisms of intergenerational transmission. We can test certain causal interpretations, because they have specific statistical implications with respect to the pattern of inequality across kins. In particular, the model described in Section 2.1 is sufficiently general to test, among other models, a standard genetic

---

[14] Moreover, we can decompose the sibling correlation into components that affect the socioeconomic success of future generations, and those components that do not. Only some of the factors shared by siblings will be transmitted within families (and therefore also affect cousins, second cousins, and so on). By comparing different degrees of kinship, we are able to disentangle these two components. Our approach provides therefore a more comprehensive, and a more detailed account of the importance of family background.

model of transmission. We therefore test how a purely genetic model fits the data compared to our more general model.

3 Data and Calibration

We describe our data sources and estimation procedures in this section. We will calibrate the model from Section 2.1 for two different countries, *Sweden* and *Spain*. The comparison will contribute to the literature on mobility differences across countries, which contains little evidence on Spain (see Black and Devereux, 2011). More importantly for our purposes, we can show how different components of the transmission process contribute to those observed differences in mobility levels.

The comparison between Swedish and Spanish sources is interesting also from a methodological perspective, to demonstrate the feasibility of our approach in different settings. To apply our methodology we need data on extended families. The Swedish registers are very extensive, include family links over multiple generations, and cover a large share of the population. In contrast, the Spanish data are limited to a single cross-section, contain only educational outcomes and names, and lack direct family links to define kinship.

Our baseline outcome is *educational attainment*, which is observed for both countries. In the Swedish registers we consider two additional outcomes, *income* and *height*. We study variation in the overall rate of persistence, but also if certain components – such as the role of assortative mating in observable or latent factors – are more important for some outcomes than others. Because strong external evidence exist on some of these components, the comparison serves as a sanity check of our approach, and will link our findings to research designs that aim to quantify the role of genetic factors.

3.1 Swedish Multigenerational Registers

Our sample is based on a random 35 percent draw of the Swedish population born between 1932 and 1967, as well as their biological parents, siblings and children. Family links are biological links, with a man and woman considered to be spouses if they have a child together. We match data on individual characteristics from bi-decennial censuses (starting from 1960), official registers, and military enlistment tests.

Variables

Education. Educational registers were compiled in 1970, 1990, and about every third year thereafter, containing detailed information on each individual's educational attainment. We consider the highest schooling level recorded across these years, and translate it into years of education, with 7 years for the old compulsory school being the minimum, and 20 years for a doctoral degree the maximum.[15] Educational information in 1970 is available only for those born 1911 and later. It may be missing also if parents had died or emigrated before 1970, but the share of affected observations is small. As the data are collected from official registers they are not subject to standard non-response problems. Our data contain records up to 2007, and therefore reliable information on educational attainment up to cohorts born in the late 1970s, after which it becomes less reliable at the top of the attainment distribution.

Income. We construct a measure of long-run income status by averaging over multiple annual incomes, which are observed for the years 1968-2007. We use total (pre-tax) income, which is the sum of an individual's labor (and labor-related) earnings, early-age pensions, and net income from business and capital realizations, and express all incomes in 2005 prices. Incomes for parents are necessarily measured at a later age than incomes for their offspring, which may bias estimates of the intergenerational correlation of lifetime income. To reduce this bias we construct ten-year average incomes measured at age 30-39 for children and age 45-54 for parents.[16] To reduce the influence of outliers we further censor those income averages, separately for each child cohort, at the 1st and 99th percentiles. For comparison we estimate our model also based on 5-year average incomes (measured at age 30-34 for children and age 45-49 for parents) and annual incomes (measured at age 32 and age 50, respectively).

Height. We observe height from military enlistment data, for male individuals born between approximately 1950 and 1980. Military enlistment took place at age 18 or 19 and was at the time universal for all men. Height was recorded as part of the medical examination. Because we observe height only for birth cohorts spanning three decades, we can consider parent-child and other vertical correlations only for parents who were

---

[15] In the 1970 Census, we impute 7 years for the (old) primary school, 9 years for (new) compulsory schooling, 9 years for post-primary school (realskola), 11 years for short high school, 12 years for long high school, 14 years for short university, 16 years for long university, and 20 years for a PhD. Schooling levels are recorded in more details in later registers.

[16] Nybom and Stuhler (2017) study the magnitude of attenuation and life-cycle biases from the approximation of long-run income with short average incomes in the same data source. The magnitude of these biases are small in our chosen age range.

sufficiently young at the birth of their child. We show below that such selectivity with respect to parental age is less problematic for height than for other outcomes.

Preparation. Kins can be born in different cohorts, or their outcomes being measured in different years. To abstract from this source of variation we de-mean all outcomes, separately by gender and birth cohort. These preparatory steps are performed in the full sample, before selecting kinship pairs. Income ranks are defined separately within each birth cohort and gender.

Cohort Selection

We next select subsamples for each kinship moment. This step is non-trivial, as multiple sources of selection need to be taken into account, separately for each outcome, and separately for horizontal and vertical kinship types. We first select cohorts for which the *outcome* is reliably observed, as described above. We then assess which *kinship types* can be reliably identified within those cohorts. For example, the identification of siblings requires observation of their parents, while for the identification of cousins we need to observe grandparents, and so on. In principle, our data contain family links up to four generations. However, we are more likely to observe great-grandparents if individuals became parent at a comparatively young age, which could introduce a selection bias in our estimates (see also Hällsten 2014). We therefore abstain from kinship types that depend on the identification of great-grandparents, such as second-degree cousins, and consider only cohorts for which three generations are reliably observed.

Finally, we minimize selectivity with respect to the *age difference* between kins. For example, parent-child correlations tend to be lower for young parents, for whom the age difference to their children is small. We avoid this problem by considering a range of child cohorts that is sufficiently narrow, for whom parental outcomes are observed irrespective of parental age-at-birth. Horizontal moments, such as sibling and spousal correlations, differ less systematically with difference in age. However, the issue could become more severe for the very distant in-law moments, because the age difference between kins tends to increase with the degree of separation – i.e. siblings tend to be more closely spaced than cousins. We avoid this problem by considering a broad range of cohort for all horizontal moments.

Intergenerational and assortative processes may vary across birth cohorts, which represents a problem for the analysis of distributional models (see Section 4.6). We address this issue in two ways. First, we select similar cohort ranges for different kinship types, as far as possible given the other constraints mentioned above. Second, we examine explicitly if kinship correlations change over birth cohorts during our period of analysis. Most problematic in this respect are cousins, because the requirement to identify grandparents forces us to consider more recent birth cohorts (such that grandparents are observed irrespective of parental age-at-birth). We return to this issue below.

A final issue arises with respect to very distant kins, as the long chain of siblings and spouses connecting them may contain duplicate entries. For example, an individual may be his own second-degree brother in-law if two families are connected via more than one spousal link, or a person may feature multiple times because he or she has children with different partners. In principle, the observation of duplicate entries are a reflection of the assortative process, and should be retained in the analysis. For example, inequality would be more persistent if in-law relations "circle" within groups defined by geography, race, or other characteristics. However, because we observe only a subset of the Swedish population, duplicate entries occur at a higher rate in our sample than in the full population. We therefore drop all chains with duplicate entries, which has however only a small effect on the most distant kinship correlations.

3.2 Spanish Census Data

The 2001 population census for Spain, which is available nationwide, does not allow to identify families unless they are living in the same household. However, for the Spanish region of *Cantabria* we obtained information on the full name of each person, and we can use this information to identify parents and children. The Census also reports, among other variables, the gender, age and educational level of all individuals living in the region ($526,339$ persons). We define the t-generation as all persons born in Cantabria between 1956 and 1976 ($71,479$ males and $68,830$ females) and the $(t-1)$-generation as their parents.

Matching

Surnames in Spain are passed from parents to children according to the following rule: A newborn person, regardless of gender, receives two surnames that are kept for life. The first surname is the father's first surname and the second the mother's first surname. This name convention allows us to identify fathers and mothers. For each person i in generation t we define the set of potential parents as all the couples born before 1956

Table 1: Descriptive Statistics in Spanish Census

	Men				Women			
	Matched		Unmatched		Matched		Unmatched	
	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev	Mean	Std. Dev
Age	33.61	5.91	35.42	6.16	33.70	5.92	35.50	6.15
Years of schooling	10.53	3.71	9.71	3.64	10.99	3.71	10.11	3.69
Observations	25,860		45,619		24,610		44,220	

such that the husband first surname coincides with person i first surname and the wife first surname coincides with person i second surname. Then, we say that we identify the parents if there is only one couple in the set of potential parents and the age difference between both parents and the child is at least 16 years. We identify the parents for $25,860$ males and $24,610$ females which is approximately 36.2% and 35.8% of the male and female population, respectively.

To assess how well our strategy to identify parents and children works, we exploit the fact that we can directly identify parents and children when they live together (without using surnames). We use this information to estimate the percentage of incorrect matchings derived from our identification strategy. We identify 51,923 pairs of parents-child using the surnames, 23,694 of these pairs are living together and 28,229 are living in different households. For the sub-sample of parents-child living together, the percentage of identification mistakes is 6.1%. We exclude these 1,453 pairs from our sample and the final sample size is 50,470. If the percentage of incorrect identifications for the sub-sample of parents-child not living together were also 6.1% we would expect 1,722 mistakes (3.4%) in the total sample.

Once we have identified parents and children, siblings are immediately identified, and when children are married we also identify siblings in-law. Finally, we identify siblings in the parents generation when there are at most four individuals in the over 25 population sharing the same two surnames. Once siblings in the parents generation are identified, uncles and nephews, and cousins are immediately identified. Again it is important to estimate how well our strategy to identify siblings in the parents' generation works. We repeat the same exercise as above but restricting the sample to children with surnames bear by between two and four individuals in the over 25 population. As expected, the percentage of incorrect identifications is now lower, 2.5%.

Variables

Education. We use the information on each individual's educational attainment and convert it to years of schooling following Calero et al. (2007).[17] We measure the years of schooling as deviations from the corresponding mean in each generation. Table 1 shows some basic descriptive statistics. The matched sample is almost two years younger than the unmatched one. The reason is that the older a person is, the more likely the parents are not living together or one of them has died. Since the matched sample is younger it is also more educated (0.8 more years of schooling than the unmatched sample).

3.3 Estimation and Calibration

We estimate each moment based on the sample restrictions described above. Because the number of family members varies across families we need to decide how to weight large compared to small families. A family or "cluster" is defined by the most recent common ancestor (such as the common grandparents shared by cousins) for biological kins, or by the linking spouse for in-laws. We considered four different sets of weights, ranging from uniform to weights that are proportional to the number of kinship pairs per family (see Solon et al. 2000). The sample correlations are not very sensitive to this choice, even though the number of kinship pairs per family varies strongly for distant kins (e.g. the number of cousins varies more strongly than the number of siblings).[18] We therefore picked an intermediate scheme, weighting each family by the square root of their number of distinct pairs.

Since we can directly estimate σ_{y^m} and σ_{y^f} from the data, we have 20 unknown parameters that we write as the vector v,

$$\begin{aligned} v \quad = \quad & \{\beta^m, \gamma^m, \sigma_{z^m}, \sigma_{x^m}, \beta^f, \gamma^f, \sigma_{z^f}, \sigma_{x^f}, \rho_{x^m x^f}, \rho_{z^m z^f}, \rho_{y^m y^f}, \rho_{z^m y^f}, \rho_{y^m z^f}, \\ & \alpha^m_y, \alpha^m_z, \alpha^f_y, \alpha^f_z, \sigma_{e^m}, \sigma_{e^f}, \rho_{e^m e^f}\}, \end{aligned}$$

and therefore we need at least 20 correlations between relatives of different kinship to calibrate these parameters. We calibrate the parameters in v by solving the following

---

[17] We assign 2 years of education to those who did not complete primary education, 5 years to primary education, 8 to compulsory education, 10 to vocational training, 12 to secondary education, 15 to sort university degrees, 17 to long university degrees other than engineering and medicine, 18 for engineers and medical doctors and 19 for a Ph.D. All our results are robust to other reasonable ways to assign years of education as, for example, assigning 0 years of education to those who did not complete primary education, 4 years to primary education, 9 to vocational training and 11 to secondary education.

[18] The correlation between the set of sample moments estimated under the two most extreme weighting schemes is greater than 0.99 (0.98) in the Swedish (Spanish) sample.

minimization problem,[19]

$$Min_{v\in F} \quad \sum_{i\in C} p_i(\rho_i - \overline{\rho}_i)^2, \tag{6}$$

where ρ_i are the theoretical correlations, $\overline{\rho}_i$ the empirical correlations, p_i the weight given to each term, F is the set of feasible values for the unknown parameters, and C denotes the set of correlations.[20] In our benchmark case the set C will contain 105 different kinship correlations. In most cases we give the same weight to all the terms so that $p_i = 1$ for all $i \in C$, but the results are very similar if we weight each moment by the number of families used to calculate the sample correlation.

4 Kinship Correlations in Sweden

In this section we report our baseline results for Sweden, considering years of schooling as our dependent variable. Table 2 lists the kinship types that we can distinguish in our data, as well as the number of moments within each group. Considering siblings in-law up to the five degrees of separation, we observe 141 distinct moments.[21] The formulas for each of these correlations as a function of the model parameters are presented in Appendix A.

4.1 Sample Moments

Table 3 reports the estimated (sample) correlation in years of schooling for each 141 kinship moments. The moments are sorted by kinship type, from closely related to more distant kins. Columns (1) and (2) report the number of pairs and sample correlations. The pairs are weighted inversely by the square root of family size, as described in Section 3.3. The sample correlations in years of schooling span between one half for close kins,

---

[19] We have used Mathematica 11.3 to solve the minimization problem. The code is in the [Online Appendix]. We have used the Simulated Annealing algorithm, which is a stochastic function minimizer. In most exercises we have used a minimum of 10,000 random starting points from the set of feasible values F. In most of our main exercises, and in particular in our benchmark case, we reach the same minimum for most of the starting points, so that we are confident that we have found a global minimum. We also tried other algorithms for constrained global optimization (Nelder-Mead, Differential Evolution and Random Search) and never found a different global minimum.

[20] The parameters $\beta^m, \gamma^m, \beta^f, \gamma^f, \alpha_y^m, \alpha_z^m, \alpha_y^f, \alpha_z^f$, have to be between 0 and 1, but the correlations can take negative values.

[21] We observe 205 moments, but some moments coincide. For example, to consider the spouse of a man's brother in-law or to consider the spouse of a woman's sister in-law yields the same exact set of pairs. Because the matching of distant kins involves multiple steps, we estimated all 205 moments and used the duplicate moments to check the veracity of our matching procedures.

Table 2: List of Kinships

	kinship	kinship type	# correlations
a–x	spouses	direct, horizontal	1
x–b	siblings	direct, horizontal	3
ax–by	cousins	direct, horizontal	10
ax–a	child-parent	direct, vertical	4
ax–b	child-uncle/aunt	direct, vertical	8
a–b	siblings in-law (degree 1)	affinity, horizontal	4
a–y	spouse of sib-in-law (dg 1)	affinity, horizontal	3
x–c	sibling of sib-in-law (dg 1)	affinity, horizontal	4
a–c	siblings in-law (degree 2)	affinity, horizontal	8
a–z	spouse of sib-in-law (dg 2)	affinity, horizontal	4
x–d	sibling of sib-in-law (dg 2)	affinity, horizontal	10
a–d	siblings in-law (degree 3)	affinity, horizontal	16
...	...	affinity, horizontal	...
ax–y	child-sibling in law (dg 1)	affinity, vertical	8
...	...	"	...

such as spouses or parents, to only a fraction of that for the most distant kinship types. Owing to the large number of observations, all correlations are precisely estimated. As far as they overlap, they appear consistent with estimates from the previous literature.[22]

4.2 Calibrated Moments

In our baseline calibration we include siblings in-law up to three degrees of separation, but do not include cousins or higher-order siblings in-laws. With these restrictions our baseline calibration is based on 105 distinct kinships, grouped into fourteen different kinship types. We calibrate the model as described in Section 3, and report the predicted moments as well as the percentage deviation between the observed and predicted moments in columns (3) and (4) of Table 3. Moments that were not included in the calibration are printed in italics to distinguish the in- vs. out-of-sample fit of the model.

Figure 2 illustrates the in-sample fit graphically, by plotting both sample moments (orange) and predicted moments from the calibrated model (blue dots). The model explains the data well, for both vertical and horizontal moments, and for both direct ("blood") and affinity ("in-law") kinships. The mean absolute error across all moments used in the calibration is 3.x [update] percent, and across all moments 4.5 percent.

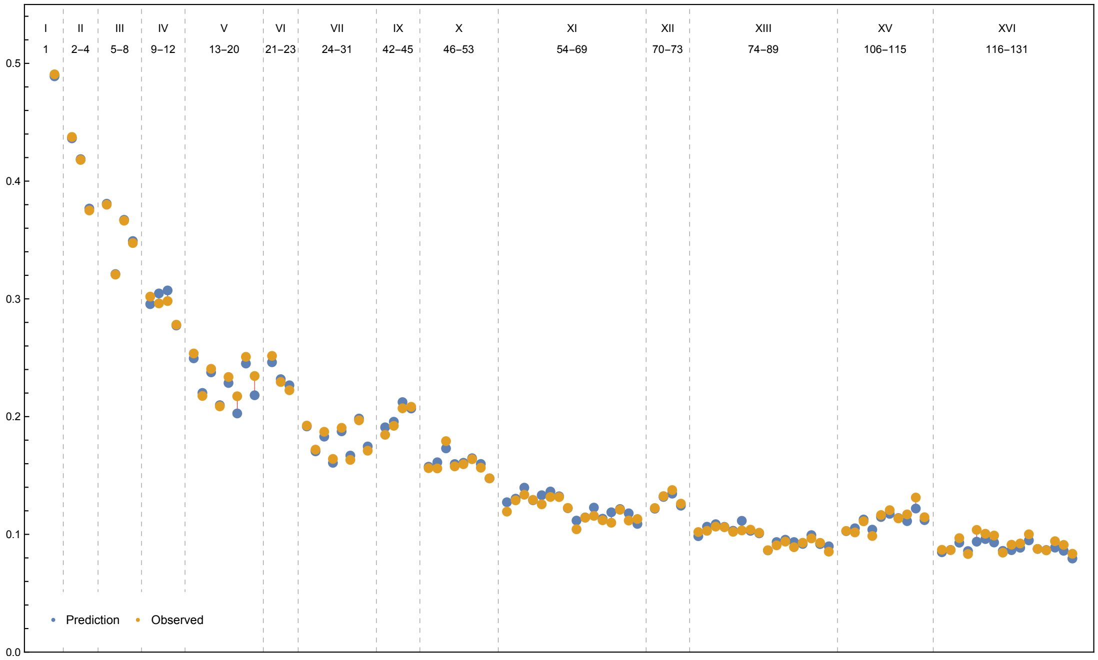
---

[22] For example, Björklund et al. (2009) report that the correlation in years of schooling between brothers in Sweden is slightly below 0.5 (cf. 0.44 in our sample), and Hällsten (2014) estimates that the corresponding correlation for cousins is about 0.15 (slightly below our estimates).

Table 3: Estimated and Calibrated Moments in Swedish Registers

#	##	name	number of pairs (1)	sample correlation (2)	predicted correlation (3)	percent error (4)
I	1	HUSB-WIFE	399,861	0.491	0.489	-0.3
II	2	BROTHERS	49,327	0.438	0.436	-0.3
	3	SISTERS	44,924	0.418	0.419	0.2
	4	BRO-SIS	87,548	0.375	0.377	0.4
III	5	FATH-SON	320,020	0.380	0.381	0.2
	6	FATH-DAUG	306,933	0.321	0.321	0.1
	7	MOTH-SON	342,038	0.366	0.367	0.2
	8	MOTH-DAUG	327,809	0.347	0.349	0.5
IV	9	BL-HS	379,000	0.302	0.296	-2.1
	10	BSL-WB	356,447	0.296	0.305	2.9
	11	BSL-HS	392,641	0.298	0.307	3.0
	12	SL-WB	366,584	0.278	0.277	-0.2
V	13	NEU-BF	177,515	0.254	0.249	-1.6
	14	NIU-BF	172,660	0.218	0.220	1.2
	15	NEU-BM	198,086	0.241	0.238	-1.2
	16	NIU-BM	191,862	0.209	0.210	0.5
	17	NEA-SF	182,561	0.234	0.229	-2.2
	18	NIA-SF	176,859	0.217	0.203	-6.7
	19	NEA-SM	209,942	0.251	0.245	-2.3
	20	NIA-SM	203,208	0.234	0.218	-7.0
VI	21	BL-WWS	177,884	0.252	0.246	-2.2
	22	SL-HHB	160,453	0.229	0.232	1.0
	23	BSL-HWBS	287,866	0.222	0.227	1.9
VII	24	NEWUBF	111,634	0.192	0.192	-0.4
	25	NIWUBF	107,661	0.172	0.170	-0.9
	26	NEWUBM	125,802	0.187	0.183	-2.2
	27	NIWUBM	120,767	0.164	0.161	-2.0
	28	NEHASF	111,486	0.190	0.188	-1.6
	29	NIHASF	106,771	0.163	0.167	2.2
	30	NEHASM	128,352	0.197	0.198	0.8
	31	NIHASM	123,120	0.171	0.175	2.1
VIII	*32*	*MC-B*	*31,353*	*0.167*	*0.159*	*-4.8*
	33	*MC-S*	*36,602*	*0.172*	*0.160*	*-6.8*
	34	*MC-BS*	*62,210*	*0.161*	*0.151*	*-6.0*
	35	*FC-B*	*29,581*	*0.135*	*0.126*	*-6.8*
	36	*FC-S*	*34,025*	*0.158*	*0.124*	*-21.3*
	37	*FC-BS*	*58,410*	*0.142*	*0.118*	*-16.5*
	38	*MFC-B*	*53,357*	*0.143*	*0.141*	*-1.2*
	39	*MFC-S*	*62,522*	*0.157*	*0.141*	*-10.3*
	40	*MFC-BS*	*60,335*	*0.143*	*0.133*	*-6.9*
	41	*MFC-SB*	*60,200*	*0.147*	*0.135*	*-8.1*
IX	42	XMMM	235,777	0.185	0.191	3.5
	43	XMMF	243,507	0.192	0.196	1.8
	44	XMFM	238,944	0.207	0.212	2.5
	45	XFMM	226,953	0.208	0.207	-0.7
X	46	SIL-MMM	176,482	0.156	0.157	0.8
	47	SIL-MMF	182,223	0.156	0.161	3.4
	48	SIL-MFM	177,514	0.179	0.173	-3.4
	49	SIL-MFF	182,284	0.158	0.160	1.3
	50	SIL-FMM	170,271	0.160	0.161	0.8
	51	SIL-FMF	174,103	0.164	0.165	0.5
	52	SIL-FFM	162,353	0.157	0.160	2.0
	53	SIL-FFF	166,637	0.148	0.148	0.0
XI	54	M-MMM	77,468	0.119	0.127	6.7
	55	M-MMF	79,318	0.129	0.130	1.1
	56	M-MFM	71,889	0.134	0.140	4.5
	57	M-MFF	73,552	0.129	0.129	-0.3
	58	M-FMM	68,623	0.125	0.133	6.2
	59	M-FMF	70,003	0.132	0.136	3.6
	60	M-FFM	60,287	0.132	0.132	0.5
	61	M-FFF	61,599	0.122	0.122	-0.2
	62	F-MMM	74,298	0.104	0.112	7.1
	63	F-MMF	76,283	0.114	0.114	0.2
	64	F-MFM	69,164	0.116	0.123	6.1
	65	F-MFF	70,732	0.112	0.113	1.2
	66	F-FMM	66,291	0.110	0.119	8.1
	67	F-FMF	67,206	0.121	0.122	0.5
	68	F-FFM	57,581	0.112	0.118	5.5
	69	F-FFF	58,969	0.113	0.109	-3.7
XII	70	MMMS	140,124	0.122	0.122	-0.5
	71	MMFS	148,936	0.132	0.132	-0.5
XII	72	MFMS	145,332	0.138	0.135	-2.2
	73	FMMS	136,212	0.126	0.124	-1.4
XIII	74	M-MMMS	63,358	0.102	0.098	-3.4
	75	M-MMFS	65,737	0.103	0.107	3.4
	76	M-MFMS	60,338	0.107	0.109	1.9
	77	M-MFFS	61,131	0.106	0.106	0.3
	78	M-FMMS	56,768	0.102	0.103	0.8
	79	M-FMFS	58,282	0.103	0.111	7.9
	80	M-FFMS	50,889	0.104	0.103	-1.0
	81	M-FFFS	51,072	0.101	0.101	-0.5
	82	F-MMMS	60,857	0.087	0.086	-0.1
	83	F-MMFS	63,330	0.091	0.093	3.1
	84	F-MFMS	57,988	0.094	0.095	1.7
	85	F-MFFS	58,777	0.089	0.093	4.9
	86	F-FMMS	54,743	0.093	0.092	-1.1
	87	F-FMFS	56,214	0.097	0.099	2.9
	88	F-FFMS	48,669	0.093	0.092	-1.2
	89	F-FFFS	49,048	0.085	0.090	5.3
XIV	*90*	*M-MMM-M*	*36,423*	*0.080*	*0.082*	*1.9*
	91	*M-MMF-M*	*41,612*	*0.086*	*0.086*	*-0.2*
	92	*M-MFM-M*	*39,212*	*0.089*	*0.090*	*1.3*
	93	*M-FMM-M*	*34,187*	*0.082*	*0.085*	*3.6*
	94	*M-MMM-F*	*35,274*	*0.074*	*0.073*	*-2.0*
	95	*M-MMF-F*	*40,595*	*0.080*	*0.076*	*-5.8*
	96	*M-MFM-F*	*38,410*	*0.079*	*0.080*	*1.7*
	97	*M-MFF-F*	*40,669*	*0.081*	*0.075*	*-6.4*
	98	*M-FMM-F*	*33,232*	*0.081*	*0.076*	*-5.9*
	99	*M-FMF-F*	*37,446*	*0.075*	*0.079*	*6.0*
	100	*M-FFM-F*	*33,288*	*0.080*	*0.076*	*-4.7*
	101	*M-FFF-F*	*34,519*	*0.070*	*0.072*	*2.5*
	102	*F-MMM-F*	*34,002*	*0.068*	*0.064*	*-6.2*
	103	*F-MMF-F*	*39,186*	*0.077*	*0.066*	*-13.6*
	104	*F-MFM-F*	*36,815*	*0.078*	*0.070*	*-9.9*
	105	*F-FMM-F*	*32,024*	*0.070*	*0.068*	*-2.8*
XV	106	XMMMM	121,556	0.103	0.103	-0.2
	107	XMMMF	124,796	0.102	0.105	3.5
	108	XMMFM	125,639	0.111	0.113	1.4
	109	XMMFF	83,189	0.099	0.104	5.5
	110	XMFMM	126,898	0.116	0.115	-1.5
	111	XMFMF	84,075	0.121	0.117	-2.6
	112	XMFFM	123,575	0.114	0.114	0.3
	113	XFMMM	117,188	0.117	0.111	-4.9
	114	XFMFM	77,370	0.131	0.122	-7.1
	115	XFFMM	75,890	0.115	0.112	-2.2
XVI	116	SIL-MMMM	93,762	0.087	0.085	-2.4
	117	SIL-MMMF	96,340	0.087	0.087	0.2
	118	SIL-MMFM	96,500	0.097	0.093	-4.0
	119	SIL-MMFF	98,974	0.083	0.086	3.0
	120	SIL-MFMM	97,247	0.104	0.094	-9.6
	121	SIL-MFMF	99,469	0.101	0.096	-4.5
	122	SIL-MFFM	94,340	0.099	0.093	-6.0
	123	SIL-MFFF	96,032	0.084	0.086	1.8
	124	SIL-FMMM	91,152	0.091	0.087	-5.1
	125	SIL-FMMF	93,883	0.092	0.089	-3.9
	126	SIL-FMFM	92,164	0.100	0.095	-5.2
	127	SIL-FMFF	94,585	0.088	0.088	0.0
	128	SIL-FFMM	89,293	0.086	0.087	0.4
	129	SIL-FFMF	90,744	0.094	0.089	-5.8
	130	SIL-FFFM	84,439	0.091	0.086	-5.5
	131	SIL-FFFF	86,665	0.084	0.079	-4.8
XII	*132*	*MMMMS*	*76,122*	*0.071*	*0.066*	*-8.3*
	133	*MMMFS*	*80,605*	*0.075*	*0.071*	*-5.9*
XIII	*134*	*MMMMM*	*53,320*	*0.047*	*0.046*	*-3.1*
	135	*FFFFF*	*48,432*	*0.054*	*0.043*	*-20.0*
XIX	*136*	*MMMMMS*	*43,805*	*0.047*	*0.035*	*-25.4*
	137	*MMMMFS*	*46,785*	*0.042*	*0.038*	*-9.8*
XX	*138*	*MMMMMM*	*31,472*	*0.031*	*0.025*	*-21.8*
	139	*FFFFFF*	*28,591*	*0.043*	*0.023*	*-46.4*
XXI	*140*	*MMMMMMS*	*26,228*	*0.032*	*0.019*	*-39.9*
	141	*MMMMMFS*	*28,051*	*0.027*	*0.021*	*-24.3*

Figure 2: Baseline Fit in Swedish Registers



In percentage terms, the out-of-sample fit is worst for cousins and extremely distant in-laws. We return to those issues below.

