Kinship Correlations and Intergenerational Mobility*

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

November 2017
Very preliminary draft

Abstract

We propose a new methodology to estimate long-run intergenerational socioeconomic mobility. Our specification takes into account assortative mating in the marriage market and is general enough to encompass the standard model as well as the specification recently proposed by Gregory Clark. Our approach does not require to have information about the variable of interest for individuals in several generations and make use of the correlations among individuals with different degrees of kinship in the same generation. In our empirical application we use census data from Spain and find a high degree of persistence that corroborates some of Clark's findings.

1 Introduction

The analysis of the degree of socioeconomic intergenerational mobility has attracted the attention of many economists in recent years (see for example Chetty et al 2014). Part of the interest on this topic is due, at least in the case of income mobility, to its possible relation with the increasing income inequality experienced recently in some economies (Corak 2013). An additional factor to explain this interest is the existence of recent studies showing that mobility in the long-run is perhaps much lower than what most economists used to think (Long and Ferrie 2013, Clark 2014, Lindahl et al. 2014). This recent literature has started to change the standard view about mobility across multiple generations, which used to assume that the correlation between grandparents and grandchildren outcomes is basically the square of the parent-offspring correlation. Since for most relevant outcomes such as income or education, parent-offspring correlations are always moderate, economists had often assumed that the correlation between individuals in one generation and their ancestors in different generations decreases really fast as we go back in time, so that after, say, three or four generations the link is already very weak. However, recent empirical studies suggest a much higher persistence rate in socioeconomic status and a significant link with grandparents and even with great-grandparents (Lindahl et al 2014).

An important contribution in this area has been the work by Clark (2014) who claims that mobility across several generations, for income as well as for other outcomes, is low due to the existence of a latent variable, the "underlying social competence" of families, which is inherited from parents and has a high persistence rate. If such latent variable indeed plays an important role in the transmission of socioeconomic status the standard regressions of offspring outcomes against parent's outcomes will be downward biased and the true persistence will be higher than suggested

*We are grateful to Gregory Clark for very helpful comments and suggestions. (a) Universidad de Alicante; (b) Universidad Carlos III

by the regression coefficient. Clark (2014) asses the role of such latent variable using a methodology based on the use of surnames. His approach requires information on the outcome of interest for individuals in several generations. Using data from a series of countries and periods of time Clark finds a very low degree of intergenerational mobility. Furthermore, the degree of mobility is very similar across countries and time. Lindahl et al (2014) and Braun and Stuhler (2016) also find a low degree of long-run intergenerational mobility but not as low as in Clark (2014).

One main problem with the approach adopted in these works is the data requirement, since for many countries it is difficult to obtain comparable information for more than two generations about outcomes such as income or educational levels. For instance, in many countries we typically find that there is very little variation in years of formal education for older generations because the majority of the population had just basic education. Thus, we propose a new approach to asses the degree of long-run intergenerational mobility that does not require information on previous generations. To apply our methodology we just need "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[1]. The idea behind our method is quite 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[2]. In particular, we compute the correlation for years of schooling for different degrees of kinship (brothers, fathers-sons, first-cousins and uncles-nephews) using census data from a Spanish region. If we have enough of these moments we can calibrate all the parameters of a reduced form model on intergenerational mobility. Our results from this calibration exercise are very much consistent with the high persistence hypothesis proposed by Clark. In particular we find that the persistence rate for the "underlying social competence" of families is around 0.8. Consistent with this result, our approach predicts 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.[3]

Our approach is also related to the literature on siblings correlations (See Solon 1999, Levine and Mazumder 2007, Björklund and Jäntti 2009 and Schnitzlein 2014). Most of the papers in this literature aim at estimating the impact of family background on an observable outcome such as income, education, etc. The family background is a latent component that accounts for all factors shared by siblings that are orthogonal to the parental outcome. We extend these models by decomposing the family background into an inheritable and a non-inheritable component. The idea is that by using correlations on outcomes of relatives of different degrees of kinship we are able to disentangle the non-inheritable part of family background that is only shared by sibling from the inheritable part that is also partially shared by cousins, second cousins, etc., through their common ancestor. 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

[1] Hällsten (2014) also computes the correlation in some observed outcomes for cousins and second cousins and points to the advantage of using data for the same generation.

[2] Güell et al. (2015) also make use of "horizontal" information to estimate intergenerational mobility, but their approach is quite different and based on the used of surnames.

[3] 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 than here individuals are classified according to 10 possible levels of education (years of schooling)

observed outcomes for different type of siblings like MZ twins, DZ twins, siblings, half siblings, adoptees, etc. (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, the first term representing the genetic factors, the second one represents the 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 in particular they 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. An advantage of our approach is that we don't 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 non-inheratable components. Furthermore, since we are not interested on measuring the direct effect of genes[4], our analysis is not based on the correlations of different types of twins.

Thus, we propose a new method to asses the degree of long-run mobility that can be seen as complementary to the one used recently by several economists. 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.

The empirical results suggest that long-run intergenerational mobility might be quite low. Because we only calibrate a reduced form model it's difficult to get policy conclusions from our findings. However, the fact that the latent variable underlying the social competence of families explains a high part of the variance in levels of education suggests that public intervention policies should pay more attention to the role of the family.

The paper proceeds as follows. Section 2 sets out the basic model and develops our method. Section 3 presents our main empirical findings and the robustness checks. Section 4 extends the basic model to account for assortative mating and the potential influence of mothers. Section 5 concludes. We include some additional information about the models in the Appendixes.

2 Theory

Suppose that y is the outcome of interest in our economy, for example income, education or wealth. Since in our empirical exercise such outcome will be the level of education henceforth we identify y with years of schooling but all our theoretical results are valid to study other outcomes as, for example, income. We want to study the link of such variable y between individuals and their ancestors. We consider a reduced form of Becker-Tomes (1979) model similar to the one in Solon

[4] In recent years, the availability of molecular genetic data has allowed the use of a new methodology to asses the influence of genetic factors. See Okbay et al. (2016) for the results of a genome-wide association study (GWAS) for years of schooling.

(2014)

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

where $t-1$ denotes the father's generation and t the children's generation, y_{t-1} denotes years of schooling of his father, z_t^i denotes a latent variable that is inherited from the parents, x_t is a shock shared by all brothers in the family which is uncorrelated with the other variables (in particular with z_t), and u_t^i is an individual's white-noise error term. In principle the variable z_t^i might include common genes and family values and depending on whether there are perfect credit markets or not father's wealth could be also part of it. The variable x_t could capture factors like the type of neighborhood, common friends and perhaps the influence of one sibling on another. These are factors that siblings might share but are not inherited from parents.

The latent variable z_t^i is often omitted in this type of analysis and it has been introduced by Clark (2014) who sees it as the "underlying social competence" of families and assumes that

$$z_t^i = \gamma z_{t-1} + e_t^i \tag{2}$$

where z_{t-1} denotes the father's value of such latent variable and e_t^i is an individual white-noise term.[5] Thus, the "underlying social competence" is passed from fathers to their sons with persistence rate γ.

The traditional approach does not consider the existence of such variable z while Clark's approach assumes that $\beta = 0$.[6] We take a more general view and a priory do not exclude any possibility and let the data determine which model is the correct one. If the traditional approach is the correct one we should find that z is zero (or close to zero) whereas if Clark's model is the correct one we should find very low values of β and significant values of z and γ. As noticed, many recent papers also allow for the possibility of an unobservable variable. The identification strategy is usually based on twins, adoptees and instrumental variables (see Holmlund 2011 for a survey of this literature). However, they neither estimate the magnitude of our heritable variable z nor the persistence parameter γ.

We suppose we are in the steady state and therefore the persistence parameters β and γ, the distribution of z_{t-1} and y_{t-1} and all the covariances remain the same across generations. Under the standard approach the parameter β is estimated by regressing child's years of schooling on parental years of schooling. Since z is unobservable, estimating γ in Clark's model requires to have observations not only on sons years of schooling and fathers years of schooling but also on grandparents years of schooling (see Clark 2014). Unfortunately, in many cases it's difficult to get good data on the outcome of interest for a large sample of individuals from more than two different generations. We propose a new methodology that only requires information on the years of schooling of individuals in two generations, and sometimes only information from one generation. The idea behind our method is quite simple: if the model specified by (1) and (2) is correct and we have the necessary data, we can compute the correlations on years of schooling for different degrees of kinship, for example the correlation for brothers, father-son, first-cousins, second-cousins, uncle-nephew and so on. If we have enough of these moments we can calibrate all the parameters of the model.[7] To compute some of these moments we need information about individuals from the same generation (brothers, first-cousins, second cousins,...) and if we have information about a previous

[5] We assume that siblings errors e_t^i are uncorrelated. Our results are robust to imposing the restriction that siblings get the same realization of e_t.

[6] Clark (2014) does not need to include x because he does not use data on brothers, cousins, etc.

[7] If we have enough moments we can consider an even more general model in which the parameter β in the current generation could be different from the one in the previous generation.

generation we might also compute correlations for father-sons and uncle-nephews. In some cases one can have data on grandparents and compute the grandfather-grandson correlation.

Write as σ_y^2, σ_z^2 and σ_x^2 the variances of y, z and x respectively. We can write the covariance in years of schooling between brothers i and j in generation t as

$$Cov_b(y_t^i, y_t^j) = \beta^2\sigma_y^2 + \frac{2\beta\gamma\sigma_z^2}{1-\beta\gamma} + \gamma^2\sigma_z^2 + \sigma_x^2$$

and the correlation as

$$\rho_b = \frac{Cov_b(y_t^i, y_t^j)}{\sigma_y^2} \tag{3}$$

In Appendix A we show the corresponding correlations for father-son, grandfather-grandson, uncle-nephew, first-cousins, second-cousins and third-cousins. If we know at least four of the previous correlations we can calibrate the model to determine the values of those unknowns. Notice that when computing the correlations for cousins i and j we assume that $Cov(x^i, x^j) = 0$ because otherwise the model would be under-identified. In Section 3 we provide some empirical evidence to justify this assumption.

We might not find an exact solution to such system of equations within the range of values that we consider feasible in our economy. In that case we determine $\beta, \gamma, \sigma_z^2$ and σ_x^2 by solving the following minimization problem

$$Min_{\{\beta,\gamma,\sigma_z^2,\sigma_x^2\}\in F} \quad \sum_{i\in C} p_i(\rho_i - \overline{\rho}_i)^2 \tag{4}$$

where $\overline{\rho}_i$ is the value of the observed correlation, p_i is the sample size used to calculate correlation ρ_i, F is the set of feasible values for the four unknowns, and C is the set of correlations for which we have reliable data (for example brothers, cousins, second-cousins, fathers-son).

3 Empirical Application I

In this section we apply the method proposed in section 2 to calibrate the model using census data from the Spanish region of Cantabria.

3.1 The data

To apply our methodology we need data on extended families. The 2001 population census for Spain, which is available nationwide, does not allow to identify families unless they are living in the same house. However, for the region of Cantabria we have information on the full name of each person and we can use this information to identify fathers and sons, brothers, uncles and nephews, and cousins. The census contains information, among other variables, on the gender, age and educational level of all individuals living in the region ($526,339$ persons). We define the t-generation as all males born in Cantabria between 1956 and 1976 ($71,479$ males and $68,830$ females) and the $(t-1)$-generation as their parents. Surnames in Spain are passed from parents to children according to the following rule: A newborn person, regardless of gender, receives two surnames that will keep 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 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 son 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. We use the information on the educational level to assign years of schooling to each person following Calero et.al.[8] 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 2 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)

Table 1

	Men				Women			
	Matched		Unmatched		Matched		Unmatched	
	Mean	St. Dev	Mean	St. Dev	Mean	St. Dev	Mean	St. 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
Number of observations	25,860		45,619		24,610		44,220	

Once we have identified parents and children, siblings are immediately identified. Finally, we identify siblings in the parents generation when there are only two individuals in that generation sharing the same two surnames. Once siblings in the parents generation are identified, uncles and nephews, and cousins are immediately identified. The strategy to identify siblings in the parents' generation is quite conservative in the sense that it is unlikely that we identify as brothers individuals who actually are not brothers, but we pay the price of having smaller sample sizes for cousins and uncles-nephews than for fathers-sons or brothers.

3.2 The benchmark case

We use the sample of males and the correlations between brothers, cousins (whose fathers are brothers), fathers and sons, and nephews and uncles (the uncle being brother of the father). The empirical covariances are first computed for each family and then averaged across families as suggested in Solon, Page and Duncan (2000). The empirical correlations are obtained by dividing the empirical covariances by the product of the standard deviations.[9] The empirical correlations and

[8] 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 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.

[9] Notice that the standard deviation of y is 3.705 for the current generation and 3.831 for the parents generations. Therefore, the empirical correlations for fathers and sons, and uncles and nephews would have been slightly larger if we would have divided the covariances by the variance of y

the number of families and pairs used to compute those correlations are presented in Table 2.

Table 2				
	Brothers	Father-son	Cousins	Uncle-nephew
Correlations	0.467	0.379	0.196	0.232
Number of families	6,022	17,663	746	1,921
Number of pairs	11,109	25,860	1,654	2,843

These correlations are within the values estimated in some other developed countries (see Hertz 2007 and Björklund and Salvanes 2011). We solve the minimization problem (4) with the four moments to obtain[10]

Table 3			
β	γ	σ_z^2	σ_x^2
0	0.790	6.586	2.303

We next compare the empirical correlations with the predicted correlations for these values of $\beta, \gamma, \sigma_z^2$ and σ_x^2

Table 4				
Correlations	Brothers	Father-son	Cousins	Uncle-nephew
Observed	0.467	0.379	0.196	0.232
Predicted	0.467	0.374	0.187	0.236
% Error	0%	-0.025%	-4.784%	1.859%

Since we don't have data on the correlation for other relatives, as grandfather-grandson, we cannot compare it with the correlation predicted by the model. However, we can compare the square of the father-son correlation with the grandfather-grandson correlation predicted by the model

Predicted grandfather-son	(father-son)^2
0.299	0.144

This result is in accordance with Clark's view and with some recent empirical evidence (Lindahl et al 2015). The grandfather-grandson correlation is much stronger than the squared of the father-son correlation.