These results suggest that it is possible to fit the pattern of inequality across very different kinship types, and for both narrow and distant relatives using a parsimonious model with a limited set of transmission mechanisms. As we will illustrate in Section 4.7, simpler models that do not allow for the transmission of latent factors would not successfully fit the data.

4.3 Intergenerational Transmission

Table 4 summarizes our baseline findings. Panel A reports the calibrated parameters for the intergeneration or "vertical" components of our model. As motivated in Section 2.2, we distinguish between the *direct* transmission of advantages that are reflected in our outcome of interest (i.e. educational attainment), and the transmission of *latent* advantages that are not necessarily reflected in that outcome, but that nevertheless contribute to the socioeconomic success of future descendants.

The direct transmission channels captured by the parameter β^k represent the causal effect of parental education (see Björklund and Salvanes 2010), but also other advantages

Table 4: Calibrated Parameters in Swedish Registers

Panel A: Intergenerational (Vertical)						
Parameters:						
	β^m	β^f	γ^m	γ^f		
	0.140	0.126	0.664	0.566		
	σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}	σ^2_{um}	σ^2_{uf}
	4.648	4.465	2.092	1.592	1.938	2.296
	α_{ym}	α_{yf}	α_{zm}	α_{zf}		
	0.386	0.000	0.658	0.779		
Parent-child correlations in z:						
	Father-Son	Father-Dau	Mother-Son	Mother-Dau		
	0.588	0.600	0.531	0.510		
Ancestor correlations in y and z:						
	Father-Son	Grandf-...	GGrandf-...	GGGrandf-...		
in y	0.381	0.210	0.122	0.072		
in z	0.588	0.346	0.203	0.119		
Panel B: Siblings (Horizontal)						
Parameters:						
	σ^2_{xm}	σ^2_{xf}	σ^2_{xmxf}	σ^2_{em}	σ^2_{ef}	σ^2_{emef}
	0.205	0.256	0.071	0.652	0.708	0.621
Variance Shares:						
in y	4.4%	5.7%	1.6%	14.0%	15.9%	13.6%
in z	-	-	-	31.2%	44.5%	34.0%
Sibling correlations in z:						
	Brothers	Sisters	Brother-Sister			
	0.674	0.812	0.704			
Panel C: Assortative Mating (Horizontal)						
Parameters:						
	r^m_{zz}	r^m_{zy}	r^m_{yz}	r^m_{yy}	$\sigma^2_{\omega m}$	$\sigma^2_{\epsilon m}$
	0.678	-0.012	0.699	0.139	0.666	2.914
	r^f_{zz}	r^f_{zy}	r^f_{yz}	r^f_{yy}		
	0.760	0.106	0.668	0.242		
Spousal correlations in y and z:						
	ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}		
	0.489	0.762	0.544	0.581		
Panel D: Variance Decomposition						
	%	y	z	x	Cov(y,z)	
	Male	0.012	0.450	0.044	0.037	
	Female	0.016	0.357	0.057	0.030	

that are closely correlated with years of schooling. With $\hat{\beta}^m = 0.14$ and $\hat{\beta}^f = 0.13$, this channel contributes very little to the overall transmission of status from one generation to the next. Only about one percent of the variation in years of schooling of the offspring is explained by parental education itself (Panel D of Table 4). This finding holds for all possible combinations of parent and child gender.

Instead, the transmission of advantages occurs predominantly via the latent factor, as measured by the parameter γ^k. We find $\hat{\gamma}^m = 0.66$ and $\hat{\gamma}^f = 0.57$, implying that the set of unobserved advantages that this factor represents are more strongly transmitted from parents to their children than educational attainment itself. Moreover, this latent factor explains a large share of the variance in child education, about 45 percent for men and 36 percent for women.

These results are consistent with the finding from recent multigenerational studies, which document that kinship correlations decline more slowly with the distance between kins than a simple iteration of the parent-child correlation would suggest.[23] To illustrate this implication, we compute the implied autocorrelation of the child and his or her ancestors on the male side (Panel A in Table 4). Because education is only an imperfect proxy for an individual's unobserved characteristics, the observed correlation in parent-child status understates the true persistence of socio-economic advantages across generations.[24]

4.4 Siblings and Horizontal Transmission

In order to incorporate the horizontal family dimension, our model allows for common shocks among siblings (that are allowed to vary with gender), in both observable and latent factors. We find that siblings share influences in both dimensions, as summarized in Panel B of Table 4. The similarity in siblings in observable characteristics (captured by x_t^k) explains about five percent of the overall variation in years of schooling, while the sibling correlation in the latent factor (captured by e_t^k) explains about 15 percent of the variation in years of schooling (and between 30 and 45 percent of the variation in the latent factor itself).

---

[23] See for example Clark (2014), Lindahl et al. (2015), or Braun and Stuhler (2018), and Hällsten (2014) for a survey of the earlier literature. Stuhler (2012) notes that this observation can be rationalized either by the influence of latent factors (as considered by Clark, 2014) or a direct influence of grandparents on their grandchildren (as considered by Mare, 2011).

[24] The correlation between grandfathers and their grandchildren as predicted by our model is very similar (0.210 vs. 0.216) to the corresponding correlation as directly measured in a Swedish data set in Lindahl et al. (2015). This similarity is notable, as this moment was not targeted by our calibration, and we do not make use of *any* higher-order autocorrelations across three or more generations.

Siblings share therefore important influences over and above what can be accounted for by parental characteristics. This finding is consistent with the literature on sibling correlations, which has shown that siblings share many additional influences that are orthogonal to the observed socioeconomic status of parents (see Jäntti and Jenkins 2014, and Section 2.4). Our results provide a richer characterization of this process. In particular, they suggest that much of the advantages that siblings share are *not* reflected in their education.

Intuitively, while the correlation in educational attainment between siblings is high, siblings must be substantially more similar to rationalize why kinship correlations decay so slowly across siblings in-law. For Sweden, the implied correlation in the latent status between siblings is between 0.7 and 0.8 (depending on their gender, see Panel B of Table 4). These implied correlations are around 50 percent higher than the observed sibling correlation in years of education (Table 3), and also much higher than sibling correlations in other socio-economic outcomes.

Siblings must therefore be far more similar in those factors that determine the educational attainment of their descendants (latent advantages) than what is visible in the siblings' own education (observed advantages). This finding implies that sibling correlations still understate the importance of family background, even though they are a more comprehensive measure than intergenerational associations.

We observe sibling correlations to be lower for mixed than for brother or sister pairs (after controlling for the mean education by gender and cohort), a pattern that has been previously observed in the literature. Our results however suggest that male and female siblings share latent influences to a similar degree (cf. $\rho_{e^m e^{m'}}$, $\rho_{e^f e^{f'}}$ and $\rho_{e^m e^f}$ in Table 4). Instead, the effect of family background varies with child gender *conditional* on their latent status – mixed siblings share only about one third of common influences in education (captured by x_t^k) compared to same-gender pairs (cf. $\sigma^2_{x_t^m}$, $\sigma^2_{x_t^f}$ and $\sigma^2_{x_t^m, x_t^f}$). A potential interpretation is that siblings share certain fundamental factors irrespectively of gender, while educational choices conditional on those factors do depend on gender – for example, because the returns to education may be different for women than men.

4.5 Assortative Mating

Spousal correlations in years of schooling are around one half in our data (see Table 3), in line with prior evidence from Sweden, and similar to other countries (Raaum et al., 2007). But while the similarity of spouses in *observable* characteristics are well quantified, our model aims to account also for assortative mating in unobservable factors that matter for child outcomes.

The calibrated parameter values for both the observed and latent dimensions of assortative mating are reported in Panel C of Table 4. In the projection of z^f_{t-1} and y^f_{t-1} on z^m_{t-1} and y^m_{t-1} ,

$$E\left(\begin{array}{c} z^f_{t-1} \\ y^f_{t-1} \end{array} \mid \begin{array}{c} z^m_{t-1} \\ y^m_{t-1} \end{array}\right) = \left(\begin{array}{cc} 0.678 & -0.012 \\ 0.699 & 0.139 \end{array}\right)\left(\begin{array}{c} z^m_{t-1} \\ y^m_{t-1} \end{array}\right), \tag{7}$$

the latent status of the mother is predominantly explained by the latent status of the father. With a coefficient just below zero, the educational attainment of the father has no additional predictive power. However, father's education does have some predictive power for the educational attainment of the mother, over and above what can be explained by the father's latent factor. The corresponding projection matrix for males is similar, with the educational attainment of mothers having slightly more predictive power for the characteristics of their spouses.

These results suggest that the spousal correlation in years of schooling is predominantly a by-product of sorting in latent characteristics. Moreover, it is those latent and more fundamental factors that determine the educational attainment of the next generation, not parental educational attainment per se (see Section 4.3). The key parameter for understanding the transmission of socioeconomic inequalities across generations is therefore the degree of assortative mating in latent factors. By exploiting variation in the magnitude of kinship correlations between close and more distant kins, our approach offers an estimate of that parameter.

The spousal correlations implied by the parametrization of our model are reported in the last block in Panel C of Table 4. The first entry is the calibrated spousal correlation in educational attainment, which at $\rho_{y^m y^f} = 0.49$ is very similar to its sample counterpart. In contrast, the implied spousal correlation in the latent factor z is substantially higher at $\rho_{z^m z^f} = 0.76$. This estimate is also far higher than spousal correlations estimated for countries with strong assortative mating, such as the United States (Raaum et al., 2007).

Our estimates therefore suggest that spouses are substantially more similar in those factors that determine the economic success of their children than they are in educational outcomes. Intuitively, while spousal correlations in educational attainment are high, they are still far too low to rationalize the slow decay of kinship correlations between distant relatives observed in our data. Spouses must therefore be much more similar to each other than what is reflected in observables. This finding is consistent also with recent evidence on the distribution of surnames (Diaz-Vidal and Clark 2015), and the pattern of intergenerational transmission across multiple generations (Braun and Stuhler 2018).

These results add an interesting perspective to recent work on the relation between assortative mating and inequality. A number of studies ask if an increased educational assortative mating may explain the rise in economic inequality that many developed countries have experienced since the early 1980s (for example, Greenwood et al. 2014, Eika et al. 2014). But while contributing to inequalities between *households*, our results suggest that trends in educational assortative mating will not necessarily affect the transmission of inequalities across *generations* – unless they reflect similar trends in the assortative mating in latent factor. But because the latter explains less than half of the variation of the former (see Panel C of Table 4), even large shifts in educational sorting may have little implication for assortative mating in those factors that really affect child outcomes. An interesting question for future research would therefore be if the assortative mating in latent factors has shifted over time, and how such shifts relate to assortative trends in observable characteristics.

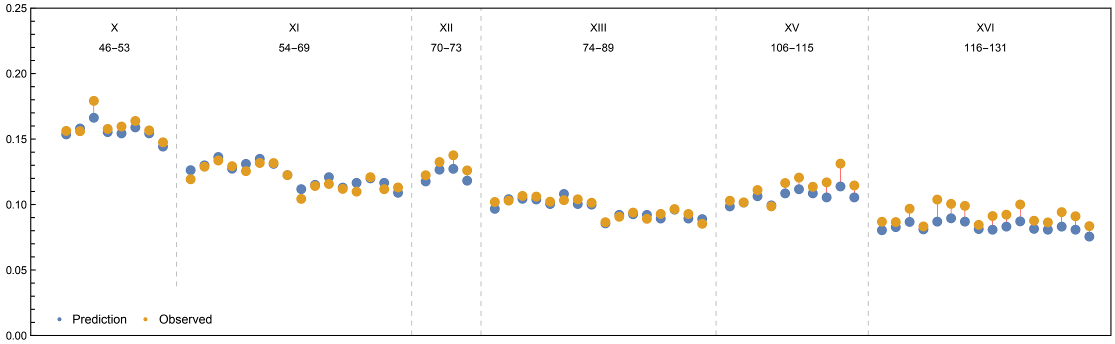
4.6 Fit and Robustness

By incorporating horizontal kins, we observe a greater set of empirical moments than prior studies in the literature. Apart from identifying a more detailed model, this also allows us to study the robustness of our results in more detail. In particular, we study their sensitivity to the choice of empirical moments, and the performance of the calibrated model in terms of predicting kinship correlations that were used (in-sample fit) or not used (out-of-sample fit) in its calibration.

In a first exercise, we exclude 32 moments that involve sibling in-laws of second degree, specifically those involving the child of the sibling in-law (group 12 in Table 4) or his or her spouse (group 13). While reducing the number of moments from 105 to 73, calibration based on this reduced set of moments implies nearly the same kinship correlations and model parameters as our baseline calibration based on the full set (results available upon request). Moreover, the model predicts also the excluded moments well.

In a second exercise, we drop nearly two thirds of our empirical moments, keeping only kinships that we also observe in the Spanish data (described in the next section).

Figure 3: Out-of-Sample Fit in Swedish Registers



Specifically, we drop all moments in moment group 10 and higher, reducing the set of empirical moments by two thirds, from 105 to 35 moments. Moreover, because we drop all distant kins, this exercise tests our model's ability to extrapolate distant kinship correlations from the more narrow ones. The exercise constitutes therefore a particularly challenging robustness test of our results for Sweden. Moreover, it illustrates if our complete model with 20 parameters can be identified from a restricted set of moments that is even more limited than the set of moments in our Spanish sample.

We find that our results remain quite stable even when using this restricted set of moments. Figure 3 plots the out-of-sample fit of the calibrated model for the 70 excluded moments. The model calibrated from the restricted set of moments successfully predicts a diverse set of kinship correlations not included in the calibration, including vertical, horizontal and distant moments. Our baseline results are therefore robust to the choice of moments. Moreover, the model appears identifiable from information that is far more limited than what we observe in the Swedish data.

However, our model does not provide a good fit for two kinship types, *cousins* and *very distant* in-laws. On average, our baseline calibration understates the correlation in years of schooling between cousins by about 10 percent, and the kinship correlations for in-laws of 4^{th} degrees of separation (i.e. separated by four siblings and four spouses) or higher by an even larger amount (although the deviations are small in absolute value). This gap closes only partially if we directly target those moments in the calibration. After examination, the apparent understatement of the kinship correlation between cousins appears to reflect a measurement problem, while the understatement of the most distant kinship correlations might reflect a limitation of our model.

A fundamental problem for the estimation of any distributional model is that the moments that we wish to explain may vary over time. Our calibration relies on the identifying assumption that these moments are in a steady-state equilibrium, but this assumption is particularly unlikely to hold in intergenerational data spanning multiple generations (see Atkinson and Jenkins 1984, Nybom and Stuhler 2014). Because the outcome for horizontal kins can be measured at approximately the same time, our approach is arguably less sensitive to this assumption than recent multigenerational studies, which exploit vertical information across three or even more generations. Still, changes in kinship correlations over time could affect our results.

Such time trends are hard to address, because various sources of selection bias – with respect to measurement of the outcome variable, the age of parents, or identification of kins via their common ancestor (see Section 3.1) – limit our ability to freely choose a cohort range for each kinship type. In particular, the identification of cousins requires the identification of grandparents, such that the sample correlations for cousins are based on comparatively younger cohorts. This would be unproblematic if the kinship correlations were stable over cohorts anyways. To study this question, we estimated the time trends in our kinship correlations based on a two-step procedure. While other kinship correlations have indeed been very stable, the sample correlation between cousins increases substantially over our period of study.[25] After adjusting for that time trend, the sample correlations for cousins are better in line with our calibration. We nevertheless exclude those moments from our baseline analysis, but including them has only negligible effects on our results.

In contrast, our model's tendency to understate the most distant kinship correlations (cf. in-laws of 4^{th} or 5^{th} degrees of separation in Table 4) cannot be due to their variability over time. Time trends for those moments – which can be estimated for a wide range of cohorts, because the identification of in-laws does not require the identification of grandparents – appear quite stable. Instead, we may understate the similarity of extremely distant family members because we allow for only a *single* latent factor in our model.

Traditionally, the latent factor is thought to summarize the influence of many transmission channels, such as cultural, genetic or social determinants (see e.g. Becker and Tomes 1986). But some of those determinants may have higher rates of persistence than others, and it is those with the strongest persistence that will increasingly explain the

---

[25] For each of the 208 kinships, we estimated in a first step the kinship correlation separately by birth year, and then estimated in the second step a liner time trend over the first-stage coefficients. The evidence is available upon request from the authors. See also Björklund et al. (2009), who show that brother correlations in years of schooling have been approximately constant in Sweden during our study period.

kinship correlation between increasingly distant kins.[26] For example, the determinants of socio-economic status that relate to race (such as discrimination) or geography (such as the local economic structure) may be quite persistent if interracial marriage is rare or geographic mobility low. This gap between the "average" persistence of the latent factor and the more persistent sub-factors will become increasingly apparent for more distant in-laws. A model with a single latent factor may therefore still understate the similarity of *very* distant kins. Consideration of a more general model that includes multiple latent factors with different rates of persistence could therefore be an interesting avenue for future research.

Finally, we consider how robust our results are to variation in the values of the observed empirical correlations. Because they are estimated from very large samples, sampling variation is not our primary concern.[27] Instead, we are concerned about a conceptual mismatch between the population that we sample from and its theoretical counterpart. One source for such mismatch is variation in kinship correlations over time (as detailed above), but deviations may also arise because of other factors not explicitly modeled (such as fertility pattern, non-linearities, and so on). To check if such deviations could have a significant effect on our findings, we test how changes in the empirical moments map into variation in the model parameters.

Specifically, we "perturb" each of the 20 parameters of our baseline calibration by multiplying it with a random variable that is uniformly distributed between 0.95 and 1.05. Implementing this perturbation one million times, we then compare the distribution of the kinship correlations based on these perturbations with the kinship correlation as predicted by the baseline calibration, and the kinship correlations as observed in the data. The model predictions are centered and smoothly distributed around the kinship correlations measured in the actual data [Figure to be included]. This observation suggests that small variation in the underlying empirical moments would only have limited effects on our findings.

---

[26] Stuhler (2012) argues that this "multiplicity" of transmission mechanisms is another explanation why multigenerational correlations do not decay as quickly as a model with a single factor would suggest. Models with persistent group membership can be seen as the extreme case of this argument. For example, Becker and Tomes (1986) include such dummy to explain why racial gaps in the United States decay much more slowly than their model would otherwise predict.

[27] Sampling variation is negligible even for the most distant moments in our baseline calibration.

4.7 Restricted Models

Because our model is comparatively general, one may ask if some of its components – i.e. the direct and indirect transmission processes in the intergenerational, assortative, and sibling dimension – could be removed without greatly reducing the model's ability to explain the horizontal (intergenerational) and vertical pattern of socioeconomic inequality. Figure 4 provides evidence on this question.

As a reference point, we first shut down the direct transmission mechanism, calibrating the model with the restriction $\beta^m = \beta^f = 0$. This restricted model explains the data nearly as well as our benchmark model, as illustrated in Figure 5a. The observation that the direct transmission mechanism is not important is consistent with the observation that it explained less than two percent in the variation of educational attainment in our benchmark model.

In contrast, it is crucial to allow for the transmission of latent characteristics. The fit of the model with the restriction $\gamma^m = \gamma^f = 0$ is very poor, as shown in Figure 5b. While some type of kinship correlations are understated (such as the spousal correlation, i.e. moment group I), others are greatly overstated (such as the parent-child and sibling in-law correlations, groups III and IV). These biases still occur among the very distant horizontal kinships – the intergenerational components of the model do affect also the horizontal dimension, and their apparent misspecification does not cancel out with distance.

We next study how important it is to allow for siblings to share more influences than what is explained by average rate of intergenerational transmission captured in β and γ. The fit of a restricted model in which all sibling components have variance zero is not as good as the benchmark model, as shown in Figure 5d. However, the consequences are not as severe as one might expect. The sibling correlations (group II) are of course heavily understated, but the predictions remain good for more distant kins.This suggests that the long-run implications of the sibling components can be "imitated" by other parts of the model. Still, it is obvious from the exercise that accounting for shared sibling components does matter, and we retain them in our benchmark model.

Finally, to study the role of assortative process we calibrate the model under the assumption that assortative mating occurs exclusively in the observed outcome, as has been a standard modeling choice in the literature. With this more restrictive assortative process the model does not explain the data well, as shown in Figure 5d. The spousal correlation is overstated by 45 percent (0.71, outside of the plot area), and the sibling (group II) and sibling in-law (group IV) correlations are substantially understated. In contrast, the distant in-law correlations are overstated, with the prediction error increasing in relative terms with distance. Our results therefore suggest that the assortative mating in latent factors is key to match the pattern of socioeconomic inequality across distant kins.[28]

In sum, these results suggest that all components of our model are important, with the exception of the direct transmission mechanism captured by β. We nevertheless retain this channel in our model because this channel has received attention in the economic literature; it is therefore desirable to not exclude it apriori. Moreover, the channel might be more important for other outcomes, as we will test below. The latent factor is the key part of the model, both in the intergenerational and assortative dimension. Based on our results, it seems doubtful that models that do not account for the role of unobservable characteristics could successfully fit the pattern of socioeconomic inequality across kins. We base this finding on our long list of moments, including distant moments. We show below that it is much harder to discriminate between transmission models with the narrow set of moments that have typically been used in the literature.

5 Kinship Correlations in Spain

We next calibrate our model for Spain, for which the intergenerational evidence has so far been quite limited (see for example Black and Devereux 2011). Because the available data sources do not report income for both parents and children, income mobility has to be approximated based on two-sample IV estimators (Cervini-Plá 2015). And while educational outcomes provide a good proxy for income mobility, the available cross-country comparisons such as Hertz et al. (2008) or Blanden (2011) do not contain evidence for Spain either.[29]

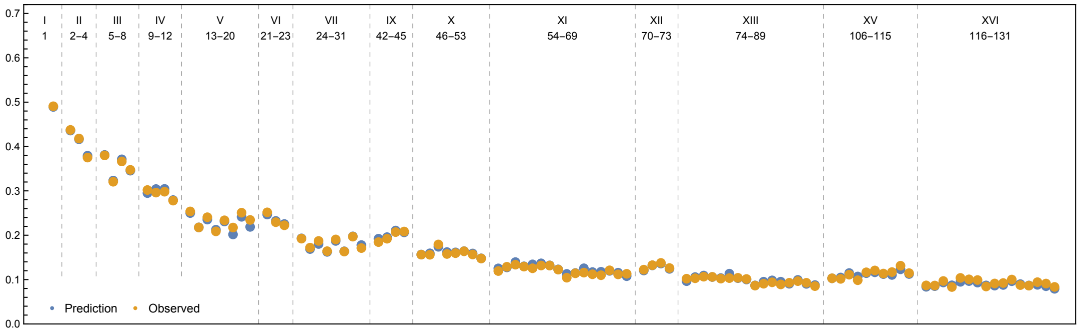
We circumvent these limitations by using *surnames* to identify kins in the 2001 Census from the Spanish region of Cantabria. This source is interesting also from a method-

---

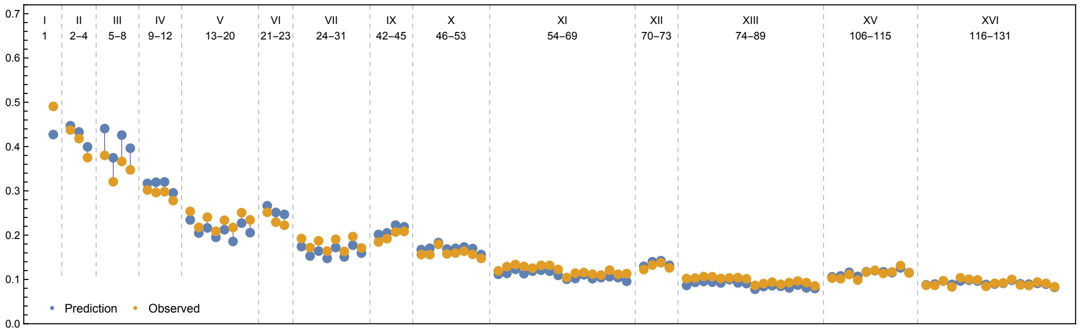
[28] We also calibrated a model in which we allowed for one-dimensional assortative mating in the latent factor, but not in observables. The results are similar to our benchmark (results available upon request). This result suggests that if researchers wish to consider a simple one-dimensional model of assortative mating the latent factor should be its key component, i.e. by interpreting the observed spousal correlation in education within an errors-in-variables model. An example for this approach is Ermisch et al. (2006).

[29] Blanden (2011) demonstrates that the ranking of countries in terms of educational mobility and income mobility is quite similar, with a pair-wise correlation between the two type of measures of around 0.7. While our evidence pertains to education, it is therefore likely to be informative about the transmission of economic advantages more generally.

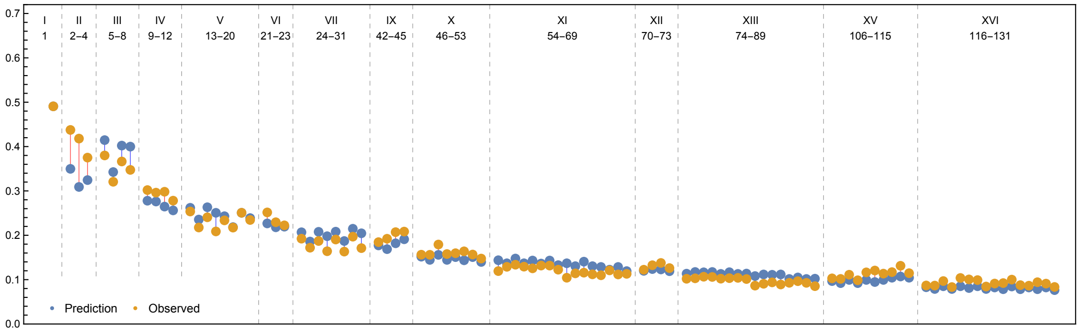
Figure 4: Restricted Models



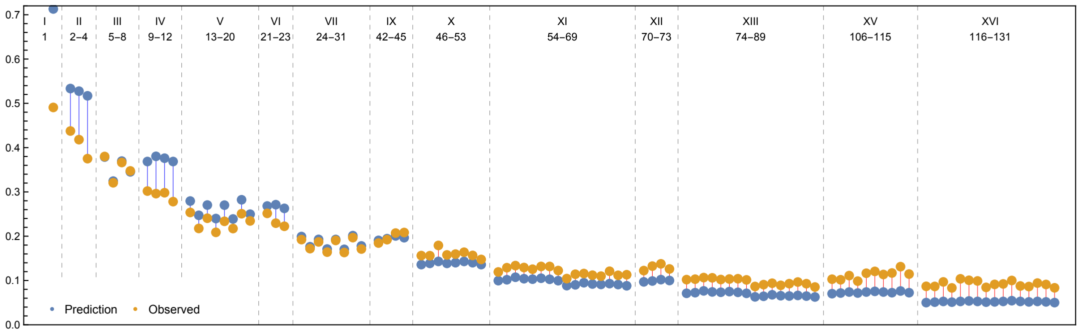
(a) No direct transmission ($\beta = 0$)



(b) No latent transmission ($\gamma = 0$)



(c) No shared sibling component



(d) Assortative mating only in observables ological perspective, to illustrate the feasibility of our approach in settings with scarce data. In contrast to the administrative registers from Sweden, the Census data are limited to a single cross-section, contain only educational outcomes and names, and lack direct family links to define kinship.

To recover family links, we exploit that children in Spain inherit surnames from both their parents. This naming convention allows us to track both maternal and paternal lines, as described further in Section 3.2. It also leads to match rates that at around 36% are far higher than the match rates that have been achieved in Census data from other countries. The Census is a complete count, allowing us to quantify educational inequalities more precisely than the prior literature.

We can therefore compare kinship correlations between Spain and Sweden, across a wide range of kinship types. More importantly, we will distinguish if differences in the direct, indirect, or assortative processes contribute to differences in intergenerational mobility rates between the two countries.

5.1 Sample and Calibrated Moments

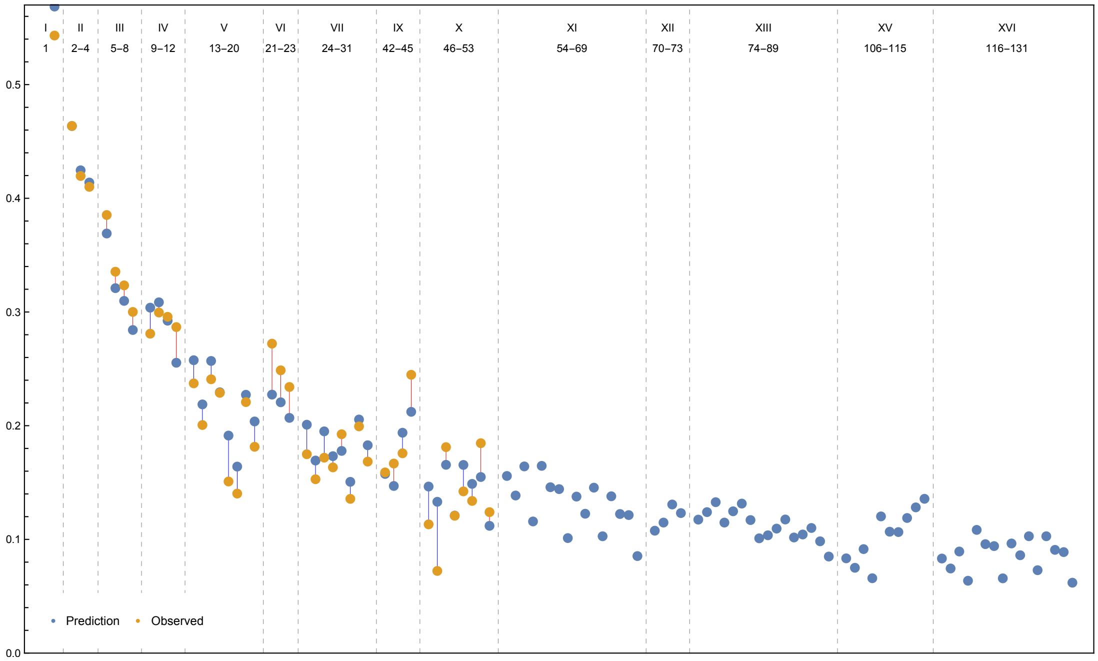
Table 5 reports the sample correlation in years of schooling for each kinship, sorted from closely related to more distant kins. Columns (1) and (2) report the number of pairs and sample correlations. The pairs are weighted inversely by the square root of family size, as described in Section 3. We observe 65 distinct moments that can be classified into groups from very close kins (such as spouses, group I) to relatively distant kins (such as second-degree siblings in-law, group X). Two kinship types that were not observed in the Swedish data are child-parents in-law (group III-b in Table 5) and the Spouse of nephew/niece-uncle/aunt (group VII-b), as its child generation was too young for this definition to be meaningful. The sample sizes are much smaller than in the Swedish sources. They are however still large enough for precise measurements of kinship types involving siblings and parents, and become noisy only for the more distant types.

The kinship correlations in educational attainment tend to be slightly larger in our Spanish data than the corresponding moments for Sweden. For example, the brother (male cousin) correlation in years of schooling in Spain is 0.46 (0.20), compared to 0.43 (0.17) for Sweden. The gap is smaller for in-law and vertical kinships, and inverses for those moments that involve females in the parent generation (such as the mother-son or aunt-nephew relationships).

Table 5: Estimated and Calibrated Moments in Spanish Census (Years of Schooling)

Kinship type			Data		Calibration	
#	##	name	number of pairs (1)	sample correlation (2)	predicted correlation (3)	percent error (4)
I	1	HUSB-WIFE	24,819	0.543	0.569	4.7
II	2	BROTHERS	11,109	0.464	0.464	0.0
	3	SISTERS	10,316	0.420	0.425	1.2
	4	BROTH-SIS	21,017	0.410	0.414	0.9
III	5	FATH-SON	25,860	0.385	0.369	-4.2
	6	FATH-DAUGH	24,610	0.335	0.321	-4.3
	7	MOTH-SON	25,860	0.323	0.310	-4.2
	8	MOTH-DAUGH	24,610	0.300	0.284	-5.3
III-b	5-b	FATHIL-SON	13,191	0.262	0.276	5.1
	6-b	FATHIL-DAUG	11,628	0.280	0.265	-5.4
	7-b	MOTHIL-SON	13,191	0.245	0.225	-8.0
	8-b	MOTHIL-DAUG	11,628	0.236	0.219	-7.1
IV	9	BL-HS	12,260	0.281	0.304	8.2
	10	BSL-WB	11,184	0.300	0.309	3.0
	11	BSL-HS	12,339	0.296	0.292	-1.2
	12	SL-WB	10,743	0.287	0.255	-10.9
V	13	NEU-BF	3,787	0.237	0.258	8.6
	14	NIU-BF	3,487	0.201	0.219	9.0
	15	NEU-BM	3,602	0.241	0.257	6.7
	16	NIU-BM	3,337	0.229	0.229	0.2
	17	NEA-SF	3,452	0.151	0.191	26.8
	18	NIA-SF	3,253	0.140	0.164	16.9
	19	NEA-SM	3,334	0.221	0.227	2.9
	20	NIA-SM	3,067	0.181	0.204	12.3
VI	21	BL-WWS	4,156	0.272	0.227	-16.5
	22	SL-HHB	3,296	0.249	0.221	-11.3
	23	BSL-HWBS	7,061	0.234	0.207	-11.6
VII	24	NEWUBF	3,787	0.175	0.201	14.9
	25	NIWUBF	3,487	0.153	0.169	10.8
	26	NEWUBM	3,602	0.172	0.195	13.5
	27	NIWUBM	3,337	0.163	0.173	6.1
	28	NEHASF	3,452	0.192	0.178	-7.6
	29	NIHASF	3,253	0.136	0.151	11.0
	...					
	30	NEHASM	3,334	0.199	0.205	3.0
	31	NIHASM	3,067	0.168	0.183	8.5
VII-b	24-b	Wife of NEWU	1,738	0.220	0.197	-10.1
	25-b	Husband of NI	1,930	0.192	0.191	-0.2
	26-b	Wife of NEWU	1,737	0.213	0.198	-6.9
	27-b	Husband of NI	1,873	0.224	0.210	-6.5
	28-b	Wife of NEHAS	1,537	0.161	0.146	-9.1
	29-b	Husband of NIH	1,746	0.112	0.140	24.3
	30-b	Wife of NEHAS	1,559	0.196	0.175	-10.9
	31-b	Husband of NIH	1,648	0.176	0.185	4.8
VIII	32	MC-B	2,053	0.202	0.200	-1.0
	33	MC-S	1,779	0.200	0.214	7.0
	34	MC-BS	3,752	0.209	0.189	-9.4
	35	FC-B	1,747	0.145	0.140	-3.5
	36	FC-S	1,523	0.182	0.167	-8.1
	37	FC-BS	3,368	0.188	0.140	-25.4
	38	MFC-B	3,817	0.187	0.167	-10.5
	39	MFC-S	3,364	0.192	0.189	-1.7
	40	MFC-BS	3,604	0.191	0.167	-12.6
	41	MFC-SB	3,625	0.172	0.159	-7.9
IX	42	XMMM	3,045	0.159	0.158	-0.8
	43	XMMF	2,924	0.167	0.147	-11.8
	44	XMFM	3,089	0.176	0.194	10.3
	45	XFMM	3,132	0.245	0.212	-13.3
X	46	SIL-MMM	1,966	0.113	0.146	29.4
	47	SIL-MMF	1,950	0.072	0.133	84.2
	48	SIL-MFM	2,009	0.181	0.166	-8.6
	49	SIL-MFF	1,881	0.121	0.121	0.1
	50	SIL-FMM	1,854	0.142	0.165	16.4
	51	SIL-FMF	1,807	0.134	0.149	11.2
	52	SIL-FFM	1,792	0.185	0.155	-16.1
	53	SIL-FFF	1,710	0.124	0.112	-9.7

Figure 5: Fit in Spanish Census



In our calibration we include all groups (including III-b and VII-b). We therefore use 65 distinct moments from 12 different kinship types. We calibrate the model as described in Section 3, and report the predicted moments as well as the percentage deviation between the observed and predicted moments in columns (3) and (4) of Table 5. Moments that were not included in the calibration are printed in italics.

Overall, the calibrated model explains the data well. Figure 5 illustrates the fit graphically, by plotting the sample moments and predicted moments from the calibrated model. For comparability, we plot the same 105 moments that are also plotted for our Swedish data in Figure 2. The mean absolute prediction error across all included moments is 9.9 percent. The prediction errors are therefore larger than in the Swedish data, consistent with the presence of sampling error from the smaller sample sizes.

The model however appears to fit the pattern of inequality transmission in Spain across vertical and horizontal kins, and across both direct and in-law relationships. Given the more restricted set of moments and number of observations, this conclusion is not as well supported as for the case of Sweden. Nevertheless, two observations suggest that our results are robust. First, we demonstrated in the Swedish sample that the full model (including the different channels via which inequalities are transmitted) could be calibrated from a set of moments that was much more more restrictive than the set available for Spain (see Section 4.6). Second, our findings for Spain appear not too sensitive to changes in the underlying set of moments or their weights.[30]

5.2 Intergenerational Transmission

Panel A of Table 6 reports the calibrated parameters for the intergenerational or "vertical" transmission in Spain. We again find that the direct transmission channels captured by the parameter β^k contribute very little to the transmission of educational inequalities. At $\hat{\beta}^m = 0.03$ and $\hat{\beta}^f = 0.11$ the estimates are close to the corresponding estimates for Sweden. Only about one percent of the variation in years of schooling is directly explained by parental education (panel D of Table 6).

As for Sweden, the transmission of advantages occurs predominantly via the latent factor. At $\hat{\gamma}^m = 0.92$ and $\hat{\gamma}^f = 0.84$, the rate by which this latent factor is transmitted from parents to children is substantially higher in Spain than in Sweden. As a consequence, the implied correlation in the latent status between parents and children is also much higher (about 30 percent higher). The share of variance explained by the latent factor is similar as in Sweden for male, but much lower for women.[31] Our results suggest that educational correlations decay more slowly across generations in Spain than in Sweden. For example, the implied correlation in educational attainment between children and their great-great-grandfathers is more than twice as large in Spain (0.16 vs. 0.07, Panel A of Tables 4 and 6).

In sum, our results have three major implications. First, the pattern of inequality transmission in Spain and Sweden are qualitatively similar, with the transmission of advantages occurring predominantly via latent variables, and only a minor role for parental education itself. Second, the stronger transmission of educational inequalities in Spain is explained by stronger transmission in the latent factor. That is, the difference in the intergenerational process between Spain and Sweden is not superficial, but due to fundamental differences in the extent of status transmission. Third, standard measures understate the difference in intergenerational transmission between Sweden and Spain.

---

[30] Because educational attainment for females has been increasing rapidly in our observation period, its distribution is quite different in the parent and child generation. The main results remain however robust to excluding moments that involve the mother, mother in-law or aunt. Our results are also robust to dropping distant kinships types for which sample sizes become very small, or to weight moments by the square root of the sample size (results available upon request).

[31] One potential explanation for this pattern could be that secular trends in females' educational attainment were stronger in Spain, such that less of the total variation in educational attainment is explained by individual- and family-specific factors.

Table 6: Calibrated Parameters in Spanish Census

Panel A: Intergenerational (Vertical)						
Parameters:						
	β^m	β^f	γ^m	γ^f		
	0.027	0.111	0.915	0.842		
	σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}	σ^2_{um}	σ^2_{uf}
	13.579	13.213	6.519	2.779	5.162	7.003
	α_{ym}	α_{yf}	α_{zm}	α_{zf}		
	0.742	0.855	0.587	0.127		
Parent-child correlations in z:						
	Father-Son	Father-Dau	Mother-Son	Mother-Dau		
	0.760	0.827	0.732	0.883		
Ancestor correlations in y and z:						
	Father-Son	Grandf-...	GGrandf-...	GGGrandf-...		
in y	0.369	0.271	0.205	0.156		
in z						
Panel B: Siblings (Horizontal)						
Parameters:						
	σ^2_{xm}	σ^2_{xf}	σ^2_{xmxf}	σ^2_{em}	σ^2_{ef}	σ^2_{emef}
	1.650	2.644	2.089	0.558	0.001	0.018
Variance Shares:						
in y	12.1%	20.0%	15.6%	4.1%	0.0%	0.1%
in z	-	-	-	8.6%	0.0%	0.4%
Sibling correlations in z:						
	Brothers	Sisters	Brother-Sister			
	0.674	0.784	0.667			
Panel C: Assortative Mating (Horizontal)						
Parameters:						
	r^m_{zz}	r^m_{zy}	r^m_{yz}	r^m_{yy}	$\sigma^2_{\omega m}$	$\sigma^2_{\epsilon m}$
	0.731	-0.139	0.418	0.357	0.381	8.369
	r^f_{zz}	r^f_{zy}	r^f_{yz}	r^f_{yy}		
	1.291	0.083	0.576	0.441		
Spousal correlations in y and z:						
	ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}		
	0.569	0.903	0.483	0.549		
Panel D: Variance Decomposition						
	%	y	z	x	Cov(y,z)	
	Male	0.001	0.480	0.121	0.009	
	Female	0.010	0.210	0.200	0.009	

While parent-child and sibling correlations are only slightly larger in Spain, the gap is larger for more distant relatives.

5.3 Siblings and Horizontal Transmission

Panel B of Table 4 summarizes our findings that pertain to siblings, which quantify what siblings share over and above the average rate of intergenerational transmission discussed in the previous section. The similarity in siblings in observable and latent characteristics captured by x^k and e^k explains between 15 percent (brothers and mixed pairs) and 20 percent (sisters) of the variation in educational attainment.

Siblings share common influences to a similar degree in Spain as in Sweden. However, in Spain most of this similarity is explained by the sibling component in observables x^k, which is about three times larger share than in Sweden. One potential explanation for this finding are location-specific shocks and trends. Because siblings grow up in the same location, structural changes in the local provision of schooling would tend to be reflected in this component. The shared sibling components in the latent factor e_t^k seem not important in the Spanish sources. However, because we lack long-distance horizontal relationships in the Spanish data, it is difficult to distinguish the two types of shared sibling components.

As in the Swedish data, siblings must be far more similar to each other than what is captured by sibling correlations in years of education. For Spain, the implied correlation in the latent status between siblings are 0.7 or even higher, about 50 percent larger than the sibling correlation in years of education.

5.4 Assortative Mating

The calibrated parameter values for both the observed and latent dimensions of assortative mating in Spain are summarized in Panel C of Table 4. Spousal correlations in years of schooling are around 0.54 in our Spanish sources, about 10 percent higher than the corresponding moment in the Swedish registers (see table 3). The expectation of z_{t-1}^f and y_{t-1}^f conditional on z_{t-1}^m and y_{t-1}^m is estimated as

$$E\begin{pmatrix} z_{t-1}^f & | & z_{t-1}^m \\ y_{t-1}^f & & y_{t-1}^m \end{pmatrix} = \begin{pmatrix} 0.731 & -0.139 \\ 0.418 & 0.357 \end{pmatrix} \begin{pmatrix} z_{t-1}^m \\ y_{t-1}^m \end{pmatrix}, \tag{8}$$

As was the case for Sweden, the latent status of the mother is predominantly explained by the latent status of the father, while his educational attainment has no additional predictive power. However, the father's education has a substantial association with maternal education, over and above what can be explained by the father's latent factor. The corresponding projection matrix for females is similar.

The spousal correlations implied by these parameters are reported in the last block of Panel C. The first entry is simply the calibrated spousal correlation in educational attainment, which at $\rho_{y^m y^f} = 0.57$ is similar to its sample counterpart. In contrast, the implied spousal correlation in the latent factor is substantially higher, which at $\rho_{z^m z^f} = 0.90$ is also more than 10 percent higher than the corresponding estimate for Sweden. Our results therefore suggest stronger assortative mating in Spain compared to Sweden – not only in educational attainment, but also in the latent determinants of socioeconomic status. Spouses in Spain appear extremely similar in those factors that ultimately determine the educational attainment of their descendants.

6 Other Outcomes

Our results point to the strong transmission of advantages, and suggest that intergenerational correlations and other standard measures understate the extent of status transmission. These results are based on educational attainment, which is seen as the key mediator for the transmission of socio-economic advantages in both sociological research (as conceptualized in the so-called *origin-education-destination triangle*) and standard economic approaches (Goldthorpe 2014). Still, the transmission of educational advantages may follow a distinct pattern that might not generalize to other socioeconomic advantages.

To study this question, we calibrated our model for other outcomes observed in the Swedish registers. We report two such exercises here. In Section 6.1, we consider the *income* of parents and children, which is arguably a more direct measure of the socioeconomic origin and destination than education.[32] Because the Swedish registers track income profiles over nearly six decades, we can construct high-quality measures

---

[32] Income has been the primary measure of socio-economic status in the economic literature, while education or occupation play a more important role in sociological research. This contrast is becoming less sharp, however. Recent sociological work emphasizes the informative content of income (Kim et al. 2018). In economics, occupational and educational measures have become the primary measures for comparative or historical studies (e.g. Hertz, 2007, or Long and Ferrie, 2013a), and the recent multigenerational literature (Stuhler, 2018).

of income for both the parent and child generation, and study how measurement error affects our findings.

In Section 6.2 we calibrate our model using information on *height* from military enlistment tests. While not a socio-economic outcome of interest per se, height is a useful reference point. Its transmission process is better understood than the transmission of socioeconomic outcomes, with genes known to be the primary source of variation in body height. This application illustrates therefore how genetic mechanisms are captured by our approach, which in turn provides a useful reference point for studying their role in socioeconomic outcomes.

Finally, the analysis of income and height is interesting from a methodological perspective. As our information on body height stems from military enlistment tests, it is available only for males. Income is available for both genders, but is a worse measure of the socioeconomic status for women given that their labor market participation is lower. We therefore ask if we can identify the complete model – including the assortative process between spouses, and the transmission processes for mothers and daughters – without observing any females in our data. Such indirect identification may be feasible because distant kinship correlations for male pairs are a function of transmission processes for females.[33]

6.1 Income

We measure income as total pre-tax income, as detailed in Section 3.1. Our primary measure is the logarithm of *ten-year* averages of annual income centered around age 35 for children and around age 45 for parents. For robustness we also consider shorter averages *(five-year* and *annual*) and Spearman rank instead of log-linear Pearson correlations. As for education, we observe 141 distinct moments, classified into 21 kinship types. Columns (1) and (2) of Table 7 report the sample size and sample correlations for a subset of those 141 moments.[34]

The correlations are systematically lower for mixed or female than for male pairs. This pattern is consistent with the observation that women were less likely to participate

---

[33] For example, we observe two different types of brothers in-law, one in which the in-laws are separated only by one female (considering the husband of the sister) and another in which they are separated by two females (considering brother in-laws whose wives are sisters).

[34] Our estimates for intergenerational and sibling correlations are broadly in line with prior evidence for Sweden. Estimates that are based on longer income spans or that adjust for measurement error are somewhat larger, see for example Nybom and Stuhler (2017) or Björklund et al. (2009).

Table 7: Estimated and Calibrated Moments in Swedish Registers (Income)

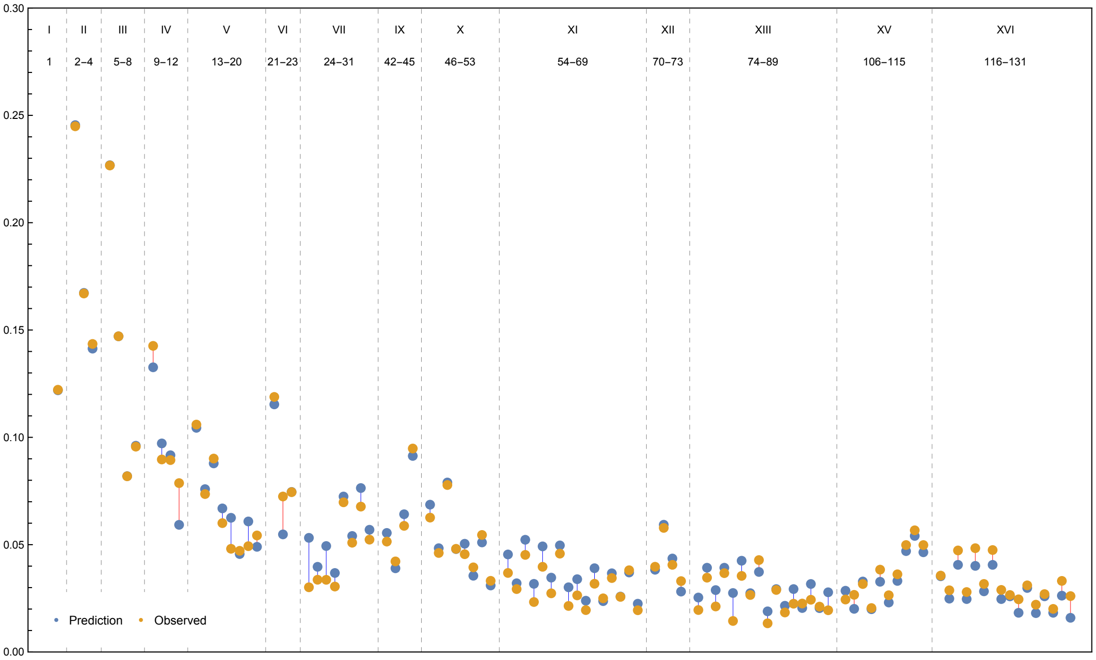
Kinship type			Data		Calibration		Kinship type			Data		Calibration	
#	##	name	number of pairs (1)	sample correlation (2)	predicted correlation (3)	percent error (4)	#	##	name	number of pairs (1)	sample correlation (2)	predicted correlation (3)	percent error (4)
I	1	HUSB-WIF	412,735	0.122	0.122	-0.2		…					
II	2	BROTHERS	99,799	0.245	0.245	0.2	XII	72	MFMS	154,974	0.041	0.044	7.3
	3	SISTERS	88,937	0.167	0.167	0.2		73	FMMS	119,219	0.033	0.028	-14.7
	4	BROTH-SIS	185,448	0.144	0.141	-1.5	XIII	74	M-MMMS	83,216	0.020	0.025	29.8
III	5	FATH-SON	388,190	0.227	0.227	0.1		75	M-MMFS	108,087	0.035	0.039	13.4
	6	FATH-DAUGH	366,590	0.147	0.147	0.0		76	M-MFMS	83,133	0.021	0.029	35.9
	7	MOTH-SON	341,891	0.082	0.082	0.1		77	M-MFFS	94,984	0.037	0.039	6.9
	8	MOTH-DAUGH	422,142	0.096	0.096	0.4		78	M-FMMS	62,612	0.014	0.027	90.1
IV	9	BL-HS	339,682	0.143	0.133	-7.0		79	M-FMFS	78,556	0.035	0.043	19.9
	10	BSL-WB	296,051	0.090	0.097	8.3		80	M-FFMS	56,750	0.027	0.027	3.1
	11	BSL-HS	362,169	0.089	0.092	2.5		81	M-FFFS	62,983	0.043	0.037	-12.9
	12	SL-WB	301,464	0.079	0.059	-24.7		82	F-MMMS	78,128	0.013	0.019	41.8
V	13	NEU-BF	149,817	0.106	0.104	-1.4		83	F-MMFS	102,222	0.029	0.029	1.4
	14	NIU-BF	142,096	0.074	0.076	3.1		84	F-MFMS	79,019	0.018	0.022	17.0
	15	NEU-BM	201,393	0.090	0.088	-2.5		85	F-MFFS	90,190	0.022	0.029	30.3
	16	NIU-BM	190,299	0.060	0.067	11.4		86	F-FMMS	59,524	0.023	0.020	-9.3
	17	NEA-SF	152,065	0.048	0.063	30.0		87	F-FMFS	75,079	0.024	0.032	30.2
	18	NIA-SF	142,912	0.047	0.046	-3.2		88	F-FFMS	53,001	0.021	0.020	-3.3
	19	NEA-SM	217,131	0.049	0.061	23.4		89	F-FFFS	59,009	0.019	0.028	43.0
	20	NIA-SM	205,748	0.054	0.049	-9.7	XIV	90	M-MMM-M	41,300	0.017	0.025	43.0
VI	21	BL-WWS	156,164	0.119	0.115	-2.9		91	M-MMF-M	61,036	0.025	0.026	4.1
	22	SL-HHB	112,873	0.072	0.055	-24.4		92	M-MFM-M	45,581	0.027	0.028	4.8
	23	BSL-HWBS	251,377	0.074	0.075	0.1		93	M-FMM-M	31,435	0.033	0.027	-20.1
VII	24	NEWUBF	120,226	0.030	0.053	76.4		94	M-MMM-F	38,113	0.022	0.018	-17.3
	25	NIWUBF	114,422	0.034	0.040	17.9		95	M-MMF-F	58,304	0.022	0.019	-10.7
	26	NEWUBM	158,116	0.034	0.049	46.8		96	M-MFM-F	43,614	0.030	0.021	-30.5
	27	NIWUBM	149,645	0.030	0.037	20.7		97	M-MFF-F	57,298	0.028	0.019	-30.1
	28	NEHASF	124,725	0.070	0.072	3.9		98	M-FMM-F	29,951	0.017	0.020	19.3
	29	NIHASF	117,191	0.051	0.054	6.2		99	M-FMF-F	43,353	0.007	0.021	201.5
	30	NEHASM	179,624	0.068	0.076	12.8		100	M-FFM-F	30,120	0.026	0.020	-24.8
	31	NIHASM	170,499	0.052	0.057	8.8		101	M-FFF-F	38,840	0.022	0.018	-15.6
	…							…					

in the labor force, in particular for the parent generation in our sample. We do not account explicitly for labor supply decisions. However, our model allows for gender differences in all its components, and for the transmission of other advantages apart from income in intergenerational and assortative processes. It is an empirical question if that model provides sufficient flexibility to fit kinship correlations for a comparatively complex outcome such as income.

In our calibration we include siblings in-law up to three degrees of separation (group XVI), but exclude more distant kinship correlations. With these restrictions our calibration is based on 129 distinct moments, grouped into sixteen different kinship types. We report a subset of the predicted moments in Table 7. Moreover, Figure 6 plots the subset of 105 moments that we also included in our baseline calibration for educational attainment. As for education, the calibrated model explains the data well, providing a close fit to both vertical and horizontal moments, and for both direct and affinity relationships. The model is also replicates the asymmetric transmission pattern across genders.

Table 8 reports the calibrated parameters for our model of income transmission, separately by intergenerational (Panel A), sibling (Panel B), and assortative processes

Figure 6: Sample and Predicted Moments (Income)



(Panel C). The results vary somewhat with the income definition (*ten-year, five-year* and *annual* incomes, *rank* or *log* incomes), but the broad pattern remains stable.

Our findings for income are qualitatively similar to our benchmark calibration based on years of schooling, but differ in magnitude. The latent advantages are more strongly transmitted than income itself, across all three dimensions of our model: the intergenerational, sibling and assortative processes. The father-son correlation in the latent factor is 0.46, twice as large as the corresponding correlation in log income, but substantially below the corresponding estimates from our benchmark calibration. Our results therefore suggest that those factors that determine educational attainment are more strongly transmitted from one generation to the next than those factors that influence earnings.

The latent determinants of income appear particularly persistent in the horizontal dimension. The sibling correlation in the latent factor averages about 0.8, and the shared sibling components in the observable and the latent factor explain a large share of the similarity between siblings (Panel B of Table 8). The shared sibling component in the observable is far more important than the direct transmission of income from parents to children, which explains less than two percent of the variation in income (see Panel D of Table 8).