It's useful to asses how much of the total variance of y_t is explained by the different components of the model. We have

$$\begin{aligned} \sigma_y^2 &= \beta^2\sigma_y^2 + \sigma_z^2 + 2\beta Cov(y_{t-1}, z_t) + \sigma_x^2 + \sigma_u^2 \\ &= \beta^2\sigma_y^2 + \sigma_z^2 + 2\frac{\beta\gamma\sigma_z^2}{1-\beta\gamma} + \sigma_x^2 + \sigma_u^2 \end{aligned} \tag{5}$$

The part of the variance σ_y^2 explained directly by the father's years of schooling is $\beta^2\sigma_y^2$. The part directly explained by the latent variable z is σ_z^2 while the part explained by the shocks shared by brothers is σ_x^2. We standardize years of schooling so that $\sigma_y^2 = 1$ and obtain the following decomposition

Table 5				
Total explained	$\beta^2\sigma_{y_{t-1}}^2$	σ_z^2	σ_x^2	$2\beta Cov(y_{t-1}, z_t)$
0.648	0	0.480	0.168	0

[10] We use Mathematica to solve all the minimization problems in this paper. The codes and the details of all the computations are available upon request.

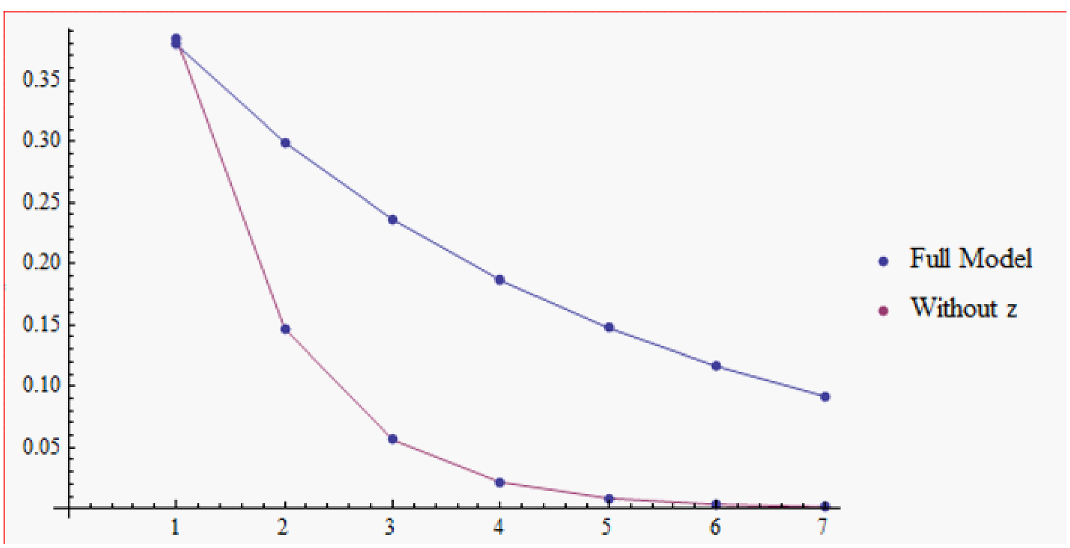
Thus, the results in our benchmark case favour Clark's view that long-run mobility is much lower than suggested by most economists, and that a large share of the persistence is explained by an inherited latent variable with a high rate of persistence ($\gamma = 0.79$).

It's important to mention that both x and z are essential to obtain a satisfactory calibration of the model. Thus, if we drop x from the model and repeat our previous procedure we again obtain a very high value of the persistence parameter γ. However, in this case the (over-identified) model performs quite poorly at predicting the correlations. This is not surprising since previous works have already shown the importance of this type of shock to understand the correlations between brothers[11] (Branigan et al. 2013). If we now drop z, we obtain a non negligible $\beta = 0.384$ but the fit regarding cousins and uncle-nephew correlations is very poor. The predictions based on these two models are presented in Table 6

Table 6

Correlations	Brothers	Father-son	Cousins	Uncle-nephew
Observed	0.467	0.379	0.196	0.232
Dropping x. Predicted	0.367	0.397	0.313	0.339
% Error	-21.47%	4.691%	59.82%	46.09%
Dropping z. Predicted	0.475	0.384	0.070	0.182
% Error	1.794%	1,190%	-64.33%	-21.42%

To better appreciate the consequences of these findings, Figure 1 shows the predicted correlations for individuals at the current generation and their ancestors, i.e. their fathers, grandfathers, great-grandfathers, etc., based on the full model and on the model without z.



[11] In the standard ACE approach the outcome of interest is decomposed in the genetic componet (A), a family or shared environment componet (C), and the idiosyncratic componet (E). Our variable x might be seen as a part of the environment component C, since x only captures the inheritated shared environment. In many studies the componet C explains a very small part of the total variance. In our case just the component x explains about 17% of the total variance. However, in the meta-analysis on educational attainment carried out by Branigan et al. (2013) the part of the variance explained by C is usually larger than 20%, what is consistent with our finding.

Figure 1

As it is very clear from the figure, the persistence based on the model without z is low, so that after a few generations the influence of ancestors vanishes almost completely. Our approach, however, provides a more pessimistic view about intergenerational mobility in the long run. Thus, we find that, under the assumption of stability of the parameters of the model, the correlation between the levels of y of individuals in the current generation and the levels of y of their ancestors seven generations back in time is still as high as 9%.

A possible concern is that our estimate of γ is biased because of the assumption that the shocks x for cousin pairs are uncorrelated. One would suspect that an important component of these shocks is related to the geographical location of the individuals, and cousins might tend to live in the same region. Remember that x captures the shocks that are shared by siblings and are not inherited from parents, so that if cousins live in the same location as their parents that is part of the inherited components and is not incorporated in x. Thus, we check in our data the location of cousins who do not live in the same location as their parents: Among the pairs of cousins such that at least one of them does not live in the same municipality than his parents, the probability that they live in the same municipality is 10%, basically the same probability as for two men taken randomly from the whole population (11%). However, such probability is 34.5% for couples of brothers that do not live in the same municipality as their parents. Thus, we are confident that our results do not critically depend on the assumption of uncorrelated non-inherited environments among cousins.

A second concern is that the minimization problem (4) might present some additional local minimum with value of the objetive function very close to the value of the global minimum but with very different values of $\beta, \gamma, \sigma_z^2$ and σ_x^2. Indeed, in our case we find a local minimum at $\beta = 0.25, \gamma = 1, \sigma_z^2 = 3.4$ and $\sigma_x^2 = 1.22$ that also yields an acceptable prediction of our four moments. However, we rule out this case because the extreme value $\gamma = 1$ implies the unreasonable situation in which the latent variable z is passed from fathers to sons without any noise. It's interesting to notice that such set of values would predict a correlation for 2nd cousins of 0.16, much higher than the correlation of 0.116 predicted by our chosen values. Unfortunately we don't have information on 2d cousins but such difference in predictions suggest that having such additional "horizontal" information could really overidentified the model[12].