The spousal correlation in log income is as low as 0.12, consistent with endogenous

Table 8: Calibrated Parameters in Swedish Registers (Income)

Panel A: Intergenerational (Vertical)						
Parameters:						
	β^{m}	β^{f}	γ^{m}	γ^{f}		
	0.133	0.104	0.824	0.550		
	σ^{2}_{ym}	σ^{2}_{yf}	σ^{2}_{zm}	σ^{2}_{zf}	σ^{2}_{um}	σ^{2}_{uf}
	0.304	0.248	0.046	0.019	0.225	0.202
	α_{ym}	α_{yf}	α_{zm}	α_{zf}		
	1.000	0.630	0.239	0.266		
Parent-child correlations in z:						
	Father-Son	Father-Dau	Mother-Son	Mother-Dau		
	0.459	0.494	0.529	0.554		
Ancestor correlations in y and z:						
	Father-Son	Grandf-…	GGrandf-…	GGGrandf-…		
in y	0.227	0.076	0.032	0.014		
in z						
Panel B: Siblings (Horizontal)						
Parameters:						
	σ^{2}_{xm}	σ^{2}_{xf}	σ^{2}_{xmxf}	σ^{2}_{em}	σ^{2}_{ef}	σ^{2}_{emef}
	0.019	0.022	0.008	0.029	0.008	0.014
Variance Shares:						
in y	6.3%	8.9%	2.8%	9.4%	3.2%	5.0%
in z	-	-	-	61.7%	41.9%	46.3%
Sibling correlations in z:						
	Brothers	Sisters	Brother-Sister			
	0.918	0.756	0.781			
Panel C: Assortative Mating (Horizontal)						
Parameters:						
	r^{m}_{zz}	r^{m}_{zy}	r^{m}_{yz}	r^{m}_{yy}	$\sigma^{2}_{\omega m}$	$\sigma^{2}_{\epsilon m}$
	0.381	0.034	0.542	0.021	0.010	0.233
	r^{f}_{zz}	r^{f}_{zy}	r^{f}_{yz}	r^{f}_{yy}		
	1.005	0.022	1.563	0.006		
Spousal correlations in y and z:						
	ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}		
	0.122	0.656	0.390	0.244		
Panel D: Variance Decomposition						
	%	y	z	x	Cov(y,z)	
	Male	0.018	0.153	0.063	0.012	
	Female	0.007	0.076	0.089	0.007	

labor supply decisions at the household level. However, the implied spousal correlation in the latent factor is 0.66. (Panel C of Table 8). This estimate is not as high as for the case of years of schooling, but still far higher than the spousal correlation in most observable characteristics. We therefore find again that spouses are much more similar in the determinants of future socioeconomic success than they are in observable characteristics.

In sum, these estimates confirm that our qualitative findings extend to socioeconomic outcomes other than educational attainment. However, the strength of different transmission processes does vary with the outcome under study. One interpretation is that different factors influence different aspects of socioeconomic status, and that some of those factors have higher persistence than others. In particular, the latent determinants of income are less strongly transmitted than the determinants of educational attainment. This finding is interesting because the recent multigenerational literature does not provide much evidence on income – because income is either not observed, or not very informative in historical sources. Our "horizontal" approach does not face such constraints and can therefore be used to study transmission across a broader set of outcomes.

6.2 Height

We observe body height from universal military enlistment tests, as described in Section 3.1. Its analysis will be interesting for a number of reasons. First, we observe body height only for males. It is an interesting question to what extent female outcomes need to be observed to identify the full model, including the assortative and gender-specific processes. Second, the underlying transmission mechanisms for height are better understood than for socioeconomic outcomes. Variation in body height is primarily due to the influence of genes, at least in populations that were not exposed to famine or undernutrition.[35] Moreover, height is known to be less affected by assortative processes, as already noted by Galton (1886). These properties make it easier to evaluate the plausibility of our results for height than for socioeconomic outcomes.

In contrast to the twin or adoption studies in behavioral genetics we use a more general source of variation, and do not attempt to separate genetic from behavioral factors. Instead, our approach separates transmission channels that are directly related to the

---

[35] The correlation in body height is much higher in biological than foster families, and can be as high as 0.99 for monozygotic twins. The proportion of the total variation in body height in a population that is due to genetic variation is estimated to be around 0.8 (Silventoinen 2003).

Table 9: Estimated and Calibrated Moments in Swedish Registers (Height)

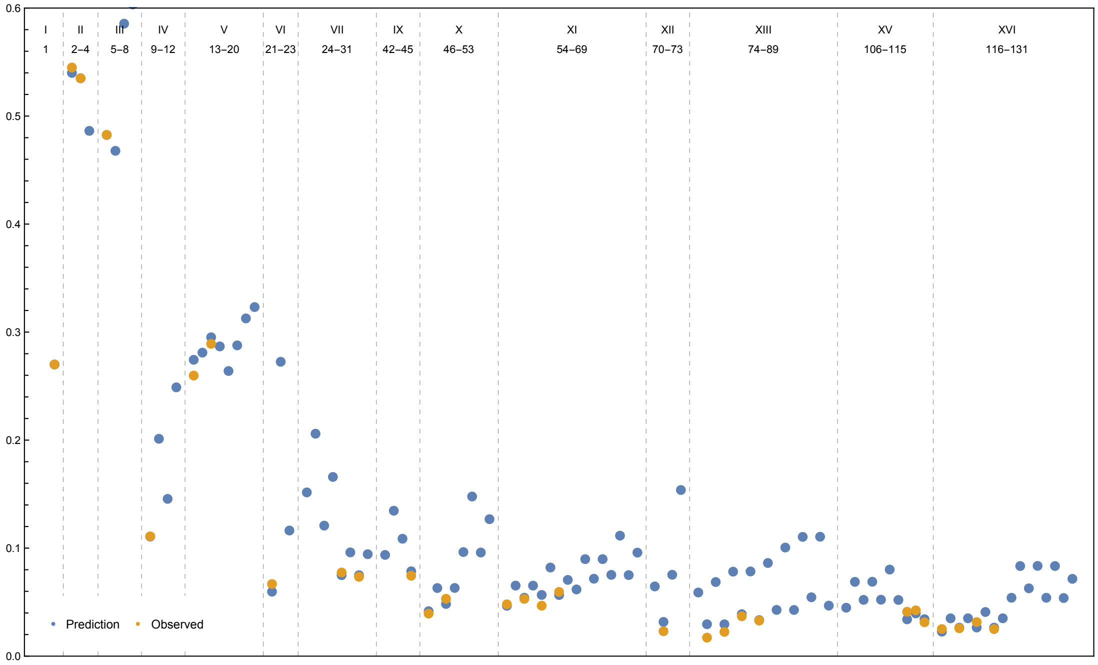
Kinship type			Data		Calibration	
#	##	name	number of pairs	sample correlation	predicted correlation general	genetic
			(1)	(2)	(3)	(4)
I	1	HUSB-WIF	External	0.270	0.270	0.287
II	2	BROTHERS	16,437	0.545	0.540	0.549
	3	SISTERS	External	0.535	0.535	0.540
	4	BROTH-SIS			0.486	0.477
III	5	FATH-SON	46,441	0.483	0.483	0.470
	6	FATH-DAUGH			0.468	0.470
	7	MOTH-SON			0.586	0.470
	8	MOTH-DAUGH			0.603	0.470
IV	9	BL-HS	135,006	0.111	0.111	0.137
	10	BSL-WB			0.201	0.158
	11	BSL-HS			0.146	0.155
	12	SL-WB			0.249	0.137
V	13	NEU-BF	52,618	0.260	0.274	0.279
	…					
	15	NEU-BM	66,270	0.289	0.295	0.288
VI	21	BL-WWS	44,034	0.067	0.060	0.045
VII	28	NEHASF	28,901	0.077	0.075	0.083
	30	NEHASM	39,634	0.073	0.075	0.089
VIII	32	MC-B	21,153	0.160	0.152	0.163
	33	MC-S	21,937	0.188	0.188	0.186
	34	MC-BS	45,689	0.167	0.153	0.175
IX	45	XFMM	99,219	0.074	0.079	0.075
X	46	SIL-MMM	59,544	0.039	0.042	0.019
	48	SIL-MFM	63,798	0.053	0.048	0.024
XI	54	M-MMM	46,771	0.048	0.047	0.040
	…					
XI	56	M-MFM	37,029	0.053	0.054	0.049
	58	M-FMM	43,068	0.047	0.057	0.038
	60	M-FFM	29,437	0.059	0.057	0.045
XII	71	MMFS	39,288	0.023	0.032	0.006
XIII	75	M-MMFS	29,311	0.017	0.030	0.013
	77	M-MFFS	21,488	0.022	0.030	0.012
	79	M-FMFS	26,207	0.037	0.039	0.012
	81	M-FFFS	16,828	0.033	0.033	0.011
XIV	90	M-MMM-M	24,274	0.025	0.034	0.024
	91	M-MMF-M	27,032	0.043	0.032	0.026
	92	M-MFM-M	27,089	0.040	0.040	0.025
	93	M-FMM-M	25,681	0.036	0.043	0.023
XV	113	XFMMM	50,330	0.041	0.034	0.010
	114	XFMFM	25,234	0.042	0.040	0.013
	115	XFFMM	29,208	0.031	0.034	0.010
XVI	116	SIL-MMMM	30,183	0.025	0.023	0.003
	118	SIL-MMFM	30,538	0.026	0.026	0.003
	120	SIL-MFMM	36,879	0.032	0.026	0.003
	122	SIL-MFFM	33,478	0.025	0.026	0.003
XII	133	MMMFS	19,846	0.016	0.018	0.001
XIII	134	MMMMM	16,980	0.020	0.013	0.000
XIX	137	MMMMFS	11,286	0.009	0.010	0.000
XX	*138*	*MMMMMM*	*10,309*	*0.015*	*0.007*	*0.000*
XXI	*141*	*MMMMMFS*	*6,873*	*0.044*	*0.006*	*0.000*

observed outcome (such as parental height) from latent transmission via factors that are not directly observed by the researcher. Our model nests a standard genetic model (replace observed/latent variables with phenotype/genotype), but in addition (i) we allow for a direct role of observables and (ii) allow for a role of other latent factors other than genes (such as cultural and latent socioeconomic variables).

Table 9 reports the sample correlations (columns (1) and (2)). We observe 39 (male) kinship correlations. Because the military enlistments tests were compulsory only for a specific range of cohorts, the number of observations are smaller than for the other outcomes. However, the moments are still precisely estimated, and in line with prior evidence. For example, the father-son correlation in height in our sample is 0.48, the same value reported by Grönqvist et al. (2017). The sample correlations become very small for the most distant kinship types, but remain always positive.

We use 37 kinship correlations up to siblings in-law of four degrees of separation for the calibration (i.e. excluding kinship groups XX and XXI). In addition, we add the correlation in height between spouses and the corresponding correlation between sisters from external sources. Price and Vandenberg (1980) reports a spousal correlation in height of 0.27 for Swedish couples, while we assume that the gap between male and female sibling pairs in the Swedish sources is as large evidence as for Norweigan siblings

Figure 7: Sample and Predicted Moments (Height)



as reported in Tambs (1980).

We report a subset of the predicted moments in column (3) of Table 9. Moreover, Figure 7 plots the subset of 105 moments that we also included in our baseline calibration for educational attainment. As for education and income, the calibrated model explains our data well, providing a close fit to both vertical and horizontal moments, and for both direct and affinity relationships. However, it appears questionable how well this model would fit moments female kins – the calibrated model predicts much larger correlations for distant kinships that involve female than those that involve male.

Table 10 summarizes our findings. On first sight, the results are less plausible than for other outcomes. In particular, the parent-child correlation in the latent status is very high for sons, but very small for daughters. However, these correlations turn out to be irrelevant, because in contrast to all other outcomes the latent factor explains hardly any of the variation in the outcome. Instead, the transmission process can be well captured by the direct transmission channels – observed height can explain the height of descendants and other family members. This direct transmission channel is very strong for height, with $\beta^m = 0.934$ and $\beta^f = 0.844$. Despite this strong intergenerational transmission, the parent-child correlations in height remain modest (approximately one half), because at $\rho_{y_m y_f} = 0.27$ the spousal correlations in height are comparatively small.

Siblings appear to share some additional environmental influences over and above what can be explained by the heigh of their parents – the shared sibling component explains about 5 percent of the variation in height.

Height follows therefore a very different transmission process than socioeconomic outcomes such as education and income. From a statistical perspective, observed height is far more informative about the transmission process than observed socioeconomic status. In particular, the kinship correlations for more distant relatives can be well approximated by extrapolating from the observed kinship correlations in height within the nuclear family. And since spouses do not sort strongly in height, those findings suggest much higher mobility in height than in socioeconomic status for kins beyond the immediate family.

7 Interpretation

Our main objective is to quantify how strongly advantages are transmitted from one generation to the next, and to distinguish the role of intergenerational, sibling and assortative processes. To capture their full strength we interpreted those processes broadly, and remained agnostic about the specific causal mechanisms that they represent. However, indirectly our results are informative about causal mechanisms, in so far as they have specific statistical implications about the pattern of inequality transmission across kinship types.

In particular, our model nests the standard model of genetic transmission (Section 7.1). We can therefore test how well the genetic model explains the transmission of body height (Section 7.2) or educational advantages (Section 7.3) – as compared to the more general model, which allows for additional, non-genetic pathways. Finally, we test if the pattern of inequality across kins could be consistent with an alternative model that allows for higher-order causal effects within the wider family, such as independent effects of grandparents on their grandchildren (Section **??**).

7.1 The Genetic Model

In this section we adopt the standard model in quantitative genetics to evaluate the correlation between relatives. The fundamental idea here is that the observed outcome or "*phenotype*" is determined by genetic and environmental factors. Each individual

Table 10: Calibrated Parameters in Swedish Registers (Height)

Panel A: Intergenerational (Vertical)						
Parameters:						
	β^{m}	β^{f}	γ^{m}	γ^{f}		
	0.934	0.844	0.781	0.018		
	σ^{2}_{ym}	σ^{2}_{yf}	σ^{2}_{zm}	σ^{2}_{zf}	σ^{2}_{um}	σ^{2}_{uf}
	1.000	1.000	0.070	0.001	0.435	0.465
	α_{ym}	α_{yf}	α_{zm}	α_{zf}		
	0.341	0.389	1.000	0.290		
Parent-child correlations in z:						
	Father-Son	Father-Dau	Mother-Son	Mother-Dau		
	0.781	0.054	0.590	0.046		
Ancestor correlations in y and z:						
	Father-Son	Grandf-...	GGrandf-...	GGGrandf-...		
in y	0.483	0.233	0.113	0.055		
in z						
Panel B: Siblings (Horizontal)						
Parameters:						
	σ^{2}_{xm}	σ^{2}_{xf}	σ^{2}_{xmxf}	σ^{2}_{em}	σ^{2}_{ef}	σ^{2}_{emef}
	0.053	0.070	0.022	0.002	0.001	0.004
Variance Shares:						
in y	5.3%	7.0%	2.2%	0.2%	0.1%	0.4%
in z	-	-	-	3.5%	54.4%	46.1%
Sibling correlations in z:						
	Brothers	Sisters	Brother-Sister			
	0.645	0.547	0.503			
Panel C: Assortative Mating (Horizontal)						
Parameters:						
	r^{m}_{zz}	r^{m}_{zy}	r^{m}_{yz}	r^{m}_{yy}	$\sigma^{2}_{\omega m}$	$\sigma^{2}_{\epsilon m}$
	0.090	-0.006	-2.131	0.264	0.000	0.611
	r^{f}_{zz}	r^{f}_{zy}	r^{f}_{yz}	r^{f}_{yy}		
	6.386	-0.153	-6.046	0.273		
Spousal correlations in y and z:						
	ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}		
	0.270	0.756	-0.187	-0.565		
Panel D: Variance Decomposition						
	%	y	z	x	Cov(y,z)	
	Male	0.534	0.070	0.053	-0.072	
	Female	0.419	0.001	0.070	-0.001	

receives half of its genetic contribution from the father and the other half from the mother. In the simplest form the model assumes that genetic and environmental factors contribute additively to the phenotype, i.e. do not interact and are uncorrelated. Assortative mating is based on similarity in the phenotype and the population is at the steady state. In Appendix E we show that such standard quantitative genetic model is a special case of our general model. Specifically, we obtain formulas for the correlations between relatives that coincide with the ones in, for example, Crow and Felsenstein (1968).

The genetic model is nested in our general model by imposing three sets of restrictions. First, parents outcome (phenotype) does not have any direct association with the child outcome (imposing the restrictions $\beta^k = 0$, $k = f, m$). Second, because the latent factor represents genes it is transmitted from parents to children as

$$z_t^k = \frac{z_{t-1}^m + z_{t-1}^f}{2} + v_t^k \tag{9}$$

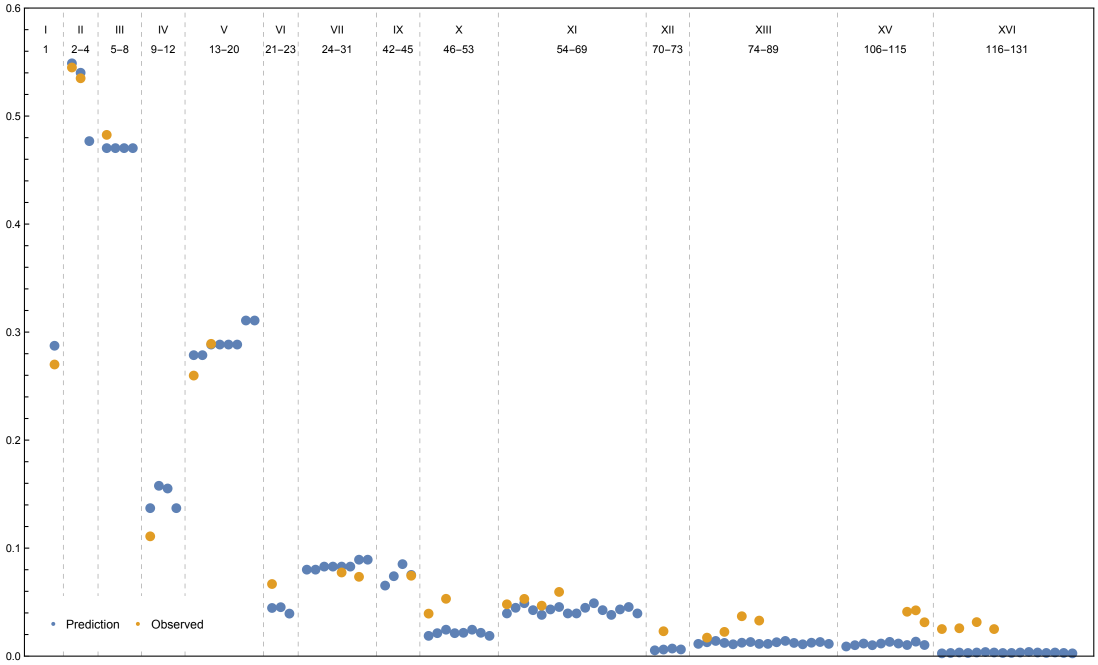
where v_t^k is uncorrelated across relatives and to z_{t-1}^m and z_{t-1}^f (imposing $\gamma^k = 1$, $\alpha_z^k = 0.5$, and $\sigma^2_{e_k} = 0$, $k = f, m$). Finally, assortative mating occurs only in the outcome (phenotype) y. As a consequence, $\rho_{z^m y^f}$, $\rho_{y^m z^f}$ and $\rho_{z^m z^f}$ are functions of $\rho_{y^m y^f}$ and some of the other parameters of the model (see Section 2.1).

The genetic model has only 5 parameters, which account for the share of variance in the phenotype explained by the genotype ($\sigma^2_{z^m} = \sigma^2_{z^f} = \sigma^2_z$), the environmental effects shared by siblings (the *heritability* $\sigma^2_{x^m}$, $\sigma^2_{x^f}$ and $\sigma_{x^m x^f}$), and the assortative mating in the phenotype ($\rho_{y^m y^f}$). Those not familiar with this model might wonder why equation (9) contains an error term, given that children inherit their genes exclusively from their parents. The explanation is that a genotype is defined by a combination of genes/alleles, and different children receive different combinations (with the exception of monozygotic twins). This result is derived formally in Section E.

7.2 The Genetic Model and *Height*

We first calibrate the genetic model for *body height*. Because genes are known to be the primary source of variation in height, the genetic model should provide a good fit to the wide set of kinship correlations that we observe (see Table 9) – even though this model is severely restricted compared to our baseline model (with 5 instead of 20 parameters).

Figure 8: Sample and Predicted Moments (Height, Genetic Model)



For calibration we use the same the 39 kinship correlations as used for calibration of our general model, including the spousal and sister correlation in height from external sources (see Section 6.2). Column (4) of Table 9 reports the predicted correlation for kinship types I-IV, as well as the more distant kinships observed in our data. For comparison, Figure plots the subset of 105 moments that we also reported for all other applications.

The genetic model fits the data worse than the general model. The mean prediction error up to second-degree sibling in-laws is only 3.7 percent for the general model, but 13.6 percent for the genetic model. However, the genetic model does fit the pattern among close relatives such as spouses and siblings quite well, and produces out-of-sample predictions that are more stable within kinship types than the widely varying predictions of the general model.

Table 11 summarizes the calibration results for the genetic model.[36] The share of variation in height explained by the latent factor is about 73 percent (see Panel A). This estimate is slightly below estimates of the heritability of height from quantitative genetics, which cluster around 80 percent (Silventoinen, 2003). In principle, our approach

---

[36] In contrast to the standard model described above, we allow $\sigma^2_{e_k}$ to be non-zero in the calibration (e.g. to account for the presence of monozygotic twins in our data). This added flexibility has only a negligible effect on the results.

Table 11: Calibrated Parameters in Swedish Registers (Height, Genetic Model)

Panel A: Intergenerational (Vertical)						
Parameters:						
	β^m	β^f	γ^m	γ^f		
	0.000	0.000	1.000	1.000		
	σ^2_{ym}	σ^2_{yf}	σ^2_{zm}	σ^2_{zf}	σ^2_{um}	σ^2_{uf}
	1.000	1.000	0.731	0.731	0.163	0.237
	α_{ym}	α_{yf}	α_{zm}	α_{zf}		
	0.000	0.000	0.500	0.500		
Parent-child correlations in z:						
	Father-Son	Father-Dau	Mother-Son	Mother-Dau		
	0.605	0.605	0.605	0.605		
Ancestor correlations in y and z:						
	Father-Son	Grandf-...	GGrandf-...	GGGrandf-...		
in y	0.470	0.285	0.172	0.104		
in z						
Panel B: Siblings (Horizontal)						
Parameters:						
	σ^2_{xm}	σ^2_{xf}	σ^2_{xmxf}	σ^2_{em}	σ^2_{ef}	σ^2_{emef}
	0.107	0.032	0.000	0.000	0.066	0.035
Variance Shares:						
in y	10.7%	3.2%	0.0%	0.0%	6.6%	3.5%
in z	-	-	-	0.0%	9.0%	4.8%
Sibling correlations in z:						
	Brothers	Sisters	Brother-Sister			
	0.605	0.695	0.653			
Panel C: Assortative Mating (Horizontal)						
Parameters:						
	r^m_{zz}	r^m_{zy}	r^m_{yz}	r^m_{yy}	$\sigma^2_{\omega m}$	$\sigma^2_{\epsilon m}$
	0.000	0.210	0.000	0.287	0.687	0.917
	r^f_{zz}	r^f_{zy}	r^f_{yz}	r^f_{yy}		
	0.000	0.210	0.000	0.287		
Spousal correlations in y and z:						
	ρ_{ymyf}	ρ_{zmzf}	ρ_{ymzf}	ρ_{zmyf}		
	0.287	0.210	0.246	0.246		
Panel D: Variance Decomposition						
	%	y	z	x	Cov(y,z)	
	Male	0.000	0.731	0.107	0.000	
	Female	0.000	0.731	0.032	0.000	

can capture the transmission of latent factors other than genes. The observation that our estimates are close to those from the genetic literature suggests therefore that the parent-child correlation in height is nearly exclusively due to genes, and not other factors.

The calibration further suggests that the remaining variation in height is predominantly explained by factors that are *not* shared by siblings (Panel A), while shared environmental factors are less important (Panel B). Finally, our results imply that the spousal correlation in phenotype height is greater than the spousal correlation in genotype height (Panel C). This result follows directly from the assumption that assortative mating occurs exclusively on the phenotype. In contrast, our general model allowing for assortative mating in another dimension, and therefore for spouses to be more similar than indicated by their observed outcomes (phenotypes). As a consequence, the genetic and general model yield substantially different predictions for the distant kinship types in our data (cf. ancestor correlations reported in Tables 10 and 11, and Table 9).

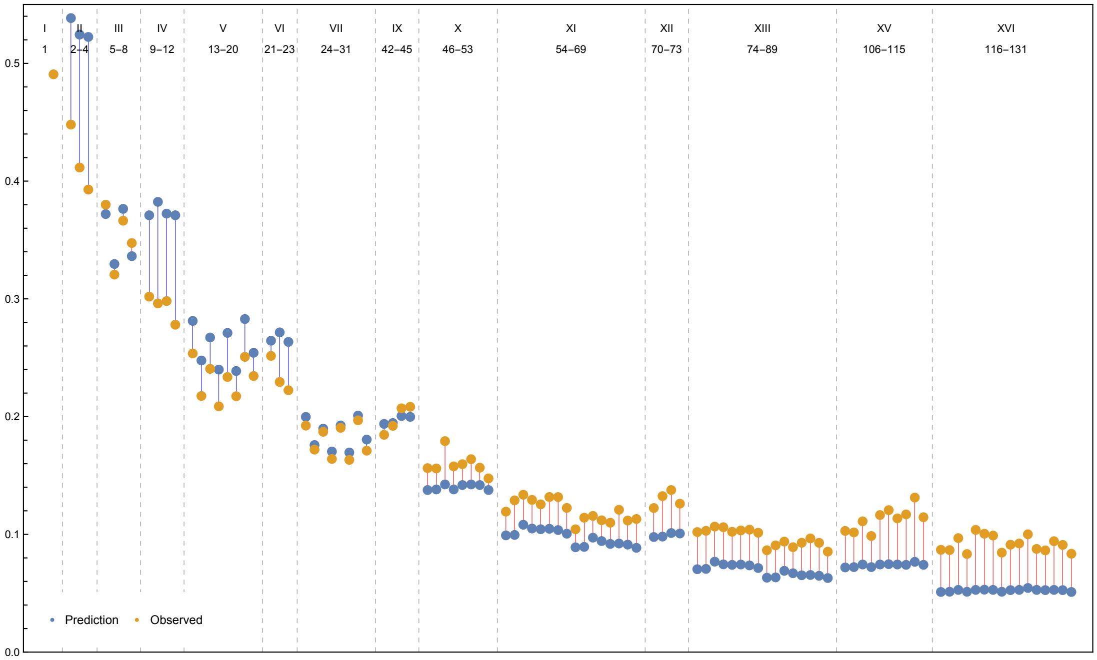
The results are therefore ambiguous. On the one hand, our approach yields parameter estimates that match estimates of the heritability of body height from quantitative genetics. However, the standard genetic model cannot explain why body height remains correlated between the most distant family members in our data. The genetic model fits a narrow set of kinship moments, but not the full set observed in this study.

A potential culprit is the assumption that assortative mating occurs only in phenotype height. If spouses match on other factors apart from the phenotype (observed) height, and those other factors have an independent association with genotype height, then the standard genetic model understates the correlation in phenotype height between relatives – even if the intergenerational transmission process is exclusively due to genetic factors. The advantage of our approach is that such failures are visible. Because we observe such distant relatives, erroneous assumptions in the assortative process can be detected even if that error has only negligible implications for the type of close kinship correlations that have been studied in the economic literature.

7.3 The Genetic Model and *Education*

We next calibrate the genetic model for *years of schooling* in the Swedish registers. For calibration, we use the same 105 kinships that we used for calibration of our baseline model (see Section 4), but impose the restrictions implied by the standard genetic model

Figure 9: Sample and Predicted Moments (Education, Genetic Model)

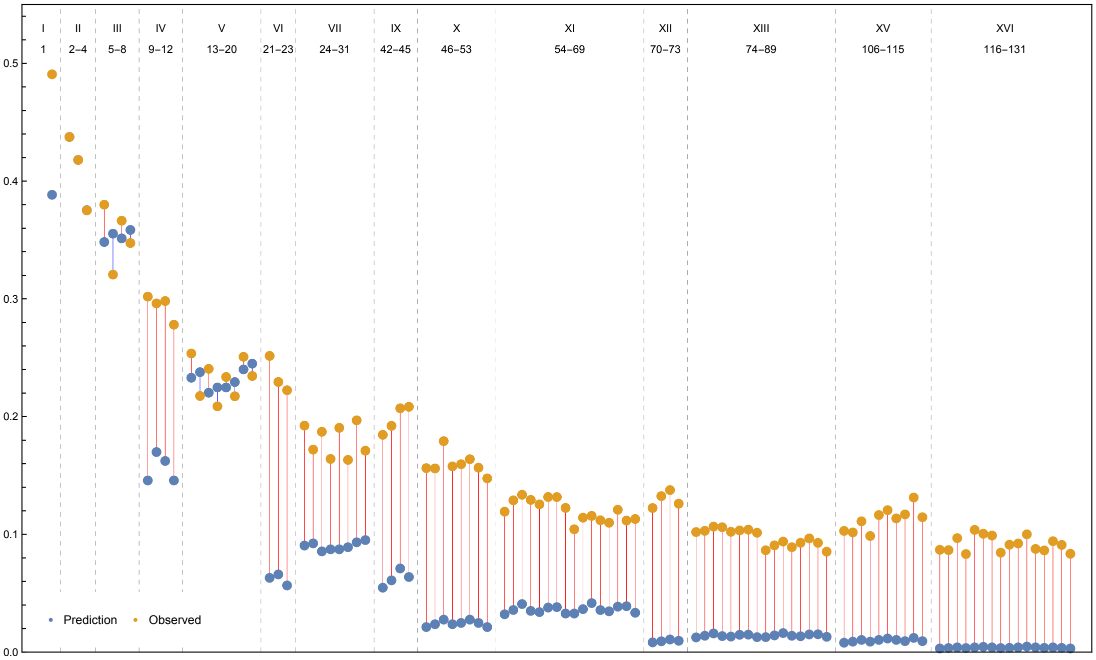


(see Section 7.1). Figure 9 illustrates the in-sample fit graphically, to be compared against the fit of our baseline model (Figure 2).

The standard genetic model cannot fit the correlation in educational advantages between relatives. Its mean prediction error is 59.7 percent, compared to only 4.5 percent for our baseline model calibrated in Section 4.2. In particular, the genetic model overstates kinship correlations within the nuclear family (e.g. for brothers, 0.538 predicted vs. 0.448 in sample), and understates them for more distant relatives (e.g. third-degree sibling in-law, 0.051 predicted vs. 0.084 in sample). The genetic model provides a much worse fit for the kinship correlations in education than in body height – in particular considering that the latter are smaller, such that the prediction errors tend to be larger in percentage terms.

Of course, one might not find it too surprising that the genetic model cannot explain the transmission of socioeconomic outcomes. After all, the belief that non-genetic pathways matter is a core motivation for mobility research in the social sciences. For example, economic models tend to be based on the assumption that human capital is subject to investment decisions, ascribing parental income a key role in the transmission process (Mogstad, 2017). However, empirical work suggests that such investment channels play only a limited role, and confirms that genetic pathways are indeed very important

Figure 10: Sample and Predicted Moments (Education, Genetic Model, 15 Moments)



(Björklund et al., 2006). And even in the social sciences, some authors argue that the correlation of socioeconomic advantages between relatives is predominantly due to genes. This argument has received new support by the observation that the transmission of socioeconomic inequalities across multiple generations can be remarkably similar across countries and time periods (Clark, 2014).

In contrast, our findings are inconsistent with a purely genetic interpretation. We base this rejection on a simple statistical observation – the standard genetic model is not able to provide a close in-sample fit to the distribution of educational advantages across kins. However, our observations are not necessarily inconsistent with the view that genes are the decisive factor in the transmission from parents to their children. Rather, we reject the *standard* genetic model from quantitative genetics, in particular the assumption that assortative mating occurs only on the phenotype. In contrast to this assumption, we find – across all applications – that spouses must be much more similar in fundamental determinants of socioeconomic success than what can be discerned from spousal correlations in socioeconomic outcomes themselves.

This finding relies heavily on the observation of distant sibling in-laws – intuitively, failure of the assortative assumptions baked into the standard genetic model become most obvious when trying to explain kinship correlations that are a function of multiple assortative matches. In contrast, the genetic model *appears* to perform well when fitting a restricted set of kinship correlations that includes only close kins.

To illustrate this point, we re-calibrated the genetic model based only on sibling correlations (group II in Table 3), parent-child correlations (group III), and nephew/niece-uncle/aunt correlations (group V). As illustrated in Figure 10, the genetic model provides an excellent in-sample fit for those 15 close kinships. However, it provides an extremely bad fit to all other moments observed in the data. This observation is important, as most economic research to date has been based on even narrower set of kinship correlations than the set used in this exercise. It suggests that in order to evaluate and discriminate between intergenerational model based on their statistical properties, researchers need to observe a sufficiently wide range of sufficiently distant kinship correlations. Our horizontal approach applied to population-wide data from administrative sources offers that opportunity.

8 Conclusions

– to be added –

References

Adermon, A., M. Lindahl, and M. Palme (2016): "Dynastic Human Capital, Inequality and Intergenerational Mobility," Discussion paper, mimeo, Uppsala University.

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

Anderson, L. R., P. Sheppard, and C. W. S. Monden (2018): "Grandparent Effects on Educational Outcomes: A Systematic Review," *Sociological Science*.

Atkinson, A. B., and S. Jenkins (1984): "The Steady-State Assumption and the Estimation of Distributional and Related Models," *Journal of Human Resources*, 19(3), 358–376.

Barone, G., and S. Mocetti (2016): "Intergenerational Mobility in the Very Long Run: Florence 1427-2011," Temi di discussione (Economic working papers) 1060, Bank of Italy, Economic Research and International Relations Area.

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

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

Behrman, J., and M. Rosenzweig (2002): "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?," *American Economic Review*, 92(1), 323–334.

Behrman, J. R., and P. Taubman (1989): "Is Schooling "Mostly in the Genes"? Nature-Nurture Decomposition Using Data on Relatives," *Journal of Political Economy*, 97(6), 1425–1446.

Björklund, A., M. Jäntti, and M. J. Lindquist (2009): "Family Background and Income During the Rise of the Welfare State: Brother Correlations in Income for Swedish Men born 1932-1968," *Journal of Public Economics*, 93(5-6), 671 – 680.

Björklund, A., M. Jäntti, and G. Solon (2005): "Influences of Nature and Nurture on Earnings Variation," in *Unequal Chances. Family Background and Economic*

Success, ed. by S. Bowles, H. Gintis, and M. O. Groves. Princeton University Press, New York.

Björklund, A., M. Lindahl, and E. Plug (2006): "The Origins of Intergenerational Associations: Lessons from Swedish Adoption Data," *Quarterly Journal of Economics*, 121(3), 999–1029.

Björklund, A., and K. Salvanes (2010): "Education and Family Background: Mechanisms and Policies," in *Handbook of the Economics of Education*, ed. by E. Hanushek, S. Machin, and L. Woessmann, vol. 3.

Black, S. E., and P. Devereux (2011): "Recent Developments in Intergenerational Mobility," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 4A. Elsevier.

Blanden, J. (2011): "Cross-Country Rankings in Intergenerational Mobility: A Comparison of Approaches from Economics and Sociology," *Journal of Economic Surveys*.

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

Breen, R. (2018): "Some Methodological Problems in the Study of Multigenerational Mobility," *European Sociological Review*, p. jcy037.

Calero, J., J. O. Escardíbul, S. Waisgrais, and M. Mediavilla (2007): "Desigualdades socioeconómicas en el sistema educativo español," Discussion paper, Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación y Ciencia.

Cervini-Plá, M. (2015): "Intergenerational Earnings and Income Mobility in Spain," *Review of Income and Wealth*, 61(4), 812–828.

Chetty, R., and N. Hendren (2018): "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *Quarterly Journal of Economics*.

Chetty, R., N. Hendren, P. Kline, and E. Saez (2014): "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States," *Quarterly Journal of Economics*, 129(4), 1553–1623.

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

——— (2017): "Nature Versus Nurture in Social Outcomes. A Lineage Study of 66,000 English Individuals, 1750-2016," Discussion paper, mimeo.

——— (2018): "Estimating Social Mobility Rates from Surnames: Social Group or Dynastic Transmission versus Family Effects," Discussion paper, mimeo.

Clark, G., and N. Cummins (2012): "What is the True Rate of Social Mobility? Surnames and Social Mobility, England 1800-2012," Unpublished working paper.

Clark, G., N. Cummins, Y. Hao, and D. D. Vidal (2015): "Surnames: A new source for the history of social mobility," *Explorations in Economic History*, 55, 3 – 24.

Collado, M. D., I. Ortuño-Ortin, and A. Romeu (2014): "Long-run Intergenerational Social Mobility and the Distribution of Surnames," Discussion Paper 2, UMUFAE Economics Working Papers.

Corak, M. (2013): "Income Inequality, Equality of Opportunity, and Intergenerational Mobility," *Journal of Economic Perspectives*, 27(3), 79–102.

Crow, J. F., and J. Felsenstein (1968): "The effect of assortative mating on the genetic composition of a population," *Eugenics Quarterly*, 15(2), 85–97.

Diaz-Vidal, D., and G. Clark (2015): "How Strong is Assortative Mating? A Surname Analysis," Discussion paper, Working Paper.

Dribe, M., and J. 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, O. D. (1969): "Contingencies in Constructing Causal Models," *Sociological Methodology*, 1, 74–112.

Eika, L., M. Mogstad, and B. Zafar (2014): "Educational Assortative Mating and Household Income Inequality," Working Paper 20271, National Bureau of Economic Research.

Ermisch, J., M. Francesconi, and T. Siedler (2006): "Intergenerational Mobility and Marital Sorting," *The Economic Journal*, 116(513), 659–679.

et al, A. O. (2016): "Genome-wide association study identifies 74 loci associated with educational attainment," *Nature*, 533, 539 EP –.

FERNÁNDEZ, R., AND R. ROGERSON (2001): "Sorting and Long-Run Inequality," *The Quarterly Journal of Economics*, 116(4), 1305–1341.

GALTON, F. (1886): "Regression towards mediocrity in hereditary stature," *The Journal of the Anthropological Institute of Great Britain and Ireland*, 15, 246–263.

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

——— (1979): "Heritability," *Economica*, 46(184), pp. 327–347.

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

GOLDTHORPE, J. H. (2014): "The role of education in intergenerational social mobility: Problems from empirical research in sociology and some theoretical pointers from economics," *Rationality and Society*, 26(3), 265–289.

GREENWOOD, J., N. GUNER, G. KOCHARKOV, AND C. SANTOS (2014): "Marry Your Like: Assortative Mating and Income Inequality," *American Economic Review: Papers and Proceedings*, 104(5), 348–53.

GRÖNQVIST, E., B. ÖCKERT, AND J. VLACHOS (2017): "The Intergenerational Transmission of Cognitive and Noncognitive Abilities," *Journal of Human Resources*, 52(4), 887–918.

GÜELL, M., J. V. R. MORA, AND G. SOLON (2018): "New Directions in Measuring Intergenerational Mobility: Introduction," *The Economic Journal.*

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(0), 19 – 33, Inequality Across Multiple Generations.

HAVNES, T., AND M. MOGSTAD (2015): "Is Universal Child Care Leveling the Playing Field?," *Journal of Public Economics*, 127, 100 – 114, The Nordic Model.

HERTZ, T. (2007): "Trends in the Intergenerational Elasticity of Family Income in the United States," *Industrial Relations*, 46(1), 22–50.

HERTZ, T., T. JAYASUNDERA, P. PIRAINO, S. SELCUK, N. SMITH, AND A. VERASHCHAGINA (2008): "The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends," *The B.E. Journal of Economic Analysis & Policy*, 7(2), 1–48.

Holmlund, H., M. Lindahl, and E. Plug (2011): "The Causal Effect of Parents' Schooling on Children's Schooling: A Comparison of Estimation Methods," *Journal of Economic Literature*, 49(3), 615–51.

Jäntti, M., and S. P. Jenkins (2014): "Income Mobility," in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, vol. 2. Elsevier.

Kim, C., C. R. Tamborini, and A. Sakamoto (2018): "The Sources of Life Chances: Does Education, Class Category, Occupation, or Short-Term Earnings Predict 20-Year Long-Term Earnings?," *Sociological Science*, 5, 206–233.

Lee, J. J., R. Wedow, A. Okbay, E. Kong, O. Maghzian, M. Zacher, T. A. Nguyen-Viet, P. Bowers, J. Sidorenko, R. Karlsson Linnér, M. A. Fontana, T. Kundu, C. Lee, H. Li, R. Li, R. Royer, P. N. Timshel, R. K. Walters, E. A. Willoughby, L. Yengo, M. Agee, B. Alipanahi, A. Auton, R. K. Bell, K. Bryc, S. L. Elson, P. Fontanillas, D. A. Hinds, J. C. McCreight., K. E. Huber, N. K. Litterman, M. H. McIntyre, J. L. Mountain, E. S. Noblin, C. A. M. Northover, S. J. Pitts, J. F. Sathirapongsasuti, O. V. Sazonova, J. F. Shelton, S. Shringarpure, C. Tian, V. Vacic, C. H. Wilson, J. P. Beauchamp, J. J. Lee, T. H. Pers, C. A. Rietveld, P. Turley, G.-B. Chen, V. Emilsson, S. F. W. Meddens, S. Oskarsson, J. K. Pickrell, K. Thom, P. Timshel, R. d. Vlaming, A. Abdellaoui, T. S. Ahluwalia, J. Bacelis, C. Baumbach, G. Bjornsdottir, J. H. Brandsma, M. P. Concas, J. Derringer, N. A. Furlotte, T. E. Galesloot, G. Girotto, R. Gupta, L. M. Hall, S. E. Harris, E. Hofer, M. Horikoshi, J. E. Huffman, K. Kaasik, I. P. Kalafati, R. Karlsson, A. Kong, J. Lahti, S. J. van der Lee, C. d. Leeuw, P. A. Lind, K.-O. Lindgren, T. Liu, M. Mangino, J. Marten, E. Mihailov, M. B. Miller, P. J. van der Most, C. Oldmeadow, A. Payton, N. Pervjakova, W. J. Peyrot, Y. Qian, O. Raitakari, R. Rueedi, E. Salvi, B. Schmidt, K. E. Schraut, J. Shi, A. V. Smith, R. A. Poot, B. St Pourcain, A. Teumer, G. Thorleifsson, N. Verweij, D. Vuckovic, J. Wellmann, H.-J. Westra, J. Yang, W. Zhao, Z. Zhu, B. Z. Alizadeh, N. Amin, A. Bakshi, S. E. Baumeister, G. Biino, K. Bønnelykke, P. A. Boyle, H. Campbell, F. P. Cappuccio, G. Davies, J.-E. De Neve, P. Deloukas, I. Demuth, J. Ding, P. Eibich, L. Eisele, N. Eklund, D. M. Evans, J. D. Faul, M. F. Feitosa, A. J. Forstner, I. Gandin, B. Gunnarsson, B. V. Halldórsson, T. B. Harris, A. C. Heath, L. J. Hocking, E. G. Holliday, G. Homuth, M. A. Horan, J.-J. Hottenga, P. L. de Jager, P. K. Joshi, A. Jugessur, M. A. Kaakinen, M. Kähönen, S. Kanoni, L. Keltigangas-Järvinen, L. A. L. M. Kiemeney, I. Kolcic, S. Koskinen, A. T. Kraja, M. Kroh, Z. Kutalik, A. Latvala, L. J. Launer, M. P. Lebreton, D. F. Levinson, P. Lichtenstein, P. Lichtner, D. C. M. Liewald, A. Loukola, LifeLines Cohort Study, P. A. Madden, R. Mägi, T. Mäki-Opas, R. E. Marioni, P. Marques-Vidal, G. A. Meddens, G. McMahon, C. Meisinger, T. Meitinger, Y. Milaneschi, L. Milani, G. W. Montgomery, R. Myhre, C. P. Nelson, D. R. Nyholt, W. E. R. Ollier, A. Palotie, L. Paternoster, N. L. Pedersen, K. E. Petrovic, D. J. Porteous, K. Räikkönen, S. M. Ring, A. Robino, O. Rostapshova, I. Rudan, A. Rustichini, V. Salomaa, A. R. Sanders, A.-P. Sarin, H. Schmidt, R. J. Scott, B. H. Smith, J. A. Smith, J. A. Staessen, E. Steinhagen-Thiessen, K. Strauch, A. Terracciano, M. D. Tobin, S. Ulivi, S. Vaccargiu, L. Quaye, F. J. A. van Rooij, C. Venturini, A. A. E. Vinkhuyzen, U. Völker, H. Völzke, J. M. Vonk, D. Vozzi, J. Waage, E. B. Ware, G. Willemsen, J. R. Attia, D. A. Bennett, K. Berger, L. Bertram, H. Bisgaard, D. I. Boomsma, I. B. Borecki, U. Bültmann, C. F. Chabris, F. Cucca, D. Cusi, I. J. Deary, G. V. Dedoussis, C. M. van Duijn, J. G. Eriksson, B. Franke, L. Franke, P. Gasparini, P. V. Gejman, C. Gieger, H.-J. Grabe, J. Gratten, P. J. F. Groenen, V. Gudnason, P. van der Harst, C. Hayward, D. A. Hinds, W. Hoffmann, E. Hyppönen, W. G. Iacono, B. Jacobsson, M.-R. Järvelin, K.-H. Jöckel, J. Kaprio, S. L. R. Kardia, T. Lehtimäki, S. F. Lehrer, P. K. E. Magnusson, N. G. Martin, M. McGue, A. Metspalu, N. Pendleton, B. W. J. H. Penninx, M. Perola, N. Pirastu, M. Pirastu, O. Polasek, D. Posthuma, C. Power, M. A. Province, N. J. Samani, D. Schlessinger, R. Schmidt, T. I. A. Sørensen, T. D. Spector, K. Stefansson, U. Thorsteinsdottir, A. R. Thurik, N. J. Timpson, H. Tiemeier, J. Y. Tung, A. Uitterlinden, V. Vitart, P. Vollenweider, D. R. Weir, J. F. Wilson, A. F. Wright, D. C. Conley, R. F. Krueger, G. D. Smith, A. Hofman, D. I. Laibson, S. E. Medland, M. N. Meyer, J. Yang, M. Johannesson, P. M. Visscher, T. Esko, P. D. Koellinger, D. Cesarini, 23andMe Research Team, C. C. G. Consortium), and S. S. G. A. Consortium (2018): "Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals," *Nature Genetics*, 50(8), 1112–1121.

Levine, D. I., and B. Mazumder (2007): "The Growing Importance of Family: Evidence from Brothers' Earnings," *Industrial Relations: A Journal of Economy and Society*, 46(1), 7–21.

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

Long, J., and J. Ferrie (2013a): "Intergenerational Occupational Mobility in Great Britain and the United States since 1850," *American Economic Review*, 103(4), 1109–37.

——— (2013b): "Intergenerational Occupational Mobility in Great Britain and the United States since 1850: Reply," *American Economic Review*, 103(5), 2041–49.

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

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

Modalsli, J. (2017): "Intergenerational Mobility in Norway, 1865–2011," *The Scandinavian Journal of Economics*, 119(1), 34–71.

Mogstad, M. (2017): "The Human Capital Approach to Intergenerational Mobility," *Journal of Political Economy*, 125(6), 1862–1868.

Neidhöfer, G., and M. Stockhausen (forthcoming): "Dynastic Inequality Compared: Multigenerational Mobility in the United States, the United Kingdom, and Germany," *Review of Income and Wealth*, 0(0).

Nybom, M., and J. Stuhler (2014): "Interpreting Trends in Intergenerational Mobility," Working Paper Series 3/2014, Swedish Institute for Social Research.

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

Piketty, T. (2000): "Theories of Persistent Inequality and Intergenerational Mobility," in *Handbook of Income Distribution*, ed. by A. B. Atkinson, and F. Bourguignon, vol. 1, pp. 429 – 476. Elsevier.

Price, R. A., and S. G. Vandenberg (1980): "Spouse similarity in American and Swedish couples," *Behavior genetics*, 10(1), 59–71.

Raaum, O., B. Bratsberg, K. Røed, E. Österbacka, T. Eriksson, M. Jäntti, and R. Naylor (2007): "Marital Sorting, Household Labor Supply, and Intergenerational Earnings Mobility across Countries," *The B.E. Journal of Economic Analysis & Policy*, 7(2), Article 7.

SACERDOTE, B. (2011): "Nature and Nurture Effects on Children's Outcomes: What Have We Learned From Studies of Twins and Adoptees?," in *Handbook of Social Economics*, ed. by J. Benhabib, A. Bisin, and M. Jackson, vol. 1. Elsevier, San Diego.

SILVENTOINEN, K. (2003): "Determinants of variation in adult body height," *Journal of biosocial science*, 35(2), 263–285.

SOLON, G. (1999): "Intergenerational Mobility in the Labor Market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, chap. 29, pp. 1761–1800. Elsevier.

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

SOLON, G., M. CORCORAN, G. GORDON, AND D. LAREN (1991): "A Longitudinal Analysis of Sibling Correlations in Economic Status," *Journal of Human Resources*, 26(3), 509–534.

SOLON, G., M. E. PAGE, AND G. J. DUNCAN (2000): "Correlations Between Neighboring Children In Their Subsequent Educational Attainment," *The Review of Economics and Statistics*, 82(3), 383–392.

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

——— (2018): "A Review of Intergenerational Mobility and its Drivers," JRC Working Papers JRC112247, Joint Research Centre (Seville site).

TORCHE, F., AND A. CORVALAN (2015): "Estimating Intergenerational Mobility with Grouped Data: A Critique of Clark's The Son Also Rises," Discussion Paper Working Paper No. 2015-22, NYU Population Center Working Paper Series.

VOSTERS, K., AND M. NYBOM (2017): "Intergenerational Persistence in Latent Socioeconomic Status: Evidence from Sweden and the United States," *Journal of Labor Economics*, 35(3), 869 – 901.

Theoretical Appendix

A General assortative mating model

We assume that the value of the outcome y for an individual from generation t is given by

$$y_t^k = \beta^k \widetilde{y}_{t-1}^k + z_t^k + x_t^k + u_t^k \tag{10}$$

where the superscript $k = f$ stands for males and $k = m$ for females. We assume that

$$\widetilde{y}_{t-1}^k = \alpha_y^k y_{t-1}^m + (1 - \alpha_y^k) y_{t-1}^f$$

and the socioeconomic status of the child, z_t^k, depends on the father z_{t-1}^m as well as on the mother z_{t-1}^f

$$\begin{aligned} z_t^k &= \gamma^k \widetilde{z}_{t-1}^k + e_t^k + v_t^k \\ \widetilde{z}_{t-1}^k &= \alpha_z^k z_{t-1}^m + (1 - \alpha_z^k) z_{t-1}^f \end{aligned} \tag{11}$$

Regarding the shocks, we assume that x_t^k and e_t^k are shared by all siblings of the same gender, can be correlated across siblings of different gender and are uncorrelated to each other and with the other variables (in particular with z_{t-1}^k and y_{t-1}^l, $l = m, f$). Finally u_t^k and v_t^k are white-noise errors.

A.1 Assortative mating process

We assume there is assortative mating both in years of schooling and in socioeconomic status (see Berhman and Rosenzweig 2002 for a related model with assortative mating in two dimensions). In particular we consider the linear projections of z_{t-1}^f and y_{t-1}^f on z_{t-1}^m and y_{t-1}^m:

$$\begin{pmatrix} z_{t-1}^f \\ y_{t-1}^f \end{pmatrix} = \begin{pmatrix} r_{zz}^m & r_{zy}^m \\ r_{yz}^m & r_{yy}^m \end{pmatrix} \begin{pmatrix} z_{t-1}^m \\ y_{t-1}^m \end{pmatrix} + \begin{pmatrix} w_{t-1}^m \\ \varepsilon_{t-1}^m \end{pmatrix}$$

where w_{t-1}^m and ε_{t-1}^m might be correlated but are uncorrelated with z_{t-1}^m and y_{t-1}^m.

The coefficients of the linear projections depend on $\rho_{z^m y^m}, \rho_{z^m z^f}$, $\rho_{z^m y^f}$, $\rho_{y^m z^f}$ and $\rho_{y^m y^f}$, as well as on the standard deviations of z_{t-1}^k and y_{t-1}^k, $k = m, f$:

$$r_{zz}^m = \frac{1}{(1 - \rho_{z^m y^m}^2)} \frac{\sigma_{z^f}}{\sigma_{z^m}} (\rho_{z^m z^f} - \rho_{z^m y^m} \rho_{y^m z^f})$$

$$r_{zy}^m = \frac{1}{(1 - \rho_{z^m y^m}^2)} \frac{\sigma_{z^f}}{\sigma_{y^m}} (\rho_{y^m z^f} - \rho_{z^m y^m} \rho_{z^m z^f})$$

$$r_{yz}^m = \frac{1}{(1-\rho_{z^m y^m}^2)} \frac{\sigma_{y^f}}{\sigma_{z^m}} (\rho_{z^m y^f} - \rho_{z^m y^m} \rho_{y^m y^f})$$

$$r_{yy}^m = \frac{1}{(1-\rho_{z^m y^m}^2)} \frac{\sigma_{y^f}}{\sigma_{y^m}} (\rho_{y^m y^f} - \rho_{z^m y^m} \rho_{z^m y^f})$$

The variance matrix of $(w_{t-1}^m, \varepsilon_{t-1}^m)$ is given by

$$Var \begin{pmatrix} w_{t-1}^m \\ \varepsilon_{t-1}^m \end{pmatrix} = Var \begin{pmatrix} z_{t-1}^f \\ y_{t-1}^f \end{pmatrix} - \begin{pmatrix} r_{zz}^m & r_{zy}^m \\ r_{yz}^m & r_{yy}^m \end{pmatrix} Var \begin{pmatrix} z_{t-1}^m \\ y_{t-1}^m \end{pmatrix} \begin{pmatrix} r_{zz}^m & r_{zy}^m \\ r_{yz}^m & r_{yy}^m \end{pmatrix}'$$

We use these matching functions to write years of schooling, y_t^k, and social status, z_t^k, as a function of father's years of schooling, y_{t-1}^m, and social status z_{t-1}^m. We can write (11) as

$$\begin{aligned} z_t^k &= \gamma^k \left(\alpha_z^k z_{t-1}^m + (1-\alpha_z^k) z_{t-1}^f \right) + e_t^k + v_t^k \\ &= \gamma^k \left(\alpha_z^k z_{t-1}^m + (1-\alpha_z^k) \left(r_{zz}^m z_{t-1}^m + r_{zy}^m y_{t-1}^m + w_{t-1}^m \right) \right) + e_t^k + v_t^k \\ &= G_{zm}^k z_{t-1}^m + G_{ym}^k y_{t-1}^m + g_m^k \omega_{t-1}^m + e_t^k + v_t^k \end{aligned}$$

where

$$\begin{aligned} G_{zm}^k &= \gamma^k (\alpha_z^k + (1-\alpha_z^k) r_{zz}^m) \\ G_{ym}^k &= \gamma^k (1-\alpha_z^k) r_{zy}^m \\ g_m^k &= \gamma^k (1-\alpha_z^k) \end{aligned}$$

and (10) as

$$\begin{aligned} y_t^k &= \beta^k \left(\alpha_y^k y_{t-1}^m + (1-\alpha_y^k) y_{t-1}^f \right) + z_t^k + x_t^k + u_t^k \\ &= \beta^k \left(\alpha_y^k y_{t-1}^m + (1-\alpha_y^k) \left(r_{yz}^m z_{t-1}^m + r_{yy}^m y_{t-1}^m + \varepsilon_{t-1}^m \right) \right) + z_t^k + x_t^k + u_t^k \\ &= \beta^k \left(\alpha_y^k y_{t-1}^m + (1-\alpha_y^k) \left(r_{yz}^m z_{t-1}^m + r_{yy}^m y_{t-1}^m + \varepsilon_{t-1}^m \right) \right) \\ &\quad + G_{zm}^k z_{t-1}^m + G_{ym}^k y_{t-1}^m + g_m^k \omega_{t-1}^m + e_t^k + v_t^k + x_t^k + u_t^k \end{aligned}$$

$$y_t^k = B_{ym}^k y_{t-1}^m + B_{zm}^k z_{t-1}^m + b_m^k \varepsilon_{t-1}^m + g_m^k \omega_{t-1}^m + e_t^k + v_t^k + x_t^k + u_t^k$$

where

$$\begin{aligned}
B^k_{ym} &= \beta^k\left(\alpha^k_y + (1-\alpha^k_y)r^m_{yy}\right) + G^k_{ym}\\
B^k_{zm} &= \beta^k(1-\alpha^k_y)r^m_{yz} + G^k_{zm}\\
b^k_m &= \beta^k(1-\alpha^k_y)
\end{aligned}$$

All these expressions will be used to compute correlations between relatives that are related through their fathers. However, when we consider relatives that are related through their mothers, we need to find expressions for y^k_t and z^k_t as functions of mother's years of schooling, y^f_{t-1}, and social status z^f_{t-1}.We then also consider the linear projections of z^m_{t-1} and y^m_{t-1} on z^f_{t-1} and y^f_{t-1}:

$$\begin{pmatrix} z^m_{t-1} \\ y^m_{t-1} \end{pmatrix} = \begin{pmatrix} r^f_{zz} & r^f_{zy} \\ r^f_{yz} & r^f_{yy} \end{pmatrix} \begin{pmatrix} z^f_{t-1} \\ y^f_{t-1} \end{pmatrix} + \begin{pmatrix} w^f_{t-1} \\ \varepsilon^f_{t-1} \end{pmatrix}$$

where w^f_{t-1} and ε^f_{t-1} might be correlated but are uncorrelated with z^f_{t-1} and y^f_{t-1}.