3.3 Robustness checks

One possible concern with our previous analysis is about the robustness of our findings to changes in the values of the observed empirical correlations. For this reason we repeat our procedure for 1,000 different sets of values of the four correlations $\overline{\rho}_b, \overline{\rho}_{fs}, \overline{\rho}_{un}, \overline{\rho}_{c1}$. These values are obtained by carrying out 1,000 random draws from our original sample, each draw selecting 75% of the original individuals in the current generation.[13] Table 7 reports the mean values of the four unknowns obtained under this procedure and compares it with the ones reported in the above benchmark case. Notice that the average value of β is 0.065 which is very close with the estimated causal effect of schooling in the survey of Holmlund et al. (2011).

[12] The 2nd cousins correlations in GPA provided in Hällsten (2014) are always lower than 0.12, except for the correlation among the wealthiest 1%.

[13] Since the number of pairs of uncle-nephew is not that large we consider that draws of 75% of the whole sample are better than draws of 50%. The results for the 50% case, which are in the vast majority of cases very similar to the ones reported here, are available upon request.

Table 7				
	β	γ	σ_z^2	σ_x^2
Mean value	0.065	0.840	5.136	2.500
Benchmark case	0	0.790	6.586	2.303

Table 8 shows the mean correlations predicted by the model and the observed ones.

Table 8				
Mean correlations	Brothers	Father-son	Cousins	Uncle-nephew
Predicted	0.460	0.377	0.215	0.257
Observed	0.471	0.383	0.207	0.246

Figure 2 shows for these 1,000 subsamples the values of β and γ and the standardized σ_z^2, σ_x^2 as well as $\beta^2\sigma_y^2$ (the part of the variance of y_t directly explained by y_{t-1}).[14] The different cases are ordered according to the obtained values of β.

[14] Notice that we have normalized σ_z^2 and σ_x^2 to σ_y^2.

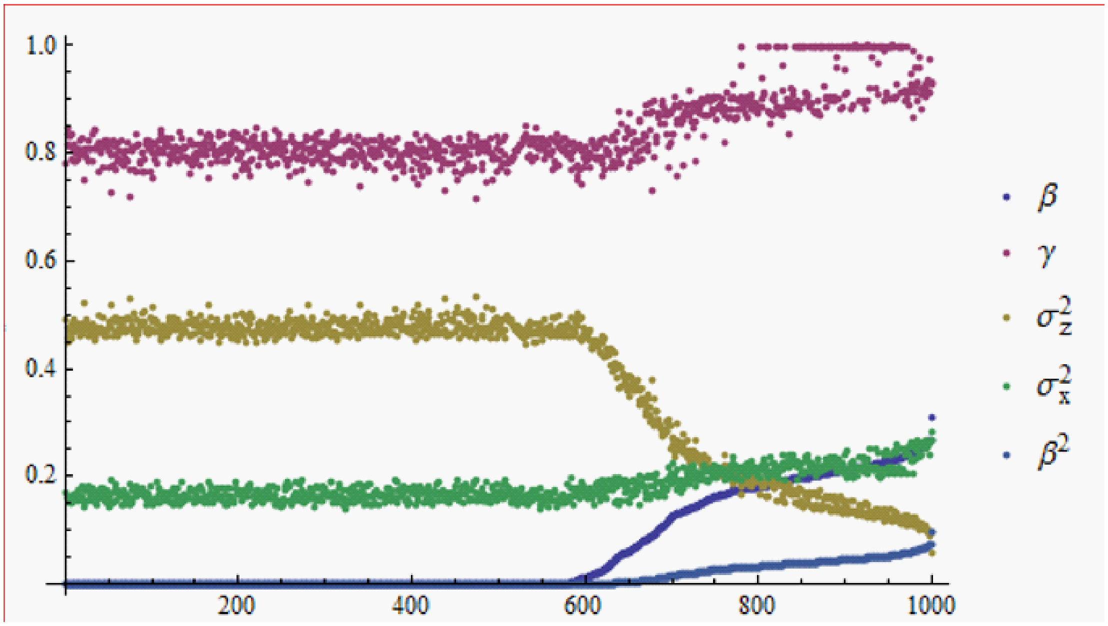


Figure 2

The basic facts that arise from this robustness check exercise are: i) The persistence parameter is always quite high and in almost all the cases greater than 0.75;[15] ii) The largest values of β are around 0.2, but even for those cases the part of the total variance of y_t explained directly by y_{t-1} is small and in all the simulations but one, smaller than the part of the variance explained by the latent variable z.

Thus, the main findings and conclusions obtained in the benchmark case are robust to these changes in the values of the observed correlations.[16]

4 A model with assortative mating

The model we were considering did not take into account the potential influence of the mother in the outcome of the children. We now extend the previous model to incorporates mothers and assortative mating. We assume that the value of the output y for an individual from generation t is given by

$$y_t^k = \beta^k \widetilde{y}_{t-1}^k + z_t^k + x_t^k + u_t^k \tag{6}$$

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

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

[15] Only 18% of the estimated γ are smaller 0.79, the value found for the benchmark case.

[16] We have carried out additional robustness checks. In particular we first have repeated the same robustness check as the one here but for draws of 50% of the original sample, and second we have solved our minimization problem for other 256 economies obtained by considering values of the correlations within a $\pm 10\%$ deviation from the benchmark case values. The results are again similar and are provided upon request.

where $\alpha_y^k \in [0,1]$,so that $\widetilde{y}_{t-1}^k$ can be seen as the weighted average years of schooling of parents. 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

$$\begin{aligned} z_t^k &= \gamma^k \widetilde{z}_{t-1}^k + e_t^k \\ \widetilde{z}_{t-1}^k &= \alpha_z^k z_{t-1}^m + (1-\alpha_z^k) z_{t-1}^f \end{aligned} \tag{7}$$

where $\alpha_z^k \in [0,1]$

Regarding the shocks, we assume that x_t^k is shared by all 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}). Finally u_t^k is an individual's white-noise error term.

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:

$$\begin{pmatrix} z_{t-1}^f \\ y_{t-1}^f \end{pmatrix} = \begin{pmatrix} r_{zz}^m & r_{zy}^m \\ r_{yz}^m & r_{yy}^m \end{pmatrix} \begin{pmatrix} z_{t-1}^m \\ y_{t-1}^m \end{pmatrix} + \begin{pmatrix} w_{t-1}^m \\ \varepsilon_{t-1}^m \end{pmatrix} \tag{8}$$

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 $\rho_{z^m y^m}, \rho_{z^m z^f}$, $\rho_{z^m y^f}$, $\rho_{y^m z^f}$, $\rho_{y^m y^f}$, σ_{z^m}, σ_{z^f}, σ_{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:

$$\begin{pmatrix} z_{t-1}^m \\ y_{t-1}^m \end{pmatrix} = \begin{pmatrix} r_{zz}^f & r_{zy}^f \\ r_{yz}^f & r_{yy}^f \end{pmatrix} \begin{pmatrix} z_{t-1}^f \\ y_{t-1}^f \end{pmatrix} + \begin{pmatrix} w_{t-1}^f \\ \varepsilon_{t-1}^f \end{pmatrix} \tag{9}$$