The coefficients of the linear projections depend on $\rho_{z^f y^f}, \rho_{z^m z^f}$, $\rho_{z^m y^f}$, $\rho_{y^m z^f}$ and $\rho_{y^m y^f}$, as well as on the standard deviations of z^k_{t-1} and y^k_{t-1}, $k = m, f$:

$$r^f_{zz} = \frac{1}{(1-\rho^2_{z^f y^f})}\frac{\sigma_{z^m}}{\sigma_{z^f}}(\rho_{z^m z^f} - \rho_{z^f y^f}\rho_{z^m y^f})$$

$$r^f_{zy} = \frac{1}{(1-\rho^2_{z^f y^f})}\frac{\sigma_{z^m}}{\sigma_{y^f}}(\rho_{z^m y^f} - \rho_{z^f y^f}\rho_{z^m z^f})$$

$$r^f_{yz} = \frac{1}{(1-\rho^2_{z^f y^f})}\frac{\sigma_{y^m}}{\sigma_{z^f}}(\rho_{y^m z^f} - \rho_{z^f y^f}\rho_{y^m y^f})$$

$$r^f_{yy} = \frac{1}{(1-\rho^2_{z^f y^f})}\frac{\sigma_{y^m}}{\sigma_{y^f}}(\rho_{y^m y^f} - \rho_{z^f y^f}\rho_{y^m z^f})$$

Using these linear projections, we can write (11) as

$$z^k_t = G^k_{zf} z^f_{t-1} + G^k_{yf} y^f_{t-1} + g^k_f \omega^f_{t-1} + e^k_t + v^k_t$$

where

$$\begin{aligned}
G^k_{zf} &= \gamma^k(\alpha^k_z r^f_{zz} + (1-\alpha^k_z))\\
G^k_{yf} &= \gamma^k \alpha^k_z r^f_{zy}\\
g^k_f &= \gamma^k \alpha^k_z
\end{aligned}$$

and (10) as

$$y_t^k = B_{yf}^k y_{t-1}^f + B_{zf}^k z_{t-1}^f + b_f^k \varepsilon_{t-1}^f + g_f^k \omega_{t-1}^f + e_t^k + v_t^k + x_t^k + u_t^k$$

where

$$\begin{aligned} B_{yf}^k &= \beta^k \left(\alpha_y^k r_{yy}^f + (1-\alpha_y^k) \right) + G_{yf}^k \\ B_{zf}^k &= \beta^k \alpha_y^k r_{yz}^f + G_{zf}^k \\ b_f^k &= \beta^k \alpha_y^k \end{aligned}$$

A.2 Steady state assumption

We assume that the second order moments of all variables are time invariant. This steady state assumption implies that $\rho_{z^m y^m}$ and $\rho_{z^f y^f}$ depend on the remaining parameters of the model as shown below.

$$Cov(y_t^m, z_t^m) = Cov(\beta^m \widetilde{y}_{t-1}^m + z_t^m, z_t^m) = \beta^m Cov(\widetilde{y}_{t-1}^m, z_t^m) + \sigma_{z^m}^2$$

$$\begin{aligned} Cov(\widetilde{y}_{t-1}^m, z_t^m) &= Cov(\alpha_y^m y_{t-1}^m + (1-\alpha_y^m) y_{t-1}^f, G_{zm}^m z_{t-1}^m + G_{ym}^m y_{t-1}^m + g_m^m \omega_{t-1}^m) \\ &= Cov(\alpha_y^m y_{t-1}^m + (1-\alpha_y^m) \left(r_{yz}^m z_{t-1}^m + r_{yy}^m y_{t-1}^m + \varepsilon_{t-1}^m \right), G_{zm}^m z_{t-1}^m + G_{ym}^m y_{t-1}^m + g_m^m \omega_{t-1}^m) \\ &= Cov(\left[\alpha_y^m + (1-\alpha_y^m) r_{yy}^m \right] y_{t-1}^m + (1-\alpha_y^m) r_{yz}^m z_{t-1}^m + (1-\alpha_y^m) \varepsilon_{t-1}^m, G_{zm}^m z_{t-1}^m + G_{ym}^m y_{t-1}^m + g_m^m \omega_{t-1}^m) \\ &= \left(\alpha_y^m + (1-\alpha_y^m) r_{yy}^m \right) G_{ym}^m \sigma_{y^m}^2 \\ &\quad + \left(\left[\alpha_y^m + (1-\alpha_y^m) r_{yy}^m \right] G_{zm}^m + (1-\alpha_y^m) r_{yz}^m G_{ym}^m \right) Cov(y_{t-1}^m, z_{t-1}^m) \\ &\quad + (1-\alpha_y^m) r_{yz}^m G_{zm}^m \sigma_{z^m}^2 + (1-\alpha_y^m) g_m^m Cov(\varepsilon_{t-1}^m, w_{t-1}^m) \end{aligned}$$

and

$$\begin{aligned} Cov(y_t^m, z_t^m) &= \beta^m \left(\alpha_y^m + (1-\alpha_y^m) r_{yy}^m \right) \gamma^m (1-\alpha_z^m) r_{zy}^m \sigma_{y^m}^2 \\ &\quad + \beta^m \left(\left[\alpha_y^m + (1-\alpha_y^m) r_{yy}^m \right] \gamma^m (\alpha_z^m + (1-\alpha_z^m) r_{zz}^m) + (1-\alpha_y^m) r_{yz}^m \gamma^m (1-\alpha_z^m) r_{zy}^m \right) Cov(y_{t-1}^m, z_{t-1}^m) \\ &\quad + (1 + \beta^m (1-\alpha_y^m) r_{yz}^m \gamma^m (\alpha_z^m + (1-\alpha_z^m) r_{zz}^m)) \sigma_{z^m}^2 + \beta^m (1-\alpha_y^m) \gamma^m (1-\alpha_z^m) Cov(\varepsilon_{t-1}^m, w_{t-1}^m) \end{aligned}$$

and substituting

$$Cov(\varepsilon_{t-1}^m, w_{t-1}^m) = Cov(y_{t-1}^f, z_{t-1}^f) - r_{yz}^m r_{zz}^m \sigma_{z^m}^2 - (r_{yz}^m r_{zy}^m + r_{zz}^m r_{yy}^m) Cov(y_{t-1}^m, z_{t-1}^m) - r_{yy}^m r_{zy}^m \sigma_{y^m}^2$$

we have

$$\begin{aligned}
Cov(y^m_t, z^m_t) \;=\;& \beta^m\alpha^m_y\gamma^m(1-\alpha^m_z)r^m_{zy}\sigma^2_{y^m}+\beta^m(1-\alpha^m_y)r^m_{yy}\gamma^m(1-\alpha^m_z)r^m_{zy}\sigma^2_{y^m}\\
&+\sigma^2_{z^m}+\beta^m(1-\alpha^m_y)r^m_{yz}\gamma^m\alpha^m_z\sigma^2_{z^m}+\beta^m(1-\alpha^m_y)r^m_{yz}\gamma^m(1-\alpha^m_z)r^m_{zz})\sigma^2_{z^m}\\
&+\beta^m\left(\left[\alpha^m_y+(1-\alpha^m_y)r^m_{yy}\right]\gamma^m(\alpha^m_z+(1-\alpha^m_z)r^m_{zz})+(1-\alpha^m_y)r^m_{yz}\gamma^m(1-\alpha^m_z)r^m_{zy}\right)Cov(y^m_{t-1},z^m_{t-1})\\
&-\beta^m(1-\alpha^m_y)\gamma^m(1-\alpha^m_z)r^m_{yz}r^m_{zz}\sigma^2_{z^m}-\beta^m(1-\alpha^m_y)\gamma^m(1-\alpha^m_z)r^m_{yy}r^m_{zy}\sigma^2_{y^m}\\
&-\beta^m(1-\alpha^m_y)\gamma^m(1-\alpha^m_z)r^m_{yz}r^m_{zy}Cov(y^m_{t-1},z^m_{t-1})-\beta^m(1-\alpha^m_y)\gamma^m(1-\alpha^m_z)r^m_{zz}r^m_{yy}Cov(y^m_{t-1},z^m_{t-1})\\
&\beta^m(1-\alpha^m_y)\gamma^m(1-\alpha^m_z)Cov(y^f_{t-1},z^f_{t-1})
\end{aligned}$$

$$\begin{aligned}
Cov(y^m_t, z^m_t) \;=\;& \beta^m\alpha^m_y\gamma^m(1-\alpha^m_z)r^m_{zy}\sigma^2_{y^m}\\
&+\sigma^2_{z^m}+\beta^m(1-\alpha^m_y)r^m_{yz}\gamma^m\alpha^m_z\sigma^2_{z^m}\\
&+\beta^m\left[\alpha^m_y\gamma^m\alpha^m_z+\alpha^m_y\gamma^m(1-\alpha^m_z)r^m_{zz}+(1-\alpha^m_y)\gamma^m\alpha^m_z r^m_{yy}\right]Cov(y^m_{t-1},z^m_{t-1})\\
&\beta^m(1-\alpha^m_y)\gamma^m(1-\alpha^m_z)Cov(y^f_{t-1},z^f_{t-1})
\end{aligned}$$

and substituting r^m_{zz}, r^m_{zy}, r^m_{yz} and r^m_{yy}

$$\begin{aligned}
Cov(y^m_t, z^m_t) \;=\;& \sigma^2_{z^m}+\beta^m\alpha^m_y\gamma^m\alpha^m_z Cov(y^m_{t-1},z^m_{t-1})+\beta^m(1-\alpha^m_y)\gamma^m(1-\alpha^m_z)Cov(y^f_{t-1},z^f_{t-1})\\
&+\beta^m\alpha^m_y\gamma^m(1-\alpha^m_z)\frac{1}{(1-\rho^2_{z^my^m})}(\rho_{y^mz^f}-\rho_{z^my^m}\rho_{z^mz^f})\sigma_{y^m}\sigma_{z^f}\\
&+\beta^m(1-\alpha^m_y)\gamma^m\alpha^m_z\frac{1}{(1-\rho^2_{z^my^m})}(\rho_{z^my^f}-\rho_{z^my^m}\rho_{y^my^f})\sigma_{z^m}\sigma_{y^f}\\
&+\beta^m\alpha^m_y\gamma^m(1-\alpha^m_z)\frac{1}{(1-\rho^2_{z^my^m})}\frac{\sigma_{z^f}}{\sigma_{z^m}}(\rho_{z^mz^f}-\rho_{z^my^m}\rho_{y^mz^f})Cov(y^m_{t-1},z^m_{t-1})\\
&+\beta^m(1-\alpha^m_y)\gamma^m\alpha^m_z\frac{1}{(1-\rho^2_{z^my^m})}\frac{\sigma_{y^f}}{\sigma_{y^m}}(\rho_{y^my^f}-\rho_{z^my^m}\rho_{z^my^f})Cov(y^m_{t-1},z^m_{t-1})
\end{aligned}$$

$$\begin{aligned}
\rho_{z^my^m} \;=\;& \frac{\sigma_{z^m}}{\sigma_{y^m}}+\beta^m\alpha^m_y\gamma^m\alpha^m_z\rho_{z^my^m}+\beta^m(1-\alpha^m_y)\gamma^m(1-\alpha^m_z)\rho_{z^fy^f}\frac{\sigma_{z^f}}{\sigma_{z^m}}\frac{\sigma_{y^f}}{\sigma_{y^m}}\\
&+\beta^m\alpha^m_y\gamma^m(1-\alpha^m_z)\rho_{y^mz^f}\frac{\sigma_{z^f}}{\sigma_{z^m}}\\
&+\beta^m(1-\alpha^m_y)\gamma^m\alpha^m_z\rho_{z^my^f}\frac{\sigma_{y^f}}{\sigma_{y^m}}
\end{aligned}$$

and

$$\begin{aligned} & (1-\beta^m\alpha_y^m\gamma^m\alpha_z^m)\rho_{z^my^m} - \beta^m(1-\alpha_y^m)\gamma^m(1-\alpha_z^m)\frac{\sigma_{z^f}}{\sigma_{z^m}}\frac{\sigma_{y^f}}{\sigma_{y^m}}\rho_{z^fy^f} \qquad (12) \\ = \; & \frac{\sigma_{z^m}}{\sigma_{y^m}} + \beta^m\alpha_y^m\gamma^m(1-\alpha_z^m)\rho_{y^mz^f}\frac{\sigma_{z^f}}{\sigma_{z^m}} + \beta^m(1-\alpha_y^m)\gamma^m\alpha_z^m\rho_{z^my^f}\frac{\sigma_{y^f}}{\sigma_{y^m}} \end{aligned}$$

analogously

$$\begin{aligned} & -\beta^f(1-\alpha_y^f)\gamma^f(1-\alpha_z^f)\frac{\sigma_{z^m}}{\sigma_{z^f}}\frac{\sigma_{y^m}}{\sigma_{y^f}}\rho_{z^my^m} + (1-\beta^f\alpha_y^f\gamma^f\alpha_z^f)\rho_{z^fy^f} \qquad (13) \\ = \; & \frac{\sigma_{z^f}}{\sigma_{y^f}} + \beta^f\alpha_y^f\gamma^f(1-\alpha_z^f)\rho_{z^my^f}\frac{\sigma_{z^m}}{\sigma_{z^f}} + \beta^f(1-\alpha_y^f)\gamma^f\alpha_z^f\rho_{y^mz^f}\frac{\sigma_{y^m}}{\sigma_{y^f}} \end{aligned}$$

In matrix form

$$\begin{aligned} & \begin{bmatrix} (1-\beta^m\alpha_y^m\gamma^m\alpha_z^m) & -\beta^m(1-\alpha_y^m)\gamma^m(1-\alpha_z^m)\frac{\sigma_{z^f}}{\sigma_{z^m}}\frac{\sigma_{y^f}}{\sigma_{y^m}} \\ -\beta^f(1-\alpha_y^f)\gamma^f(1-\alpha_z^f)\frac{\sigma_{z^m}}{\sigma_{z^f}}\frac{\sigma_{y^m}}{\sigma_{y^f}} & (1-\beta^f\alpha_y^f\gamma^f\alpha_z^f) \end{bmatrix} \begin{pmatrix} \rho_{z^my^m} \\ \rho_{z^fy^f} \end{pmatrix} \\ = \; & \begin{pmatrix} \frac{\sigma_{z^m}}{\sigma_{y^m}} + \beta^m\alpha_y^m\gamma^m(1-\alpha_z^m)\rho_{y^mz^f}\frac{\sigma_{z^f}}{\sigma_{z^m}} + \beta^m(1-\alpha_y^m)\gamma^m\alpha_z^m\rho_{z^my^f}\frac{\sigma_{y^f}}{\sigma_{y^m}} \\ \frac{\sigma_{z^f}}{\sigma_{y^f}} + \beta^f\alpha_y^f\gamma^f(1-\alpha_z^f)\rho_{z^my^f}\frac{\sigma_{z^m}}{\sigma_{z^f}} + \beta^f(1-\alpha_y^f)\gamma^f\alpha_z^f\rho_{y^mz^f}\frac{\sigma_{y^m}}{\sigma_{y^f}} \end{pmatrix} \end{aligned}$$

and $\rho_{z^my^m}$ and $\rho_{z^fy^f}$ depend on $\rho_{y^mz^f}$, $\rho_{z^my^f}$ and some other parameters of the model.

We then have that the model has 20 parameters $\gamma^k, \beta^k, \alpha_z^k, \alpha_y^k, \sigma_{z^k}^2, \sigma_{x^k}^2, \sigma_{e^k}^2, k = m, f, \sigma_{x^mx^f}, \sigma_{e^me^f}, \rho_{z^mz^f}, \rho_{y^mz^f}, \rho_{z^my^f}$ and $\rho_{y^my^f}$

A.3 Covariances

A.3.1 Main covariances

We first compute the main covariances (husband-wife, parent-child and siblings). Then, the covariances for other relatives are obtained recursively.

Husband and wife

$$\begin{aligned} Cov(y_{t-1}^m, y_{t-1}^f) &= Cov(y_{t-1}^m, r_{yz}^m z_{t-1}^m + r_{yy}^m y_{t-1}^m + \varepsilon_{t-1}^m) = r_{yz}^m Cov(y_{t-1}^m, z_{t-1}^m) + r_{yy}^m \sigma_{y^m}^2 \\ &= \rho_{z^my^m}\sigma_{z^m}\sigma_{y^m} + r_{yy}^m\sigma_{y^m}^2 \end{aligned}$$

Parent–child

Let $k = m, f$ be the gender of the child and $n = m, f$ the gender of the parent

$$\begin{aligned} Cov(z_t^k, z_{t-1}^n) &= Cov(G_{yn}^k y_{t-1}^n + G_{zn}^k z_{t-1}^n, z_{t-1}^n) \\ &= G_{yn}^k Cov(y_{t-1}^n, z_{t-1}^n) + G_{zn}^k \sigma_{z^n}^2 \end{aligned}$$

$$\begin{aligned} Cov(z_t^k, y_{t-1}^n) &= Cov(G_{yn}^k y_{t-1}^n + G_{zn}^k z_{t-1}^n, y_{t-1}^n) \\ &= G_{yn}^k \sigma_{y^n}^2 + G_{zn}^k Cov(y_{t-1}^n, z_{t-1}^n) \end{aligned}$$

$$\begin{aligned} Cov(y_t^k, z_{t-1}^n) &= Cov(B_{yn}^k y_{t-1}^n + B_{zn}^k z_{t-1}^n, z_{t-1}^n) \\ &= B_{yn}^k Cov(y_{t-1}^n, z_{t-1}^n) + B_{zn}^k \sigma_{z^n}^2 \end{aligned}$$

$$\begin{aligned} Cov(y_t^k, y_{t-1}^n) &= Cov(B_{yn}^k y_{t-1}^n + B_{zn}^k z_{t-1}^n, y_{t-1}^n) \\ &= B_{yn}^k \sigma_{y^n}^2 + B_{zn}^k Cov(y_{t-1}^n, z_{t-1}^n) \end{aligned}$$

Siblings

We denote by $k, l = m, f$ the genders of the siblings. We can compute the covariances projecting on the father ($n = m$) or on the mother ($n = f$)

$$\begin{aligned} &Cov(z_t^{k,i}, z_t^{l,j}) = Cov(G_{yn}^k y_{t-1}^n + G_{zn}^k z_{t-1}^n + g_n^k \omega_{t-1}^n + e_t^{k,i}, G_{yn}^l y_{t-1}^n + G_{zn}^l z_{t-1}^n + g_n^l \omega_{t-1}^n + e_{t-1}^{l,j}) \\ &= G_{yn}^k G_{yn}^l \sigma_{y^n}^2 + G_{zn}^k G_{zn}^l \sigma_{z^n}^2 + \left(G_{yn}^k G_{zn}^l + G_{zn}^k G_{yn}^l \right) Cov(y_{t-1}^n, z_{t-1}^n) \\ &+ g_n^k g_n^l \sigma_{w^n}^2 + \sigma_{e^k e^l} \end{aligned}$$

$$\begin{aligned} &Cov(z_t^{k,i}, y_t^{l,j}) = Cov(G_{yn}^k y_{t-1}^n + G_{zn}^k z_{t-1}^n + g_n^k \omega_{t-1}^n + e_{t-1}^{k,i}, B_{yn}^l y_{t-1}^n + B_{zn}^l z_{t-1}^n + b_n^l \varepsilon_{t-1}^n + g_n^l \omega_{t-1}^n + e_{t-1}^{l,j}) \\ &= G_{yn}^k B_{yn}^l \sigma_{y^n}^2 + G_{zn}^k B_{zn}^l \sigma_{z^n}^2 + \left(G_{yn}^k B_{zn}^l + G_{zn}^k B_{yn}^l \right) Cov(y_{t-1}^n, z_{t-1}^n) \\ &+ g_n^k g_n^l \sigma_{w^n}^2 + \sigma_{e^k e^l} + g_n^k b_n^l Cov(\varepsilon_{t-1}^n, \omega_{t-1}^n) \end{aligned}$$

$$\begin{aligned} &Cov(y_t^{k,i}, z_t^{l,j}) = Cov(B_{yn}^k y_{t-1}^n + B_{zn}^k z_{t-1}^n + b_n^k \varepsilon_{t-1}^n + g_n^k \omega_{t-1}^n + e_{t-1}^{k,i} + x_{t-1}^k, G_{yn}^l y_{t-1}^n + G_{zn}^l z_{t-1}^n + g_n^l \omega_{t-1}^n + e_{t-1}^{l,j}) \\ &= B_{yn}^k G_{yn}^l \sigma_{y^n}^2 + B_{zn}^k G_{zn}^l \sigma_{z^n}^2 + \left(B_{yn}^k G_{zn}^l + B_{zn}^k G_{yn}^l \right) Cov(y_{t-1}^n, z_{t-1}^n) \\ &+ g_n^k g_n^l \sigma_{w^n}^2 + \sigma_{e^k e^l} + b_n^k g_n^l Cov(\varepsilon_{t-1}^n, \omega_{t-1}^n) \end{aligned}$$

$$\begin{aligned} &Cov(y_t^{k,i}, y_t^{l,j}) = Cov(B_{yn}^k y_{t-1}^n + B_{zn}^k z_{t-1}^n + b_n^k \varepsilon_{t-1}^n + g_n^k \omega_{t-1}^n + e_{t-1}^{k,i} + x_{t-1}^k, B_{yn}^l y_{t-1}^n + B_{zn}^l z_{t-1}^n + b_n^l \varepsilon_{t-1}^n + g_n^l \omega_{t-1}^n + e_{t-1}^{l,j} + x_{t-1}^l) \\ &= B_{yn}^k B_{yn}^l \sigma_{y^n}^2 + B_{zn}^k B_{zn}^l \sigma_{z^n}^2 + \left(B_{yn}^k B_{zn}^l + B_{zn}^k B_{yn}^l \right) Cov(y_{t-1}^n, z_{t-1}^n) + b_n^k b_n^l \sigma_{\varepsilon^n}^2 \\ &+ g_n^k g_n^l \sigma_{w^n}^2 + \sigma_{e^k e^l} + \left(b_n^k g_n^l + g_n^k b_n^l \right) Cov(\varepsilon_{t-1}^n, \omega_{t-1}^n) + Cov(x_{t-1}^k, x_{t-1}^l) \end{aligned}$$

A.3.2 Other covariances

Before we obtain the remaining covariances for different degrees of kinship we compute the linear projections of $z_{t-1}^{i,k}$ and $y_{t-1}^{i,k}$ on $z_{t-1}^{j,l}$ and $y_{t-1}^{j,l}$, $k, l = m, f$, where i and j are siblings

$$\begin{pmatrix} z_{t-1}^{i,k} \\ y_{t-1}^{i,k} \end{pmatrix} = \begin{pmatrix} r_{zz}^{k,l} & r_{zy}^{k,l} \\ r_{yz}^{k,l} & r_{yy}^{k,l} \end{pmatrix} \begin{pmatrix} z_{t-1}^{j,l} \\ y_{t-1}^{j,l} \end{pmatrix} + \begin{pmatrix} w_{t-1}^{k,l} \\ \varepsilon_{t-1}^{k,l} \end{pmatrix}$$

where $w_{t-1}^{k,l}$ and $\varepsilon_{t-1}^{k,l}$ might be correlated but are uncorrelated with $z_{t-1}^{j,l}$ and $y_{t-1}^{j,l}$. We have that

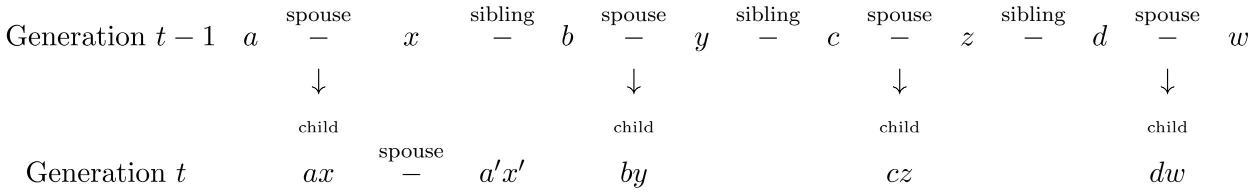
$$r_{zz}^{k,l} = \frac{1}{\sigma_{z^l}^2 \sigma_{y^l}^2 - \sigma_{z^l y^l}^2} \left(\sigma_{y^l}^2 \sigma_{z^{i,k} z^{j,l}} - \sigma_{z^l y^l} \sigma_{z^{i,k} y^{j,l}} \right)$$

$$r_{zy}^{k,l} = \frac{1}{\sigma_{z^l}^2 \sigma_{y^l}^2 - \sigma_{z^l y^l}^2} \left(\sigma_{z^l}^2 \sigma_{z^{i,k} y^{j,l}} - \sigma_{z^l y^l} \sigma_{z^{i,k} z^{j,l}} \right)$$

$$r_{yz}^{k,l} = \frac{1}{\sigma_{z^l}^2 \sigma_{y^l}^2 - \sigma_{z^l y^l}^2} \left(\sigma_{y^l}^2 \sigma_{y^{i,k} z^{j,l}} - \sigma_{z^l y^l} \sigma_{y^{i,k} y^{j,l}} \right)$$

$$r_{yy}^{k,l} = \frac{1}{\sigma_{z^l}^2 \sigma_{y^l}^2 - \sigma_{z^l y^l}^2} \left(\sigma_{z^l}^2 \sigma_{y^{i,k} y^{j,l}} - \sigma_{z^l y^l} \sigma_{y^{i,k} z^{j,l}} \right)$$

Let's now consider several couples "a" and "x", "b" and "y", "c" and "z", "d" and "w", etc., such that "x" and "b" are siblings, "y" and "c" are siblings, etc. Let "ax" be a child of "a" and "x", "by" a child of "b" and "y", etc., and "$a'x'$" the spouse of "ax".



Consanguine relatives ("blood")

Vertical covariances

Uncle/aunt (siblings of the parents)

We have to compute the covariances between "ax" and "b". Let $a^* = m, f$ be the gender of "ax" and $l = m, f$ the gender of the "b". We project ax on x (his/her father or mother) who has gender n'

$$Cov(z_t^{ax,n^*}, z_{t-1}^{b,l}) = Cov(G_{zn'}^{n^*} z_{t-1}^{x,n'} + G_{yn'}^{n^*} y_{t-1}^{x,n'}, z_{t-1}^{b,l}) = G_{zn'}^{n^*} Cov(z_{t-1}^{x,n'}, z_{t-1}^{b,l}) + G_{yn'}^{n^*} Cov(y_{t-1}^{x,n'}, z_{t-1}^{b,l})$$

$$Cov(z_t^{ax,n^*}, y_{t-1}^{b,l}) = Cov(G^{n^*}_{zn'} z^{x,n'}_{t-1} + G^{n^*}_{yn'} y^{x,n'}_{t-1}, y^{b,l}_{t-1}) = G^{n^*}_{zn'} Cov(z^{x,n'}_{t-1}, y^{b,l}_{t-1}) + G^{n^*}_{yn'} Cov(y^{x,n'}_{t-1}, y^{b,l}_{t-1})$$

$$Cov(y_t^{ax,n^*}, z_{t-1}^{b,l}) = Cov(B^{n^*}_{zn'} z^{x,n'}_{t-1} + B^{n^*}_{yn'} y^{x,n'}_{t-1}, z^{b,l}_{t-1}) = B^{n^*}_{zn'} Cov(z^{x,n'}_{t-1}, z^{b,l}_{t-1}) + B^{n^*}_{yn'} Cov(y^{x,n'}_{t-1}, z^{b,l}_{t-1})$$

$$Cov(y_t^{ax,n^*}, y_{t-1}^{b,l}) = Cov(B^{n^*}_{zn'} z^{x,n'}_{t-1} + B^{n^*}_{yn'} y^{x,n'}_{t-1}, y^{b,l}_{t-1}) = B^{n^*}_{zn'} Cov(z^{x,n'}_{t-1}, y^{b,l}_{t-1}) + B^{n^*}_{yn'} Cov(y^{x,n'}_{t-1}, y^{b,l}_{t-1})$$

where ”x” and ”b” are siblings.