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 $\rho_{z^f y^f}, \rho_{z^m z^f}$, $\rho_{z^m y^f}$, $\rho_{y^m z^f}$, $\rho_{y^m y^f}$, σ_{z^m}, σ_{z^f}, σ_{y^m} and σ_{y^f}. In Appendix B 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. Then, since we can directly estimate σ_{y^m}, σ_{y^f} and $\rho_{y^m y^f}$ from the data, we have 16 unknown parameters that write as the vector v:

$$v = \{\beta^m, \gamma^m, \sigma_{z^m}, \sigma_{x^m}, \beta^f, \gamma^f, \sigma_{z^f}, \sigma_{x^f}, \rho_{x^m x^f}, \rho_{z^m z^f}, \rho_{z^m y^f}, \rho_{y^m z^f}, \alpha_y^m, \alpha_z^m, \alpha_y^f, \alpha_z^f, \}$$

and therefore we need at least 16 correlations between relatives of different kinship to calibrate these parameters. The inclusion of females into the model allows us to use the following 22 correlations: husband and wife, brothers, sisters, brother-sister, three types of male cousins (fathers are brothers, mothers are sisters, and father and mother are brother an sister) and analogously three of female cousins, four types of male-female cousins (fathers are brothers, mothers are sisters, father of the male is brother of the mother of the female, and mother of the male is sister of the father of the female), son-father, daughter-father, son-mother, daughter-mother, two types of nephew-uncle (brother of the father and brother of the mother) and analogously two of nice-uncle.[17]

[17] The formulas for these correlations as functions of the parameters are presented in Appendix B.

4.1 Empirical Application II

We calibrate the parameters in v by solving the following minimization problem[18]

$$Min_{v \in F} \sum_{i \in C} p_i(\rho_i - \overline{\rho}_i)^2 \tag{10}$$

where ρ_i are the theoretical correlations, $\overline{\rho}_i$ the empirical correlations, p_i the number of families used to calculate each correlation (see Table 10), F is the set of feasible values for the unknown parameters[19], and C is the set of correlations we mentioned above.

The calibrated parameters are presented in Table 9

Table 9			
β^m	γ^m	$\sigma^2_{z^m}$	$\sigma^2_{x^m}$
0	0.818	6.600	2.252
β^f	γ^f	$\sigma^2_{z^f}$	$\sigma^2_{x^f}$
0	0.839	5.491	2.006
$\rho_{x^m x^f}$	$\rho_{z^m z^f}$	$\rho_{z^m y^f}$	$\rho_{y^m z^f}$
0.779	0.895	0.579	0.601
α^m_z	α^m_y	α^f_z	α^f_y
0.828	-	0.488	-

The picture we obtain is again consistent with Clark's results. Both β^m and β^f are basically zero, whereas γ^m and γ^f are around 0.8. This means that the observable outcome is transferred from parents to children indirectly through the latent variable z, which is very persistent. Another remarkable result is the large degree of assortative mating in z (correlation of 0.895 between z^m and z^f) which suggests that the possible genetic part of it cannot be that large, and probably culture and identifiable preferences and values form the largest part of z[20]. Notice that since β^m and β^f are zero α^m_y and α^f_y are not identified.

Regarding the fitting, we have computed the predicted correlations based on this parameters and we compare them with the empirical correlations. The results are presented in Table 10. The fit is remarkable taking into account that we try to match 22 moments using 12 parameters. As we expected, the empirical correlations based on a large number of pairs of observations, which are likely to be quite accurate, are very close to the predicted ones, whereas those based on a smaller number of pairs are less close. This result is not only due to the weights used, a quite similar fit arises when we use equal weights.

[18] We have used Mathematica to solve the minimization problem. This is a more complex problem than the one discussed in Section 3 since there are many local minima. The code is in the online appendix.

[19] The parameters $\beta^m, \gamma^m, \beta^f, \gamma^f, \alpha^m_y, \alpha^m_z, \alpha^f_y, \alpha^f_z$, have to be between 0 and 1, but the correlation can take negative values.

[20] Cavalli-Sforza and Feldman (1981) suugest that a high correlation in this type of variable might be explained by the existence of "horizontal transmission" of cultural attitudes, which makes that the values of z for the father and for the mother converge by the time they raise their offspring. However, it's not clear how such horizontal transmission could explain our finding since in our model z exclusively refers to vertical transmitted values.

Table 10				
	N. families	Empirical	Predicted	Error (%)
brothers	6022	0.467	0.467	0.010%
sisters	5662	0.437	0.437	0.000%
brother-sister	9525	0.414	0.414	-0.003%
male cousins				
(fathers are brothers)	746	0.196	0.190	-3.070%
(mothers are sisters)	670	0.213	0.190	-11.020%
(father and mother are brother and sister)	1146	0.209	0.188	-10.282%
female cousins				
(fathers are brothers)	670	0.170	0.174	2.616%
(mothers are sisters)	608	0.193	0.201	4.282%
(father and mother are brother and sister)	1063	0.207	0.185	-10.554%
male-female cousins				
(fathers are brothers)	1100	0.186	0.182	-2.127%
(mothers are sisters)	1003	0.168	0.195	16.254%
(father-male is brother of mother-female)	1112	0.218	0.193	-11.360%
(mother-male is sister of father-female)	1115	0.157	0.180	14.435%
son-father	17663	0.379	0.379	-0.007%
daughter-father	16982	0.360	0.360	-0.139%
son-mother	17663	0.328	0.328	-0.066%
daughter-mother	16982	0.335	0.335	0.042%
nephew-uncle (brother of the father)	1921	0.232	0.239	3.182%
nephew-uncle (brother of the mother)	1350	0.228	0.238	4.412%
nice-uncle (brother of the father)	1852	0.216	0.228	5.662%
nice-uncle (brother of the mother)	1298	0.247	0.245	-0.893%

We now decompose the variance of y into its different components as

$$\sigma^2_{y^k} = (\beta^k)^2 \sigma^2_{\widetilde{y}} + \sigma^2_{z^k} + \beta^k Cov(\widetilde{y}^k_{t-1}, z^k_t) + \sigma^2_{x^k} + \sigma^2_{u^k}$$

The results of these decompositions for males and females are presented in Table 11.[21] We can see that the model explains 64.5% of variance in years of schooling for males and 54.4% for females, with z and x accounting respectively for around 70% and 30% of the explained variance. Thus, the family factors (z and x) play a more determinant role in explaining years of schooling among men than among women.

Table 11					
	Total explained	$(\beta^k)^2\sigma^2_{\widetilde{y}^k_t}$	$\sigma^2_{z^k}$	$\sigma^2_{x^k}$	$2\beta^k Cov(\widetilde{y}^k_{t-1}, z^k_t)$
Males	0.645	0	0.481	0.164	0
Females	0.544	0	0.399	0.146	0

We can compute the predicted correlations for individuals at the current generation and their ancestors. However, we have now different possible "ancestors lines". Figure 3 shows those correlation for the male paternal line (son, fathers, grandfathers, great-grandfathers,...), and the female maternal line (daughter, mother, grandmother, great-grandmother,....). These correlations seem consistent with the ones provided in Figure 1.

[21] We standardize the different components to $\sigma^2_{y^m}$ for males and to $\sigma^2_{y^f}$ for females.