Horizontal covariances

Cousins

We have to compute the covariances between ”ax” and ”by”. Let $a^* = m, f$ be the gender of ”ax” and $l^* = m, f$ the gender of the ”ay”. We project by on b (his/her father or mother) who has gender l

$$Cov(z_t^{ax,n^*}, z_t^{by,l^*}) = Cov(z_t^{ax,n^*}, G^{l^*}_{zl} z^{b,l}_{t-1} + G^{l^*}_{yl} y^{b,l}_{t-1}) = G^{l^*}_{zl} Cov(z_t^{ax,n^*}, z^{b,l}_{t-1}) + G^{l^*}_{yl} Cov(z_t^{ax,n^*}, y^{b,l}_{t-1})$$

$$Cov(z_t^{ax,n^*}, y_t^{by,l^*}) = Cov(z_t^{ax,n^*}, B^{l^*}_{zl} z^{b,l}_{t-1} + B^{l^*}_{yl} y^{b,l}_{t-1}) = B^{l^*}_{zl} Cov(z_t^{ax,n^*}, z^{b,l}_{t-1}) + B^{l^*}_{yl} Cov(z_t^{ax,n^*}, y^{b,l}_{t-1})$$

$$Cov(y_t^{ax,n^*}, z_t^{by,l^*}) = Cov(y_t^{ax,n^*}, G^{l^*}_{zl} z^{b,l}_{t-1} + G^{l^*}_{yl} y^{b,l}_{t-1}) = G^{l^*}_{zl} Cov(y_t^{ax,n^*}, z^{b,l}_{t-1}) + G^{l^*}_{yl} Cov(y_t^{ax,n^*}, y^{b,l}_{t-1})$$

$$Cov(y_t^{ax,n^*}, y_t^{by,l^*}) = Cov(y_t^{ax,n^*}, B^{l^*}_{zl} z^{b,l}_{t-1} + B^{l^*}_{yl} y^{b,l}_{t-1}) = B^{l^*}_{zl} Cov(y_t^{ax,n^*}, z^{b,l}_{t-1}) + B^{l^*}_{yl} Cov(y_t^{ax,n^*}, y^{b,l}_{t-1})$$

where ”b” is the uncle/aunt of ”ax”.

Affinity relatives (”in-law”)

Vertical covariances

Parents in-law

We have to compute the covariances between ”$a'x'$” and ”a”. Let $a^{*'} = m, f$ be the gender of ”$a'x'$” and $n = m, f$ the gender of the ”a”. We project ”$a'x'$” on his/her spouse ”ax”

$$Cov(z_t^{a'x',a^{*'}}, z^{a,n}_{t-1}) = Cov(r^{n^*}_{zz} z^{ax,n^*}_{t-1} + r^{n^*}_{zy} y^{ax,n^*}_{t-1}, z^{a,n}_{t-1}) = r^{n^*}_{zz} Cov(z_t^{ax,n^*}, z^{a,n}_{t-1}) + r^{n^*}_{zy} Cov(y^{ax,n^*}_{t-1}, z^{a,n}_{t-1})$$

$$Cov(z_t^{a'x',a^{*'}}, y^{a,n}_{t-1}) = Cov(r^{n^*}_{zz} z^{ax,n^*}_{t-1} + r^{n^*}_{zy} y^{ax,n^*}_{t-1}, y^{a,n}_{t-1}) = r^{n^*}_{zz} Cov(z_t^{ax,n^*}, y^{a,n}_{t-1}) + r^{n^*}_{zy} Cov(y^{ax,n^*}_{t-1}, y^{a,n}_{t-1})$$

$$Cov(y_t^{a'x',a^{*'}}, z^{a,n}_{t-1}) = Cov(r^{n^*}_{yz} z^{ax,n^*}_{t-1} + r^{n^*}_{yy} y^{ax,n^*}_{t-1}, z^{a,n}_{t-1}) = r^{n^*}_{yz} Cov(z_t^{ax,n^*}, z^{a,n}_{t-1}) + r^{n^*}_{yy} Cov(y^{ax,n^*}_{t-1}, z^{a,n}_{t-1})$$

$$Cov(y_t^{a'x',a^{*'}}, y^{a,n}_{t-1}) = Cov(r^{n^*}_{yz} z^{ax,n^*}_{t-1} + r^{n^*}_{yy} y^{ax,n^*}_{t-1}, y^{a,n}_{t-1}) = r^{n^*}_{yz} Cov(z_t^{ax,n^*}, y^{a,n}_{t-1}) + r^{n^*}_{yy} Cov(y^{ax,n^*}_{t-1}, y^{a,n}_{t-1})$$

where ”a” is the father/mother of ”ax”.

Spouse of the uncle/aunt (spouses of the siblings of the parents)

We have to compute the covariances between ”ax” and ”y”. Let $a^* = m, f$ be the gender of ”ax” and

$l' = m, f$ the gender of the "y". We project y on his/her spouse b

$$Cov(z_t^{ax,n^*}, z_{t-1}^{y,l'}) = Cov(z_t^{ax,n^*}, r_{zz}^{l} z_{t-1}^{b,l} + r_{zy}^{l} y_{t-1}^{b,l}) = r_{zz}^{l} Cov(z_t^{ax,n^*}, z_{t-1}^{b,l}) + r_{zy}^{l} Cov(z_t^{ax,n^*}, y_{t-1}^{b,l})$$

$$Cov(z_t^{ax,n^*}, y_{t-1}^{y,l'}) = Cov(z_t^{ax,n^*}, r_{yz}^{l} z_{t-1}^{b,l} + r_{yy}^{l} y_{t-1}^{b,l}) = r_{yz}^{l} Cov(z_t^{ax,n^*}, z_{t-1}^{b,l}) + r_{yy}^{l} Cov(z_t^{ax,n^*}, y_{t-1}^{b,l})$$

$$Cov(y_t^{ax,n^*}, z_{t-1}^{y,l'}) = Cov(y_t^{ax,n^*}, r_{zz}^{l} z_{t-1}^{b,l} + r_{zy}^{l} y_{t-1}^{b,l}) = r_{zz}^{l} Cov(y_t^{ax,n^*}, z_{t-1}^{b,l}) + r_{zy}^{l} Cov(y_t^{ax,n^*}, y_{t-1}^{b,l})$$

$$Cov(y_t^{ax,n^*}, y_{t-1}^{y,l'}) = Cov(y_t^{ax,n^*}, r_{yz}^{l} z_{t-1}^{b,l} + r_{yy}^{l} y_{t-1}^{b,l}) = r_{yz}^{l} Cov(y_t^{ax,n^*}, z_{t-1}^{b,l}) + r_{yy}^{l} Cov(y_t^{ax,n^*}, y_{t-1}^{b,l})$$

where "b" is uncle/aunt of "ax".

Uncles/Aunts of the spouse

We have to compute the covariances between "$a'x'$" and "b". Let $a^{*'} = m, f$ be the gender of "$a'x'$" and $l = m, f$ the gender of the "b". We project "$a'x'$" on his/her spouse "ax"

$$Cov(z_t^{a'x',a^{*'}}, z_{t-1}^{b,l}) = Cov(r_{zz}^{n^*} z_{t-1}^{ax,n^*} + r_{zy}^{n^*} y_{t-1}^{ax,n^*}, z_{t-1}^{b,l}) = r_{zz}^{n^*} Cov(z_t^{ax,n^*}, z_{t-1}^{b,l}) + r_{zy}^{n^*} Cov(y_{t-1}^{ax,n^*}, z_{t-1}^{b,l})$$

$$Cov(z_t^{a'x',a^{*'}}, y_{t-1}^{b,l}) = Cov(r_{zz}^{n^*} z_{t-1}^{ax,n^*} + r_{zy}^{n^*} y_{t-1}^{ax,n^*}, y_{t-1}^{b,l}) = r_{zz}^{n^*} Cov(z_t^{ax,n^*}, y_{t-1}^{b,l}) + r_{zy}^{n^*} Cov(y_{t-1}^{ax,n^*}, y_{t-1}^{b,l})$$

$$Cov(y_t^{a'x',a^{*'}}, z_{t-1}^{b,l}) = Cov(r_{yz}^{n^*} z_{t-1}^{ax,n^*} + r_{yy}^{n^*} y_{t-1}^{ax,n^*}, z_{t-1}^{b,l}) = r_{yz}^{n^*} Cov(z_t^{ax,n^*}, z_{t-1}^{b,l}) + r_{yy}^{n^*} Cov(y_{t-1}^{ax,n^*}, z_{t-1}^{b,l})$$

$$Cov(y_t^{a'x',a^{*'}}, y_{t-1}^{b,l}) = Cov(r_{yz}^{n^*} z_{t-1}^{ax,n^*} + r_{yy}^{n^*} y_{t-1}^{ax,n^*}, y_{t-1}^{b,l}) = r_{yz}^{n^*} Cov(z_t^{ax,n^*}, y_{t-1}^{a,n}) + r_{yy}^{n^*} Cov(y_{t-1}^{ax,n^*}, y_{t-1}^{b,l})$$

where "b" is the uncle/aunt of "ax".

Siblings of the siblings in law of the parents

We have to compute the covariances between "ax" and "c". Let $a^* = m, f$ be the gender of "ax" and $k = m, f$ the gender of the "c". We project c on his/her sibling y

$$Cov(z_t^{ax,n^*}, z_{t-1}^{c,k}) = Cov(z_t^{ax,n^*}, r_{zz}^{k,l'} z_{t-1}^{y,l'} + r_{zy}^{k,l'} y_{t-1}^{y,l'}) = r_{zz}^{k,l'} Cov(z_t^{ax,n^*}, z_{t-1}^{y,l'}) + r_{zy}^{k,l'} Cov(z_t^{ax,n^*}, y_{t-1}^{y,l'})$$

$$Cov(z_t^{ax,n^*}, y_{t-1}^{c,k}) = Cov(z_t^{ax,n^*}, r_{yz}^{k,l'} z_{t-1}^{y,l'} + r_{yy}^{k,l'} y_{t-1}^{y,l'}) = r_{yz}^{k,l'} Cov(z_t^{ax,n^*}, z_{t-1}^{y,l'}) + r_{yy}^{k,l'} Cov(z_t^{ax,n^*}, y_{t-1}^{y,l'})$$

$$Cov(y_t^{ax,n^*}, z_{t-1}^{c,k}) = Cov(y_t^{ax,n^*}, r_{zz}^{k,l'} z_{t-1}^{y,l'} + r_{zy}^{k,l'} y_{t-1}^{y,l'}) = r_{zz}^{k,l'} Cov(y_t^{ax,n^*}, z_{t-1}^{y,l'}) + r_{zy}^{k,l'} Cov(y_t^{ax,n^*}, y_{t-1}^{y,l'})$$

$$Cov(y_t^{ax,n^*}, y_{t-1}^{c,k}) = Cov(y_t^{ax,n^*}, r_{yz}^{k,l'} z_{t-1}^{y,l'} + r_{yy}^{k,l'} y_{t-1}^{y,l'}) = r_{yz}^{k,l'} Cov(y_t^{ax,n^*}, z_{t-1}^{y,l'}) + r_{yy}^{k,l'} Cov(y_t^{ax,n^*}, y_{t-1}^{y,l'})$$

where "y" is the spouse of the uncle/aunt of "ax". We can analogously compute the covariance between "ax" and "z", "d", etc.

Horizontal covariances

Siblings in law

We have to compute the covariances between ”a” and ”b”. Let $n = m, f$ be the gender of ”a” and $l = m, f$ the gender of the ”b”. We project a on his/her spouse x

$$Cov(z^{a,n}_{t-1}, z^{b,l}_{t-1}) = Cov(r^{n'}_{zz} z^{x,n'}_{t-1} + r^{n'}_{zy} y^{x,n'}_{t-1}, z^{b,l}_{t-1}) = r^{n'}_{zz} Cov(z^{x,n'}_{t-1}, z^{b,l}_{t-1}) + r^{n'}_{zy} Cov(y^{x,n'}_{t-1}, z^{b,l}_{t-1})$$

$$\begin{aligned}
Cov(z^{a,n}_{t-1}, y^{b,l}_{t-1}) &= Cov(r^{n'}_{zz} z^{x,n'}_{t-1} + r^{n'}_{zy} y^{x,n'}_{t-1}, y^{b,l}_{t-1}) = r^{n'}_{zz} Cov(z^{x,n'}_{t-1}, y^{b,l}_{t-1}) + r^{n'}_{zy} Cov(y^{x,n'}_{t-1}, y^{b,l}_{t-1}) \\
Cov(y^{a,n}_{t-1}, z^{b,l}_{t-1}) &= Cov(r^{n'}_{yz} z^{x,n'}_{t-1} + r^{n'}_{yy} y^{x,n'}_{t-1}, z^{b,l}_{t-1}) = r^{n'}_{yz} Cov(z^{x,n'}_{t-1}, z^{b,l}_{t-1}) + r^{n'}_{yy} Cov(y^{x,n'}_{t-1}, z^{b,l}_{t-1}) \\
Cov(y^{a,n}_{t-1}, y^{b,l}_{t-1}) &= Cov(r^{n'}_{yz} z^{x,n'}_{t-1} + r^{n'}_{yy} y^{x,n'}_{t-1}, y^{b,l}_{t-1}) = r^{n'}_{yz} Cov(z^{x,n'}_{t-1}, y^{b,l}_{t-1}) + r^{n'}_{yy} Cov(y^{x,n'}_{t-1}, y^{b,l}_{t-1})
\end{aligned}$$

where ”x” and ”b” are siblings. Notice that since ”x” is the spouse of ”a”, $n' = f$ when $n = m$ and viceversa.

Spouse of the siblings in law

We have to compute the covariances between ”a” and ”y”. Let $n = m, f$ be the gender of ”a” and $l' = m, f$ the gender of the ”y”. We project y on his/her spouse b

$$Cov(z^{a,n}_{t-1}, z^{y,l'}_{t-1}) = Cov(z^{a,n}_{t-1}, r^{l}_{zz} z^{b,l}_{t-1} + r^{l}_{zy} y^{b,l}_{t-1}) = r^{l}_{zz} Cov(z^{a,n}_{t-1}, z^{b,l}_{t-1}) + r^{l}_{zy} Cov(z^{a,n}_{t-1}, y^{b,l}_{t-1})$$

$$\begin{aligned}
Cov(z^{a,n}_{t-1}, y^{y,l'}_{t-1}) &= Cov(z^{a,n}_{t-1}, r^{l}_{yz} z^{b,l}_{t-1} + r^{l}_{yy} y^{b,l}_{t-1}) = r^{l}_{yz} Cov(z^{a,n}_{t-1}, z^{b,l}_{t-1}) + r^{l}_{yy} Cov(z^{a,n}_{t-1}, y^{b,l}_{t-1}) \\
Cov(y^{a,n}_{t-1}, z^{y,l'}_{t-1}) &= Cov(y^{a,n}_{t-1}, r^{l}_{zz} z^{b,l}_{t-1} + r^{l}_{zy} y^{b,l}_{t-1}) = r^{l}_{zz} Cov(y^{a,n}_{t-1}, z^{b,l}_{t-1}) + r^{l}_{zy} Cov(y^{a,n}_{t-1}, y^{b,l}_{t-1}) \\
Cov(y^{a,n}_{t-1}, y^{y,l'}_{t-1}) &= Cov(y^{a,n}_{t-1}, r^{l}_{yz} z^{b,l}_{t-1} + r^{l}_{yy} y^{b,l}_{t-1}) = r^{l}_{yz} Cov(y^{a,n}_{t-1}, z^{b,l}_{t-1}) + r^{l}_{yy} Cov(y^{a,n}_{t-1}, y^{b,l}_{t-1})
\end{aligned}$$

where ”a” and ”b” are siblings in law. Notice that since ”b” is the spouse of ”y”, $l = f$ when $l' = m$ and viceversa.

Sibling of the sibling in law

We have to compute the covariances between ”x” and ”c”. Let $n' = m, f$ be the gender of ”x” and $k = m, f$ the gender of the ”c”. We project x on his/her sibling b who has gender l

$$Cov(z^{x,n'}_{t-1}, z^{c,k}_{t-1}) = Cov(r^{n',l}_{zz} z^{b,l}_{t-1} + r^{n',l}_{zy} y^{b,l}_{t-1}, z^{c,k}_{t-1}) = r^{n',l}_{zz} Cov(z^{b,l}_{t-1}, z^{c,k}_{t-1}) + r^{n',l}_{zy} Cov(y^{b,l}_{t-1}, z^{c,k}_{t-1})$$

$$\begin{aligned}
Cov(z^{x,n'}_{t-1}, y^{c,k}_{t-1}) &= Cov(r^{n',l}_{zz} z^{b,l}_{t-1} + r^{n',l}_{zy} y^{b,l}_{t-1}, y^{c,k}_{t-1}) = r^{n',l}_{zz} Cov(z^{b,l}_{t-1}, y^{c,k}_{t-1}) + r^{n',l}_{zy} Cov(y^{b,l}_{t-1}, y^{c,k}_{t-1}) \\
Cov(y^{x,n'}_{t-1}, z^{c,k}_{t-1}) &= Cov(r^{n',l}_{yz} z^{b,l}_{t-1} + r^{n',l}_{yy} y^{b,l}_{t-1}, z^{c,k}_{t-1}) = r^{n',l}_{yz} Cov(z^{b,l}_{t-1}, z^{c,k}_{t-1}) + r^{n',l}_{yy} Cov(y^{b,l}_{t-1}, z^{c,k}_{t-1}) \\
Cov(y^{x,n'}_{t-1}, y^{c,k}_{t-1}) &= Cov(r^{n',l}_{yz} z^{b,l}_{t-1} + r^{n',l}_{yy} y^{b,l}_{t-1}, y^{c,k}_{t-1}) = r^{n',l}_{yz} Cov(z^{b,l}_{t-1}, y^{c,k}_{t-1}) + r^{n',l}_{yy} Cov(y^{b,l}_{t-1}, y^{c,k}_{t-1})
\end{aligned}$$

where ”b” and ”c” are siblings in law.

Sibling in law of degree 2

We have to compute the covariances between ”a” and ”c”. Let $n = m, f$ be the gender of ”a” and $k = m, f$

the gender of the ”b”. We project a on his/her spouse x

$$Cov(z_{t-1}^{a,n}, z_{t-1}^{c,l}) = Cov(r_{zz}^{n'} z_{t-1}^{x,n'} + r_{zy}^{n'} y_{t-1}^{x,n'}, z_{t-1}^{c,l}) = r_{zz}^{n'} Cov(z_{t-1}^{x,n'}, z_{t-1}^{c,l}) + r_{zy}^{n'} Cov(y_{t-1}^{x,n'}, z_{t-1}^{c,l})$$

$$Cov(z_{t-1}^{a,n}, y_{t-1}^{c,l}) = Cov(r_{zz}^{n'} z_{t-1}^{x,n'} + r_{zy}^{n'} y_{t-1}^{x,n'}, y_{t-1}^{c,l}) = r_{zz}^{n'} Cov(z_{t-1}^{x,n'}, y_{t-1}^{c,l}) + r_{zy}^{n'} Cov(y_{t-1}^{x,n'}, y_{t-1}^{c,l})$$

$$Cov(y_{t-1}^{a,n}, z_{t-1}^{c,l}) = Cov(r_{yz}^{n'} z_{t-1}^{x,n'} + r_{yy}^{n'} y_{t-1}^{x,n'}, z_{t-1}^{c,l}) = r_{yz}^{n'} Cov(z_{t-1}^{x,n'}, z_{t-1}^{c,l}) + r_{yy}^{n'} Cov(y_{t-1}^{x,n'}, z_{t-1}^{c,l})$$

$$Cov(y_{t-1}^{a,n}, y_{t-1}^{c,l}) = Cov(r_{yz}^{n'} z_{t-1}^{x,n'} + r_{yy}^{n'} y_{t-1}^{x,n'}, y_{t-1}^{c,l}) = r_{yz}^{n'} Cov(z_{t-1}^{x,n'}, y_{t-1}^{c,l}) + r_{yy}^{n'} Cov(y_{t-1}^{x,n'}, y_{t-1}^{c,l})$$

where ”x” is the sibling of the sibling in law of ”b”. Notice that since ”x” is the spouse of ”a”, $n' = f$ when $n = m$ and viceversa. We can compute siblings in law of any degree analogously.

Spouse of the sibling in law of degree 2

We have to compute the covariances between ”a” and ”z”. Let $n = m, f$ be the gender of ”a” and $k' = m, f$ the gender of the ”z”. We project z on his/her spouse c

$$Cov(z_{t-1}^{a,n}, z_{t-1}^{z,k'}) = Cov(z_{t-1}^{a,n}, r_{zz}^{k} z_{t-1}^{c,k} + r_{zy}^{k} y_{t-1}^{c,k}) = r_{zz}^{k} Cov(z_{t-1}^{a,n}, z_{t-1}^{c,k}) + r_{zy}^{k} Cov(z_{t-1}^{a,n}, y_{t-1}^{c,k})$$

$$Cov(z_{t-1}^{a,n}, y_{t-1}^{z,k'}) = Cov(z_{t-1}^{a,n}, r_{yz}^{k} z_{t-1}^{c,k} + r_{yy}^{k} y_{t-1}^{c,k}) = r_{yz}^{k} Cov(z_{t-1}^{a,n}, z_{t-1}^{c,k}) + r_{yy}^{k} Cov(z_{t-1}^{a,n}, y_{t-1}^{c,k})$$

$$Cov(y_{t-1}^{a,n}, z_{t-1}^{z,k'}) = Cov(y_{t-1}^{a,n}, r_{zz}^{k} z_{t-1}^{c,k} + r_{zy}^{k} y_{t-1}^{c,k}) = r_{zz}^{k} Cov(y_{t-1}^{a,n}, z_{t-1}^{c,k}) + r_{zy}^{k} Cov(y_{t-1}^{a,n}, y_{t-1}^{c,k})$$

$$Cov(y_{t-1}^{a,n}, y_{t-1}^{z,k'}) = Cov(y_{t-1}^{a,n}, r_{yz}^{k} z_{t-1}^{c,k} + r_{yy}^{k} y_{t-1}^{c,k}) = r_{yz}^{k} Cov(y_{t-1}^{a,n}, z_{t-1}^{c,k}) + r_{yy}^{k} Cov(y_{t-1}^{a,n}, y_{t-1}^{c,k})$$

where ”a” and ”c” are siblings in law of degree 2. Notice that since ”c” is the spouse of ”z”, $k = f$ when $k' = m$ and viceversa. We can compute spouses of siblings in law of any degree analogously.

Sibling of the sibling in law of degree 2

We have to compute the covariances between ”x” and ”d”. Let $n' = m, f$ be the gender of ”x” and $j = m, f$ the gender of the ”d”. We project x on his/her sibling b who has gender l

$$Cov(z_{t-1}^{x,n'}, z_{t-1}^{d,j}) = Cov(r_{zz}^{n',l} z_{t-1}^{b,l} + r_{zy}^{n',l} y_{t-1}^{b,l}, z_{t-1}^{d,j}) = r_{zz}^{n',l} Cov(z_{t-1}^{b,l}, z_{t-1}^{d,j}) + r_{zy}^{n',l} Cov(y_{t-1}^{b,l}, z_{t-1}^{d,j})$$

$$Cov(z_{t-1}^{x,n'}, y_{t-1}^{d,j}) = Cov(r_{zz}^{n',l} z_{t-1}^{b,l} + r_{zy}^{n',l} y_{t-1}^{b,l}, y_{t-1}^{d,j}) = r_{zz}^{n',l} Cov(z_{t-1}^{b,l}, y_{t-1}^{d,j}) + r_{zy}^{n',l} Cov(y_{t-1}^{b,l}, y_{t-1}^{d,j})$$

$$Cov(y_{t-1}^{x,n'}, z_{t-1}^{d,j}) = Cov(r_{yz}^{n',l} z_{t-1}^{b,l} + r_{yy}^{n',l} y_{t-1}^{b,l}, z_{t-1}^{d,j}) = r_{yz}^{n',l} Cov(z_{t-1}^{b,l}, z_{t-1}^{d,j}) + r_{yy}^{n',l} Cov(y_{t-1}^{b,l}, z_{t-1}^{d,j})$$

$$Cov(y_{t-1}^{x,n'}, y_{t-1}^{d,j}) = Cov(r_{yz}^{n',l} z_{t-1}^{b,l} + r_{yy}^{n',l} y_{t-1}^{b,l}, y_{t-1}^{d,j}) = r_{yz}^{n',l} Cov(z_{t-1}^{b,l}, y_{t-1}^{d,j}) + r_{yy}^{n',l} Cov(y_{t-1}^{b,l}, y_{t-1}^{d,j})$$

where ”b” and ”d” are siblings in law of degree 2. We can compute siblings of siblings in law of any degree analogously.

Cousins in law

We have to compute the covariances between ”ax” and ”cz”. Let $a^* = m, f$ be the gender of ”ax” and

$k^* = m, f$ the gender of the "cz". We project cz on c (his/her father or mother) who has gender k

$$Cov(z_t^{ax,n^*}, z_t^{cz,k^*}) = Cov(z_t^{ax,n^*}, G_{zk}^{k^*} z_{t-1}^{c,k} + G_{yk}^{k^*} y_{t-1}^{c,k}) = G_{zk}^{k^*} Cov(z_t^{ax,n^*}, z_{t-1}^{c,k}) + G_{yk}^{k^*} Cov(z_t^{ax,n^*}, y_{t-1}^{c,k})$$

$$Cov(z_t^{ax,n^*}, y_t^{cz,k^*}) = Cov(z_t^{ax,n^*}, B_{zk}^{k^*} z_{t-1}^{c,k} + B_{yk}^{k^*} y_{t-1}^{c,k}) = B_{zk}^{k^*} Cov(z_t^{ax,n^*}, z_{t-1}^{c,k}) + B_{yk}^{k^*} Cov(z_t^{ax,n^*}, y_{t-1}^{c,k})$$

$$Cov(y_t^{ax,n^*}, z_t^{cz,k^*}) = Cov(y_t^{ax,n^*}, G_{zk}^{k^*} z_{t-1}^{c,k} + G_{yk}^{k^*} y_{t-1}^{c,k}) = G_{zk}^{k^*} Cov(y_t^{ax,n^*}, z_{t-1}^{c,k}) + G_{yk}^{k^*} Cov(y_t^{ax,n^*}, y_{t-1}^{c,k})$$

$$Cov(y_t^{ax,n^*}, y_t^{cz,k^*}) = Cov(y_t^{ax,n^*}, B_{zk}^{k^*} z_{t-1}^{c,k} + B_{yk}^{k^*} y_{t-1}^{c,k}) = B_{zk}^{k^*} Cov(y_t^{ax,n^*}, z_{t-1}^{c,k}) + B_{yk}^{k^*} Cov(y_t^{ax,n^*}, y_{t-1}^{c,k})$$

where "c" is the sibling in law of the uncle/aunt of "ax". We can compute cousins in law of any degree analogously.

B Assortative mating only in z

If we assume that there is assortative mating only in z, that is $r_{zy}^m = r_{yy}^m = r_{zy}^f = r_{yy}^f = 0$, we have that

$$r_{zy}^m = \frac{1}{(1-\rho_{z^m y^m}^2)} \frac{\sigma_{z^f}}{\sigma_{y^m}} (\rho_{y^m z^f} - \rho_{z^m y^m} \rho_{z^m z^f}) = 0 \Rightarrow \rho_{y^m z^f} = \rho_{z^m y^m} \rho_{z^m z^f} \tag{14}$$

$$r_{yy}^m = \frac{1}{(1-\rho_{z^m y^m}^2)} \frac{\sigma_{y^f}}{\sigma_{y^m}} (\rho_{y^m y^f} - \rho_{z^m y^m} \rho_{z^m y^f}) = 0 \Rightarrow \rho_{y^m y^f} = \rho_{z^m y^m} \rho_{z^m y^f} \tag{15}$$

$$r_{zy}^f = \frac{1}{(1-\rho_{z^f y^f}^2)} \frac{\sigma_{z^m}}{\sigma_{y^f}} (\rho_{z^m y^f} - \rho_{z^f y^f} \rho_{z^m z^f}) = 0 \Rightarrow \rho_{z^m y^f} = \rho_{z^f y^f} \rho_{z^m z^f} \tag{16}$$

$$r_{yy}^f = \frac{1}{(1-\rho_{z^f y^f}^2)} \frac{\sigma_{y^m}}{\sigma_{y^f}} (\rho_{y^m y^f} - \rho_{z^f y^f} \rho_{y^m z^f}) = 0 \Rightarrow \rho_{y^m y^f} = \rho_{z^f y^f} \rho_{y^m z^f}$$

and the other coefficients of the linear projections

$$r_{zz}^m = \frac{\sigma_{z^f}}{\sigma_{z^m}} \rho_{z^m z^f}$$

$$r_{yz}^m = \frac{\sigma_{y^f}}{\sigma_{z^m}} \rho_{z^m y^f} \underset{\text{Using (16)}}{=} \frac{\sigma_{y^f}}{\sigma_{z^m}} \rho_{z^f y^f} \rho_{z^m z^f}$$

$$r_{zz}^f = \frac{\sigma_{z^m}}{\sigma_{z^f}} \rho_{z^m z^f}$$

$$r_{yz}^f = \frac{\sigma_{y^m}}{\sigma_{z^f}} \rho_{y^m z^f} \underset{\text{Using (14)}}{=} \frac{\sigma_{y^m}}{\sigma_{z^f}} \rho_{z^m y^m} \rho_{z^m z^f}$$

only depend on $\rho_{z^m z^f}$, $\rho_{z^m y^m}$ and $\rho_{z^f y^f}$.