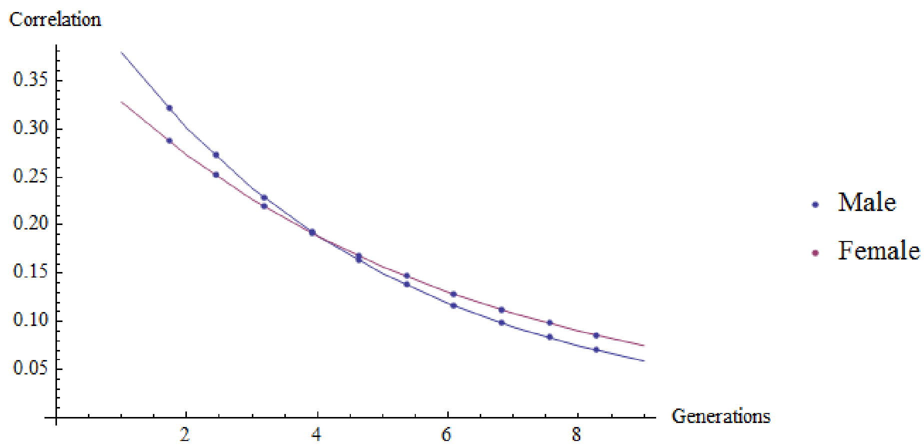


Figure 3

One might wonder here about the existence of other minimum in the problem (10). Indeed, we find a local minimum that yields a very close value of the objective function as the one above and takes values of the parameters $\beta^m = 0.28, \beta^f = 0.26$, $\gamma^m = 1, \gamma^f = 0.96$.As in our benchmark case without assortative mating, we rule out this case because the value of $\gamma = 1$ implies that the latent variable z is passed from fathers to sons without any noise.

5 Conclusions

We have proposed a method to asses the degree of intergenerational mobility which takes into account the possibility that a substantial part of the persistence in socioeconomic status might be due to the existence of a latent variable that is inherited from parents. The method is based on the correlations between a series of relatives and does not demand much information about individuals in previous generations. Our findings suggest that indeed such latent variable plays a very important role and is the reason why persistence in socioeconomic status is much stronger than what is commonly thought. Thus, our results are in line with Clark's claims about the low degree of social mobility in the long run. However, our exercise does not provide any new information in favor or against the possibility that the degree of intergenerational mobility is constant across different economies and time. We have applied our method to asses the degree of intergenerational mobility in a Spanish region and the extension to other regions and countries is an important task which is left for future research.

References

[1] Becker, Gary S., and Nigel Tomes (1979) "An equilibrium theory of the distribution of income and intergenerational mobility." *Journal of Political Economy* 87, pp 1153-1189.

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

[3] Behrman, 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.

[4] Björklund, A., Jäntti, M. and G. Solon, (2005) "Influences of natura and nurture on earnings variation", in *Unequal Chances. Family Bacground and Economic Success*, eds. S. Bowles, H. Gintis and M. O. Groves, Princeton University Press, New York.

[5] Björklund and Salvanes (2011) "Education and Family Background: Machanisms and Policies", in Eric A. Hanushek, S. Machin and L. Woessmann, eds. *Handbook of the Economics of Education*, vol3, pp 201-47, Amsterdam: North-Holland.

[6] Branigan, A. R., McCallum, K.J., and J. Freese (2013), "Variation in the Heritability of Educational Attainment: An International Meta-Analysis", *Social Forces* 92(1), 109-140

[7] Braun, S. and J. Stuhler (2016), "The Transmission of Inequality Across Multiple Generations: Testing Recent Theories with Evidence from Germany"

[8] Calero, J. J. O. Escardíbul, S. Waisgrais and M. Mediavilla, "Desigualdades socioeconómicas en el sistema educativo español". Secretaría General Técnica. Centro de Publicaciones. Ministerio de Educación y Ciencia (https://sede.educacion.gob.es/publiventa/desigualdades-socioeconomicas-en-el-sistema-educativo-espanol/investigacion-educativa/12330)

[9] Cavalli-Sforza, L.L. and M. Feldman (1981), *Cultural Transmission and Evolution: A Quantitative Approach*, Princeton University Press, New Jersey.

[10] Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, (2014) "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *Quarterly Journal of Economics*, pp 2633-2679.

[11] Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. (2014b). "Is the United States Still a Land of Opportunity?" *American Economic Review* 104 (5), pp. 141 47.

[12] Clark, Gregory (2014) *The son also rises: Surnames and the history of social mobility.* Princeton: Princeton University Press.

[13] Collado, M.D., I. Ortuño Ortín, and A. Romeu (2014), "Long-run intergenerational social mobility and the distribution of surnames".

[14] Corak, Miles (2013), Income Inequality, Equality of Opportunity, and Intergenerational Mobility", *Journal of Economic Perspectives*, 27,3, pp79-192

[15] Goldberger, A. S. (1979). "Heritability". *Economica* 46: 327-47.

[16] Guell, M.C. Telmer and J. Rodriguez Mora, (2015), "The Informational Content of Surnames, the Evolution of Intergenerational Mobility and Assortative Mating", *Review of Economic Studies*, 82 (2): 693-735

[17] Hällsten, M., (2014), "Inequality across three and four generations in Egalitarian Sweden: 1st and 2nd cousin correlations in socio-economic outcomes", *Research in Socila Stratification and Mobility*, 35, 19-33.

[18] Hertz et al (2007) "The inheritance of educational inequality: International comparisons and fifty-year trends", *The B.E. Journal of Economic Analysis & Policy*

[19] Holmlund, H., Lindahl, M. 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, 615-651.

[20] Jäntti, M and S.P. Jenkins (2015) "Income mobility", Chapter 10, pp. 807–935, in Handbook of Income Distribution, Volume 2A, edited by A. B. Atkinson and F. Bourguignon, Elsevier.

[21] Levine, D.I. and B. Mazumder, (2007), "The growing importance of family: Evidence from brothers' earnings", *Industrial Relations*, 46(1), 7-21

[22] Lindahl, Mikale, Marten Palme, Sofia Massih, and Anna Sjorgen (2015) "Long-Term Intergenerational Persistence of Human Capital: An Empirical Analysis of Four Generations", *Journal of Human Resources*. 50, 1, 1-33.

[23] Long, J. and J. Ferrie, (2013), "Intergenerational occupational mobility in Great Britain and the United States since 1850", *American Economic Review*, 103(4): 1109-1137.

[24] Okbay, Aysu et al. (2016), "Genome-wide association study identifies 74 loci associated with educational attainment", *Nature*, 533, pp 539-542

[25] Olivetti, C. and D. Paserman (2015), "In the name of the son (and the daughter): Intergenerational mobility in the United States, 1850-1940", *American Economic Review*, 105,8, pp 2695-2724

[26] Sacerdote, Bruce, (2011) "Sacerdote, Bruce, (2011) Nature and Nurture Effects On Children's Outcomes: What Have We Learned From Studies of Twins And Adoptees?, In: J. Benhabib, A. Bisin and M.O. Jackson (eds.) *Handbook of Social Economics* , vol 1. Elsevier, San Diego, pp 1-30

[27] Solon, Gary (1999), "Intergenerational mobility in the labour market," In: Ashenfelter, O. Card, D. (eds), *Handbook of Labor Economics*, vol 3. Elsevier Science B.V. , New York, pp. 1761-1800.

[28] Solon, Gary.(2014) "Theoretical models of inequality transmission across multiple generations", *Research in Social Stratification and Mobility*, 35, pp 13-18.

[29] Solon, Gary.(2015) "What Do We Know So Far about Multigenerational Mobility?", NBER, Working Paper 21053.

[30] 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

[31] Stuhler, Jan (2015), "Mobility Across Multiple Generations: The Iterated Regression Fallacy", IZA Discussion Paper 7072.

Appendix A

Consider a reduced form of Becker-Tomes (1979) model similar to the one in Solon (2014)

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

where $t-1$ denotes the father's generation and t the children's generation, y_{t-1} is father's years of schooling, y_t^i is child i's years of schooling, z_t^i is the child i's status, x_t is not inherited, is uncorrelated with y_{t-1} and z_t^i but is shared among brothers, and u_t^i is a random error that is uncorrelated with y_{t-1} and z_t^i.

Status is partially inherited so that

$$z_t^i = \gamma z_{t-1} + e_t^i$$

where e_t^i that is not correlated across brothers. Notice that when $\beta = 0$ we are in Clark's model.

We assume that the second order moments of all variables are time invariant. We present below the formulas for the covariances in years of scholing for relatives of different degrees of kinship. The correlations are computed by dividing the covariances by the variance of y.

Covariances

Brothers

We first compute the covariances between y_{t-1} and z_{t-1}

$$\begin{aligned} Cov(y_t^i, z_t^i) &= Cov(\beta y_{t-1} + z_t^i, z_t^i) = \beta Cov(y_{t-1}, z_t^i) + \sigma_z^2 \\ &= \beta Cov(y_{t-1}, \gamma z_{t-1}) + \sigma_z^2 = \beta\gamma Cov(y_{t-1}, z_{t-1}) + \sigma_z^2 \end{aligned}$$

and in the steady state we have $Cov(y_t, z_t) = Cov(y_{t-1}, z_{t-1})$, so that

$$Cov(y_{t-1}, z_{t-1}) = \frac{\sigma_z^2}{1 - \beta\gamma} \tag{11}$$

and the covariance between brothers is

$$Cov_b(y_t^i, y_t^j) = \beta^2 \sigma_y^2 + 2\beta\gamma Cov(y_{t-1}, z_{t-1}) + \gamma^2 \sigma_z^2 + \sigma_x^2 \tag{12}$$

Cousins

We first compute the following covariances for their fathers (who are brothers)

$$Cov_b(z_{t-1}^i, z_{t-1}^j) = \gamma^2 \sigma_z^2$$

and

$$Cov_b(y_{t-1}^i, z_{t-1}^j) = \beta\gamma Cov(y_{t-2}, z_{t-2}) + \gamma^2 \sigma_z^2$$

The covariance for male cousins whose fathers are brothers is

$$Cov_c(y_t^i, y_t^j) = \beta^2 Cov_b(y_{t-1}^i, y_{t-1}^j) + 2\beta\gamma Cov_b(y_{t-1}^i, z_{t-1}^j) + \gamma^2 Cov_b(z_{t-1}^{m,i}, z_{t-1}^{m,j})$$

Son-Father

$$Cov_{sf}(y_t^i, y_{t-1}) = \beta\sigma_y^2 + \gamma Cov(y_{t-1}, z_{t-1})$$

Nephew and uncle (brother of the father)

$$Cov_{neph-u}(y_t^i, y_{t-1}^j) = \beta Cov_b(y_{t-1}^i, y_{t-1}^j) + \gamma Cov_b(y_{t-1}^i, z_{t-1}^j)$$

Second cousins

We first compute the following covariances for their fathers (who are cousins)

$$Cov_c(z_{t-1}^i, z_{t-1}^j) = \gamma^2 Cov_b(z_{t-2}^i, z_{t-2}^j)$$

and

$$Cov_c(y_{t-1}^i, z_{t-1}^j) = \beta\gamma Cov_b(y_{t-2}, z_{t-2}) + \gamma^2 Cov_b(z_{t-2}^i, z_{t-2}^j)$$

The covariance for second cousins whose fathers are brothers is

$$Cov_{c2}(y_t^i, y_t^j) = \beta^2 Cov_c(y_{t-1}^i, y_{t-1}^j) + 2\beta\gamma Cov_c(y_{t-1}^i, z_{t-1}^j) + \gamma^2 Cov_c(z_{t-1}^i, z_{t-1}^j)$$

Third cousins

We first compute the following covariances for their fathers (who are second cousins)

$$Cov_{c2}(z_{t-1}^i, z_{t-1}^j) = \gamma^2 Cov_c(z_{t-2}^i, z_{t-2}^j)$$

and

$$Cov_{c2}(y_{t-1}^i, z_{t-1}^j) = \beta\gamma Cov_c(y_{t-2}, z_{t-2}) + \gamma^2 Cov_c(z_{t-2}^i, z_{t-2}^j)$$

The covariance for second cousins whose fathers are brothers is

$$Cov_{c3}(y_t^i, y_t^j) = \beta^2 Cov_{c2}(y_{t-1}^i, y_{t-1}^j) + 2\beta\gamma Cov_{c2}(y_{t-1}^i, z_{t-1}^j) + \gamma^2 Cov_{c2}(z_{t-1}^i, z_{t-1}^j)$$

Appendix B

The model we were considering did not take into account the potential influence of the mother in the outcome of the children. We now extend the previous model to incorporates mothers and assortative mating. We assume that the value of the output y for an individual from generation t is given by

$$y_t^k = \beta^k \widetilde{y}_{t-1}^k + z_t^k + x_t^k + u_t^k \tag{13}$$

where the superscript k stands for males ($k = m$) and for females ($k = f$). 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

$$\begin{aligned} z_t^k &= \gamma^k \widetilde{z}_{t-1}^k + e_t^k \\ \widetilde{z}_{t-1}^k &= \alpha_z^k z_{t-1}^m + (1 - \alpha_z^k) z_{t-1}^f \end{aligned} \tag{14}$$

Regarding the shocks, we assume that x_t^k is shared by all 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}). Finally u_t^k is an individual's white-noise error term.