Moreover, substituting the expressions for $\rho_{y^m z^f}$ and $\rho_{z^m y^f}$ in (14) and (16), in the steady state equations for $\rho_{z^m y^m}$ and $\rho_{z^f y^f}$ ((12) and (13)), we have that

$$\begin{aligned}&(1-\beta^m\alpha_y^m\gamma^m\alpha_z^m)\rho_{z^m y^m}-\beta^m(1-\alpha_y^m)\gamma^m(1-\alpha_z^m)\frac{\sigma_{z^f}}{\sigma_{z^m}}\frac{\sigma_{y^f}}{\sigma_{y^m}}\rho_{z^f y^f}\\ = \;&\frac{\sigma_{z^m}}{\sigma_{y^m}}+\beta^m\alpha_y^m\gamma^m(1-\alpha_z^m)\rho_{z^m y^m}\rho_{z^m z^f}\frac{\sigma_{z^f}}{\sigma_{z^m}}+\beta^m(1-\alpha_y^m)\gamma^m\alpha_z^m\rho_{z^f y^f}\rho_{z^m z^f}\frac{\sigma_{y^f}}{\sigma_{y^m}}\end{aligned}$$

$$\begin{aligned}&-\beta^f(1-\alpha_y^f)\gamma^f(1-\alpha_z^f)\frac{\sigma_{z^m}}{\sigma_{z^f}}\frac{\sigma_{y^m}}{\sigma_{y^f}}\rho_{z^m y^m}+(1-\beta^f\alpha_y^f\gamma^f\alpha_z^f)\rho_{z^f y^f}\\ = \;&\frac{\sigma_{z^f}}{\sigma_{y^f}}+\beta^f\alpha_y^f\gamma^f(1-\alpha_z^f)\rho_{z^f y^f}\rho_{z^m z^f}\frac{\sigma_{z^m}}{\sigma_{z^f}}+\beta^f(1-\alpha_y^f)\gamma^f\alpha_z^f\rho_{z^m y^m}\rho_{z^m z^f}\frac{\sigma_{y^m}}{\sigma_{y^f}}\end{aligned}$$

which can be written as:

$$\left[\begin{array}{ll} 1-\beta^m\alpha_y^m\gamma^m\left(\alpha_z^m+(1-\alpha_z^m)\rho_{z^m z^f}\frac{\sigma_{z^f}}{\sigma_{z^m}}\right) & -\beta^m(1-\alpha_y^m)\gamma^m\frac{\sigma_{y^f}}{\sigma_{y^m}}\left((1-\alpha_z^m)\frac{\sigma_{z^f}}{\sigma_{z^m}}+\alpha_z^m\rho_{z^m z^f}\right) \\ -\beta^f(1-\alpha_y^f)\gamma^f\frac{\sigma_{y^m}}{\sigma_{y^f}}\left((1-\alpha_z^f)\frac{\sigma_{z^m}}{\sigma_{z^f}}+\alpha_z^f\rho_{z^m z^f}\right) & 1-\beta^f\alpha_y^f\gamma^f\left(\alpha_z^f+(1-\alpha_z^f)\rho_{z^m z^f}\frac{\sigma_{z^m}}{\sigma_{z^f}}\right)\end{array}\right]\left(\begin{array}{c}\rho_{z^m y^m}\\ \rho_{z^f y^f}\end{array}\right)$$
$$=\left(\begin{array}{c}\frac{\sigma_{z^m}}{\sigma_{y^m}}\\ \frac{\sigma_{z^f}}{\sigma_{y^f}}\end{array}\right)$$

and therefore the coefficients of the linear projections can be written as a function of $\rho_{z^m z^f}$ (and the γs, βs, etc.). The model then has 17 parameters instead of 20.

C Assortative mating only in y

If we assume that there is assortative mating only in y, that is $r_{zz}^m = r_{yz}^m = r_{zz}^f = r_{yz}^f = 0$, we have that

$$r_{zz}^m = \frac{1}{(1-\rho_{z^m y^m}^2)}\frac{\sigma_{z^f}}{\sigma_{z^m}}(\rho_{z^m z^f}-\rho_{z^m y^m}\rho_{y^m z^f}) = 0 \Rightarrow \rho_{z^m z^f} = \rho_{z^m y^m}\rho_{y^m z^f}$$

$$r_{yz}^m = \frac{1}{(1-\rho_{z^m y^m}^2)}\frac{\sigma_{y^f}}{\sigma_{z^m}}(\rho_{z^m y^f}-\rho_{z^m y^m}\rho_{y^m y^f}) = 0 \Rightarrow \rho_{z^m y^f} = \rho_{z^m y^m}\rho_{y^m y^f}$$

$$r_{zz}^f = \frac{1}{(1-\rho_{z^f y^f}^2)}\frac{\sigma_{z^m}}{\sigma_{z^f}}(\rho_{z^m z^f}-\rho_{z^f y^f}\rho_{z^m y^f}) = 0 \Rightarrow \rho_{z^m z^f} = \rho_{z^f y^f}\rho_{z^m y^f}$$

$$r_{yz}^f = \frac{1}{(1-\rho_{z^f y^f}^2)}\frac{\sigma_{y^m}}{\sigma_{z^f}}(\rho_{y^m z^f}-\rho_{z^f y^f}\rho_{y^m y^f}) = 0 \Rightarrow \rho_{y^m z^f} = \rho_{z^f y^f}\rho_{y^m y^f}$$

and the coefficients of the linear projections

$$r_{zy}^{m} = \frac{\sigma_{z^f}}{\sigma_{y^m}}\rho_{y^m z^f} = \frac{\sigma_{z^f}}{\sigma_{y^m}}\rho_{z^f y^f}\rho_{y^m y^f}$$

$$r_{yy}^{m} = \frac{\sigma_{y^f}}{\sigma_{y^m}}\rho_{y^m y^f}$$

$$r_{zy}^{f} = \frac{\sigma_{z^m}}{\sigma_{y^f}}\rho_{z^m y^f} = \frac{\sigma_{z^m}}{\sigma_{y^f}}\rho_{z^m y^m}\rho_{y^m y^f}$$

$$r_{yy}^{f} = \frac{\sigma_{y^m}}{\sigma_{y^f}}\rho_{y^m y^f}$$

only depend on $\rho_{y^m y^f}$, $\rho_{z^m y^m}$ and $\rho_{z^f y^f}$.

Moreover, substituting $\rho_{y^m z^f}$ and $\rho_{z^m y^f}$ in the steady state equations for $\rho_{z^m y^m}$ and $\rho_{z^f y^f}$ ((12) and (13))

$$\begin{aligned} & (1-\beta^m\alpha_y^m\gamma^m\alpha_z^m)\rho_{z^m y^m} - \beta^m(1-\alpha_y^m)\gamma^m(1-\alpha_z^m)\frac{\sigma_{z^f}}{\sigma_{z^m}}\frac{\sigma_{y^f}}{\sigma_{y^m}}\rho_{z^f y^f} \\ = \; & \frac{\sigma_{z^m}}{\sigma_{y^m}} + \beta^m\alpha_y^m\gamma^m(1-\alpha_z^m)\rho_{z^f y^f}\rho_{y^m y^f}\frac{\sigma_{z^f}}{\sigma_{z^m}} + \beta^m(1-\alpha_y^m)\gamma^m\alpha_z^m\rho_{z^m y^m}\rho_{y^m y^f}\frac{\sigma_{y^f}}{\sigma_{y^m}} \end{aligned}$$

$$\begin{aligned} & -\beta^f(1-\alpha_y^f)\gamma^f(1-\alpha_z^f)\frac{\sigma_{z^m}}{\sigma_{z^f}}\frac{\sigma_{y^m}}{\sigma_{y^f}}\rho_{z^m y^m} + (1-\beta^f\alpha_y^f\gamma^f\alpha_z^f)\rho_{z^f y^f} \\ = \; & \frac{\sigma_{z^f}}{\sigma_{y^f}} + \beta^f\alpha_y^f\gamma^f(1-\alpha_z^f)\rho_{z^m y^m}\rho_{y^m y^f}\frac{\sigma_{z^m}}{\sigma_{z^f}} + \beta^f(1-\alpha_y^f)\gamma^f\alpha_z^f\rho_{z^f y^f}\rho_{y^m y^f}\frac{\sigma_{y^m}}{\sigma_{y^f}} \end{aligned}$$

which can be written as:

$$\begin{bmatrix} 1-\beta^m\alpha_z^m\gamma^m\left(\alpha_y^m+(1-\alpha_y^m)\rho_{y^m y^f}\frac{\sigma_{y^f}}{\sigma_{y^m}}\right) & -\beta^m(1-\alpha_z^m)\gamma^m\frac{\sigma_{z^f}}{\sigma_{z^m}}\left((1-\alpha_y^m)\frac{\sigma_{y^f}}{\sigma_{y^m}}+\alpha_y^m\rho_{y^m y^f}\right) \\ -\beta^f(1-\alpha_z^f)\gamma^f\frac{\sigma_{z^m}}{\sigma_{z^f}}\left((1-\alpha_y^f)\frac{\sigma_{y^m}}{\sigma_{y^f}}+\alpha_y^f\rho_{y^m y^f}\right) & 1-\beta^f\alpha_z^f\gamma^f\left(\alpha_y^f+(1-\alpha_y^f)\rho_{y^m y^f}\frac{\sigma_{y^m}}{\sigma_{y^f}}\right) \end{bmatrix}\begin{pmatrix} \rho_{z^m y^m} \\ \rho_{z^f y^f} \end{pmatrix}$$
$$= \begin{pmatrix} \frac{\sigma_{z^m}}{\sigma_{y^m}} \\ \frac{\sigma_{z^f}}{\sigma_{y^f}} \end{pmatrix}$$

and therefore the coefficients of the linear projections can be written as a function of $\rho_{y^m y^f}$ (and the γs, βs, etc.). The model then has 17 parameters instead of 20.

D The reduced form model

We now consider a reduced form model where the outcome y for an individual from generation t only depends on his father, and is given by

$$y_t = \beta y_{t-1}^m + z_t + x_t + u_t \tag{17}$$

Moreover, the socioeconomic status of the child, z_t, only depends on the father z_{t-1}^m

$$z_t = \gamma z_{t-1}^m + e_t + v_t \tag{18}$$

Substituting (18) in (17)

$$y_t = \beta y_{t-1}^m + \gamma z_{t-1}^m + e_t + v_t + x_t + u_t \tag{19}$$

Regarding the shocks, as in the general model, we assume that x_t and e_t are shared by all siblings and are uncorrelated to each other and with the other variables (in particular with z_{t-1}^m and y_{t-1}^m). Finally u_t and v_t are white-noise errors.

We can now compare (18) and (19) with the expression for z_t^k and y_t^k as a function of the father obtained in Section 1. For the general model we have that

$$z_t^k = G_{zm}^k z_{t-1}^m + G_{ym}^k y_{t-1}^m + g_m^k \omega_{t-1}^m + e_t^k + v_t^k \tag{20}$$

$$y_t^k = B_{ym}^k y_{t-1}^m + B_{zm}^k z_{t-1}^m + b_m^k \varepsilon_{t-1}^m + g_m^k \omega_{t-1}^m + e_t^k + v_t^k + x_t^k + u_t^k \tag{21}$$

where

$$\begin{aligned} G_{zm}^k &= \gamma^k(\alpha_z^k + (1-\alpha_z^k) r_{zz}^m) \\ G_{ym}^k &= \gamma^k (1-\alpha_z^k) r_{zy}^m \\ g_m^k &= \gamma^k (1-\alpha_z^k) \end{aligned}$$

$$\begin{aligned} B_{ym}^k &= \beta^k \left(\alpha_y^k + (1-\alpha_y^k) r_{yy}^m\right) + G_{ym}^k \\ B_{zm}^k &= \beta^k (1-\alpha_y^k) r_{yz}^m + G_{zm}^k \\ b_m^k &= \beta^k (1-\alpha_y^k) \end{aligned}$$

It is then trivial to see that there are two key differences between the two models:

1. The errors in (18) are assumed to be orthogonal to y_{t-1}^m and z_{t-1}^m, whereas in (20), z_t^k depends on y_{t-1}^m and therefore, unless $G_{ym}^k = 0$, if we project y_{t-1}^m on z_{t-1}^m, the new error will be correlated to y_{t-1}^m.
2. In the reduced form model, z_{t-1}^m has the same coefficient in (18) and (17), whereas in the general model, the coefficient of z_{t-1}^m is different in (20) and (21).

Then, we have that the general model can be written as a reduced form model if and only if

1. $G^k_{ym} = 0 \Longleftrightarrow \gamma^k = 0$, or $\alpha^k_z = 1$ or $r^m_{zy} = 0$
2. $B^k_{zm} = G^k_{zm} \Longleftrightarrow \beta^k = 0$, or $\alpha^k_y = 1$ or $r^m_{yz} = 0$

We then have that the general model can be written as a reduced form model

1. In the trivial case when just the father matters ($\alpha^k_z = \alpha^k_y = 1$, or $\alpha^k_z = 1$ and $\beta^k = 0$, or $\alpha^k_y = 1$ and $\gamma^k = 0$).
2. When $\beta^k = 0$ and y^m_{t-1} does not influence z^f_{t-1} once the effect of z^m_{t-1} has been netted out ($r^m_{zy} = 0$).
3. When $\gamma^k = 0$ and z^m_{t-1} does not influence y^f_{t-1} once the effect of y^m_{t-1} has been netted out ($r^m_{yz} = 0$).
4. When y^m_{t-1} does not influence z^f_{t-1} once the effect of z^m_{t-1} has been netted out ($r^m_{zy} = 0$), and z^m_{t-1} does not influence y^f_{t-1} once the effect of y^m_{t-1} has been netted out ($r^m_{yz} = 0$).

Notice that Case 2 corresponds to a latent factor model with assortative mating only in z and Case 3 to a direct effect model with assortative mating only in y.

E The genetic model

The genetic model is nested in our general model by imposing the following restrictions:

- There is no a direct effect of parents outcome on children outcome ($\beta^k = 0$, $k = f, m$)
- The latent factor is genetic and therefore it is transmitted from parents to children as

$$z^k_t = \frac{z^m_{t-1} + z^f_{t-1}}{2} + v^k_t$$

  where v^k_t is uncorrelated across relatives and to z^m_{t-1} and z^f_{t-1} ($\gamma^k = 1$ and $\sigma^2_{e_k} = 0$, $k = f, m$)
- The share of the variance explained by the latent factor is equal across genders ($\sigma^2_{z^k} = \sigma^2_z$, $k = f, m$)
- There is assortative mating only in the observed outcome y ($\rho_{z^m y^f}$, $\rho_{y^m z^f}$ and $\rho_{z^m z^f}$ are functions of $\rho_{y^m y^f}$ and some of the other parameters of the model (see Section 3)).

The genetic model has only 5 parameters: $\sigma^2_z, \sigma^2_{x^m}, \sigma^2_{x^f}, \sigma_{x^m x^f}, \rho_{y^m y^f}$.

E.1 Genetic transmission

Suppose that each person has only one gene with two alleles, A and B. One of the allele is inherited from the father and the other one from the mother. Let X_A and X_B be the random variables representing the potential values each allele may take and suppose that the outcome of interest Y depends on $Z = X_A + X_B$

$$Y = Z + U$$

where U is mean independent of Z.

We have to compute

$$E\left(Z_t \,|\, Z^m_{t-1} = z^m,\, Z^f_{t-1} = z^f\right)$$

The distribution of the child Z_t conditional on the parents $X^m_{A,t-1} = x^m_A$, $X^m_{B,t-1} = x^m_B$, $X^f_{A,t-1} = x^f_A$, $X^f_{B,t-1} = x^f_B$ is multinomial with the following probability mass function

$$Z_t = \begin{cases} x^m_A + x^f_A, \text{ with probability } \frac{1}{4} \\ x^m_A + x^f_B, \text{ with probability } \frac{1}{4} \\ x^m_B + x^f_A, \text{ with probability } \frac{1}{4} \\ x^m_B + x^f_B, \text{ with probability } \frac{1}{4} \end{cases}$$

Then, the distribution of Z_t conditional on $X^m_{A,t-1} = x^m_A$, $Z^m_{t-1} = z^m$, $X^f_{A,t-1} = x^f_A$, $Z^f_{t-1} = z^f$ is also multinomial, and the probability mass function is

$$Z_t = \begin{cases} x^m_A + x^f_A, \text{ with probability } \frac{1}{4} \\ x^m_A + z^f - x^f_A, \text{ with probability } \frac{1}{4} \\ z^m - x^m_A + x^f_A, \text{ with probability } \frac{1}{4} \\ z^m - x^m_A + z^f - x^f_A, \text{ with probability } \frac{1}{4} \end{cases}$$

Then,

$$\begin{aligned} & E\left(Z_t \,|\, X^m_{A,t-1},\ Z^m_{t-1},\ X^f_{A,t-1},\ Z^f_{t-1}\right) \\ = \ & \frac{1}{4}\left(X^m_{A,t-1} + X^f_{A,t-1}\right) + \frac{1}{4}\left(X^m_{A,t-1} + Z^f_{t-1} - X^f_{A,t-1}\right) \\ & + \frac{1}{4}\left(Z^m_{t-1} - X^m_{A,t-1} + X^f_{A,t-1}\right) \frac{1}{4}\left(Z^m_{t-1} - X^m_{A,t-1} + Z^f_{t-1} - X^f_{A,t-1}\right) \\ = \ & \frac{1}{4}Z^f_{t-1} + \frac{1}{4}Z^m_{t-1} + \frac{1}{4}\left(Z^m_{t-1} + Z^f_{t-1}\right) = \frac{1}{2}\left(Z^m_{t-1} + Z^f_{t-1}\right) \end{aligned}$$

Since $E\left(Z_t \,|\, X^m_{A,t-1},\ Z^m_{t-1},\ X^f_{A,t-1},\ Z^f_{t-1}\right)$ does not depend on $X^m_{A,t-1}$ and $X^f_{A,t-1}$, using the law of iterated expectations

$$E\left(Z_t \,|\, Z^m_{t-1},\ Z^f_{t-1}\right) = \frac{1}{2}\left(Z^m_{t-1} + Z^f_{t-1}\right)$$

If we now consider that each person has n genes, each with two alleles, A_i and B_i. For each gene, one of the allele is inherited from the father and the other one from the mother. Let X_{A_i} and X_{B_i} be the random variables representing the potential values each allele may take and suppose that the outcome of interest Y depends on $Z = \sum_{i=1}^{n} \left(X_{A_i} + X_{B_i} \right) = \sum_{i=1}^{n} Z_i$, where $Z_i = X_{A_i} + X_{B_i}$. For each gene i, given $X^m_{A_i,t-1} = x^m_{A_i}$, $Z^m_{i,t-1} = z^m_i$, $X^f_{A_i,t-1} = x^f_{A_i}$, there are 4 possible realizations of the child alleles of gene i, and therefore 4^n possible genomes. Then, the distribution of the child Z_t conditional on the parents $X^m_{A_1,t-1} = x^m_{A_1}$, $Z^m_{1,t-1} = z^m_1$, $X^f_{A_1,t-1} = x^f_{A_1}$, $Z^f_{1,t-1} = z^f_1$, ...,$X^m_{A_n,t-1} = x^m_{A_n}$, $Z^m_{n,t-1} = z^m_n$, $X^f_{A_n,t-1} = x^f_{A_n}$, $Z^f_{n,t-1} = z^f_n$ is also multinomial, and the probability mass function is

$$Z_t = \begin{cases} \left(x^m_{A_1} + x^f_{A_1}\right) + \left(x^m_{A_2} + x^f_{A_2}\right) + .. + \left(x^m_{A_n} + x^f_{A_n}\right), \text{ with probability } \frac{1}{4^n} \\ \left(x^m_{A_1} + z^f_1 - x^f_{A_1}\right) + \left(x^m_{A_2} + x^f_{A_2}\right) + .. + \left(x^m_{A_n} + x^f_{A_n}\right), \text{ with probability } \frac{1}{4^n} \\ \left(z^m_1 - x^m_{A_1} + x^f_{A_1}\right) + \left(x^m_{A_2} + x^f_{A_2}\right) + .. + \left(x^m_{A_n} + x^f_{A_n}\right), \text{ with probability } \frac{1}{4^n} \\ \left(z^m_1 - x^m_{A_1} + z^f_1 - x^f_{A_1}\right) + \left(x^m_{A_2} + x^f_{A_2}\right) + .. + \left(x^m_{A_n} + x^f_{A_n}\right), \text{ with probability } \frac{1}{4^n} \\ \vdots \\ \left(z^m_1 - x^m_{A_1} + z^f_1 - x^f_{A_1}\right) + \left(z^m_2 - x^m_{A_2} + z^f_2 - x^f_{A_2}\right) + .. + \left(z^m_n - x^m_{A_n} + z^f_n - x^f_{A_n}\right), \text{ with probability } \frac{1}{4^n} \end{cases}$$

As in the model with just one gene, all the $x^k_{A_i}$ cancel when we compute the conditional mean

$$\begin{aligned} & E\left(Z_t \,|\, X^m_{A_1,t-1}, Z^m_{1,t-1}, X^f_{A_1,t-1}, Z^f_{1,t-1}, ..., X^m_{A_n,t-1}, Z^m_{n,t-1}, X^f_{A_n,t-1}, Z^f_{n,t-1}\right) \\ = \; & \frac{1}{2} Z^m_{1,t-1} + ... + \frac{1}{2} Z^m_{n,t-1} + \frac{1}{2} Z^f_{1,t-1} + ... + \frac{1}{2} Z^f_{n,t-1} = \frac{1}{2}\left(Z^m_{t-1} + Z^f_{t-1}\right) \end{aligned}$$

Then, since the conditional expectation above only depends on Z^m_{t-1} and Z^f_{t-1}, using the law of iterated expectations

$$E\left(Z_t \,|\, Z^m_{t-1},\ Z^f_{t-1}\right) = \frac{1}{2}\left(Z^m_{t-1} + Z^f_{t-1}\right)$$

Lets now consider two siblings i and j. We have to compute

$$E\left(Z_{it} Z_{jt} \,|\, Z^m_{t-1} = z^m,\, Z^f_{t-1} = z^f\right)$$

The distribution of $Z_{it} Z_{jt}$ conditional on the parents $X^m_{A,t-1} = x^m_A$, $X^m_{B,t-1} = x^m_B$, $X^f_{A,t-1} = x^f_A$, $X^f_{B,t-1} = x^f_B$

is multinomial with the following probability mass function

$$Z_{it}Z_{jt} = \begin{cases} \left(x_A^m + x_A^f\right)^2, \text{ with probability } \frac{1}{16} \\ \left(x_A^m + x_A^f\right)\left(x_A^m + x_B^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_A^m + x_A^f\right)\left(x_B^m + x_A^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_A^m + x_A^f\right)\left(x_B^m + x_B^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_A^m + x_B^f\right)^2, \text{ with probability } \frac{1}{16} \\ \left(x_A^m + x_B^f\right)\left(x_B^m + x_A^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_A^m + x_B^f\right)\left(x_B^m + x_B^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_B^m + x_A^f\right)^2, \text{ with probability } \frac{1}{16} \\ \left(x_B^m + x_A^f\right)\left(x_B^m + x_B^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_B^m + x_B^f\right)^2, \text{ with probability } \frac{1}{16} \end{cases}$$

Then, the distribution of $Z_{it}Z_{jt}$ conditional on $X_{A,t-1}^m = x_A^m$, $Z_{t-1}^m = z^m$, $X_{A,t-1}^f = x_A^f$, $Z_{t-1}^f = z^f$ is also multinomial, and the probability mass function is

$$Z_{it}Z_{jt} = \begin{cases} \left(x_A^m + x_A^f\right)^2, \text{ with probability } \frac{1}{16} \\ \left(x_A^m + x_A^f\right)\left(x_A^m + z^f - x_A^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_A^m + x_A^f\right)\left(z^m - x_A^m + x_A^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_A^m + x_A^f\right)\left(z^m - x_A^m + z^f - x_A^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_A^m + z^f - x_A^f\right)^2, \text{ with probability } \frac{1}{16} \\ \left(x_A^m + z^f - x_A^f\right)\left(z^m - x_A^m + x_A^f\right), \text{ with probability } \frac{1}{8} \\ \left(x_A^m + z^f - x_A^f\right)\left(z^m - x_A^m + z^f - x_A^f\right), \text{ with probability } \frac{1}{8} \\ \left(z^m - x_A^m + x_A^f\right)^2, \text{ with probability } \frac{1}{16} \\ \left(z^m - x_A^m + x_A^f\right)\left(z^m - x_A^m + z^f - x_A^f\right), \text{ with probability } \frac{1}{8} \\ \left(z^m - x_A^m + z^f - x_A^f\right)^2, \text{ with probability } \frac{1}{16} \end{cases}$$

Then,

$$
\begin{aligned}
& E\left(Z_{it}Z_{jt}\,|\,X^m_{A,t-1},\ Z^m_{t-1},\ X^f_{A,t-1},\ Z^f_{t-1}\right)\\
= & \frac{1}{16}\left(X^m_A+X^f_A\right)\left(\left(X^m_A+X^f_A\right)+\left(X^m_A+Z^f-X^f_A\right)+\left(Z^m-X^m_A+X^f_A\right)+\left(Z^m-X^m_A+Z^f-X^f_A\right)\right)\\
& +\frac{1}{16}\left(X^m_A+Z^f-X^f_A\right)\left(\left(X^m_A+X^f_A\right)+\left(X^m_A+Z^f-X^f_A\right)+\left(Z^m-X^m_A+X^f_A\right)+\left(Z^m-X^m_A+Z^f-X^f_A\right)\right)\\
& +\frac{1}{16}\left(Z^m-X^m_A+X^f_A\right)\left(\left(X^m_A+X^f_A\right)+\left(X^m_A+Z^f-X^f_A\right)+\left(Z^m-X^m_A+X^f_A\right)+\left(Z^m-X^m_A+Z^f-X^f_A\right)\right)\\
& +\frac{1}{16}\left(Z^m-X^m_A+Z^f-X^f_A\right)\Big(\left(X^m_A+X^f_A\right)+\left(X^m_A+Z^f-X^f_A\right)+\left(Z^m-X^m_A+X^f_A\right)+\Big(Z^m-X^m_A+Z^f-X^f_A\\
= & \frac{1}{8}\left(X^m_A+X^f_A\right)\left(Z^f+Z^m\right)+\frac{1}{8}\left(X^m_A+Z^f-X^f_A\right)\left(Z^f+Z^m\right)\\
& +\frac{1}{8}\left(Z^m-X^m_A+X^f_A\right)\left(Z^f+Z^m\right)+\frac{1}{8}\left(Z^m-X^m_A+Z^f-X^f_A\right)\left(Z^f+Z^m\right)\\
= & \frac{1}{4}\left(Z^m_{t-1}+Z^f_{t-1}\right)^2
\end{aligned}
$$

Since $E\left(Z_{it}Z_{jt}\,|\,X^m_{A,t-1},\ Z^m_{t-1},\ X^f_{A,t-1},\ Z^f_{t-1}\right)$ does not depend on $X^m_{A,t-1}$ and $X^f_{A,t-1}$, using the law of iterated expectations

$$E\left(Z_{it}Z_{jt}|\,Z^m_{t-1},\ Z^f_{t-1}\right)=\frac{1}{4}\left(Z^m_{t-1}+Z^f_{t-1}\right)^2$$

We then have that

$$E\left(Z_{it}Z_{jt}|\,Z^m_{t-1},\ Z^f_{t-1}\right)=E\left(Z_{it}|\,Z^m_{t-1},\ Z^f_{t-1}\right)E\left(Z_{jt}|\,Z^m_{t-1},\ Z^f_{t-1}\right)$$

and Z_{it} and Z_{jt} are uncorrelated conditional on Z^m_{t-1}, Z^f_{t-1}. Then, we can write

$$Z_{it}=\frac{1}{2}\left(Z^m_{t-1}+Z^f_{t-1}\right)+e_{it}$$

where e_{it} and e_{jt} are uncorrelated across siblings.