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, and

$$\begin{aligned}
\begin{pmatrix} r_{zz}^m & r_{zy}^m \\ r_{yz}^m & r_{yy}^m \end{pmatrix}' &= \begin{pmatrix} \sigma_{z^m}^2 & \sigma_{z^m y^m} \\ \sigma_{z^m y^m} & \sigma_{y^m}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{z^m z^f} & \sigma_{z^m y^f} \\ \sigma_{y^m z^f} & \sigma_{y^m y^f} \end{pmatrix} \\
&= \frac{1}{\sigma_{z^m}^2 \sigma_{y^m}^2 - \sigma_{z^m y^m}^2} \begin{pmatrix} \sigma_{y^m}^2 & -\sigma_{z^m y^m} \\ -\sigma_{z^m y^m} & \sigma_{z^m}^2 \end{pmatrix} \begin{pmatrix} \sigma_{z^m z^f} & \sigma_{z^m y^f} \\ \sigma_{y^m z^f} & \sigma_{y^m y^f} \end{pmatrix} \\
&= \frac{1}{\sigma_{z^m}^2 \sigma_{y^m}^2 (1-\rho_{z^m y^m}^2)} \begin{pmatrix} \sigma_{y^m}^2 \sigma_{z^m z^f} - \sigma_{z^m y^m} \sigma_{y^m z^f} & \sigma_{y^m}^2 \sigma_{z^m y^f} - \sigma_{z^m y^m} \sigma_{y^m y^f} \\ \sigma_{z^m}^2 \sigma_{y^m z^f} - \sigma_{z^m y^m} \sigma_{z^m z^f} & \sigma_{z^m}^2 \sigma_{y^m y^f} - \sigma_{z^m y^m} \sigma_{z^m y^f} \end{pmatrix} \\
&= \frac{1}{(1-\rho_{z^m y^m}^2)} \begin{pmatrix} \frac{\sigma_{z^f}}{\sigma_{z^m}} (\rho_{z^m z^f} - \rho_{z^m y^m} \rho_{y^m z^f}) & \frac{\sigma_{y^f}}{\sigma_{z^m}} (\rho_{z^m y^f} - \rho_{z^m y^m} \rho_{y^m y^f}) \\ \frac{\sigma_{z^f}}{\sigma_{y^m}} (\rho_{y^m z^f} - \rho_{z^m y^m} \rho_{z^m z^f}) & \frac{\sigma_{y^f}}{\sigma_{y^m}} (\rho_{y^m y^f} - \rho_{z^m y^m} \rho_{z^m y^f}) \end{pmatrix}
\end{aligned}$$

We then have that

$$\begin{pmatrix} r_{zz}^m & r_{zy}^m \\ r_{yz}^m & r_{yy}^m \end{pmatrix} = \frac{1}{(1-\rho_{z^m y^m}^2)} \begin{pmatrix} \frac{\sigma_{z^f}}{\sigma_{z^m}} (\rho_{z^m z^f} - \rho_{z^m y^m} \rho_{y^m z^f}) & \frac{\sigma_{z^f}}{\sigma_{y^m}} (\rho_{y^m z^f} - \rho_{z^m y^m} \rho_{z^m z^f}) \\ \frac{\sigma_{y^f}}{\sigma_{z^m}} (\rho_{z^m y^f} - \rho_{z^m y^m} \rho_{y^m y^f}) & \frac{\sigma_{y^f}}{\sigma_{y^m}} (\rho_{y^m y^f} - \rho_{z^m y^m} \rho_{z^m y^f}) \end{pmatrix} \tag{15}$$

and

$$\begin{aligned}
r_{zz}^m &= \frac{1}{(1-\rho_{z^m y^m}^2)} \frac{\sigma_{z^f}}{\sigma_{z^m}} (\rho_{z^m z^f} - \rho_{z^m y^m} \rho_{y^m z^f}) \\
r_{zy}^m &= \frac{1}{(1-\rho_{z^m y^m}^2)} \frac{\sigma_{z^f}}{\sigma_{y^m}} (\rho_{y^m z^f} - \rho_{z^m y^m} \rho_{z^m z^f}) \\
r_{yz}^m &= \frac{1}{(1-\rho_{z^m y^m}^2)} \frac{\sigma_{y^f}}{\sigma_{z^m}} (\rho_{z^m y^f} - \rho_{z^m y^m} \rho_{y^m y^f}) \\
r_{yy}^m &= \frac{1}{(1-\rho_{z^m y^m}^2)} \frac{\sigma_{y^f}}{\sigma_{y^m}} (\rho_{y^m y^f} - \rho_{z^m y^m} \rho_{z^m y^f})
\end{aligned}$$

and 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 $\rho_{y^m y^f}$.

Notice that 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 write (14) as

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

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 (13) as

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

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

where

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

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 consider y_t^k and z_t^k as functions of mother's years of schooling, y_{t-1}^f, and social status z_{t-1}^f. We then also consider the linear projections of z_{t-1}^m and y_{t-1}^m on z_{t-1}^f and y_{t-1}^f:

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

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

$$\begin{aligned}
\begin{pmatrix} r_{zz}^f & r_{zy}^f \\ r_{yz}^f & r_{yy}^f \end{pmatrix}' &= \begin{pmatrix} \sigma_{z^f}^2 & \sigma_{z^f y^f} \\ \sigma_{z^f y^f} & \sigma_{y^f}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{z^f z^m} & \sigma_{z^f y^m} \\ \sigma_{y^f z^m} & \sigma_{y^f y^m} \end{pmatrix} \\
&= \frac{1}{\sigma_{z^f}^2 \sigma_{y^f}^2 - \sigma_{z^f y^f}^2} \begin{pmatrix} \sigma_{y^f}^2 & -\sigma_{z^f y^f} \\ -\sigma_{z^f y^f} & \sigma_{z^f}^2 \end{pmatrix} \begin{pmatrix} \sigma_{z^f z^m} & \sigma_{z^f y^m} \\ \sigma_{y^f z^m} & \sigma_{y^f y^m} \end{pmatrix} \\
&= \frac{1}{\sigma_{z^f}^2 \sigma_{y^f}^2 (1-\rho_{z^f y^f}^2)} \begin{pmatrix} \sigma_{y^f}^2 \sigma_{z^f z^m} - \sigma_{z^f y^f} \sigma_{y^f z^m} & \sigma_{y^f}^2 \sigma_{z^f y^m} - \sigma_{z^f y^f} \sigma_{y^f y^m} \\ \sigma_{z^f}^2 \sigma_{y^f z^m} - \sigma_{z^f y^f} \sigma_{z^f z^m} & \sigma_{z^f}^2 \sigma_{y^f y^m} - \sigma_{z^f y^f} \sigma_{z^f y^m} \end{pmatrix} \\
&= \frac{1}{(1-\rho_{z^f y^f}^2)} \begin{pmatrix} \frac{\sigma_{z^m}}{\sigma_{z^f}} (\rho_{z^f z^m} - \rho_{z^f y^f} \rho_{y^f z^m}) & \frac{\sigma_{y^m}}{\sigma_{z^f}} (\rho_{z^f y^m} - \rho_{z^f y^f} \rho_{y^f y^m}) \\ \frac{\sigma_{z^m}}{\sigma_{y^f}} (\rho_{y^f z^m} - \rho_{z^f y^f} \rho_{z^f z^m}) & \frac{\sigma_{y^m}}{\sigma_{y^f}} (\rho_{y^f y^m} - \rho_{z^f y^f} \rho_{z^f y^m}) \end{pmatrix}
\end{aligned}$$

We then have that

$$\begin{pmatrix} r_{zz}^f & r_{zy}^f \\ r_{yz}^f & r_{yy}^f \end{pmatrix} = \frac{1}{(1-\rho_{z^f y^f}^2)} \begin{pmatrix} \frac{\sigma_{z^m}}{\sigma_{z^f}} (\rho_{z^f z^m} - \rho_{z^f y^f} \rho_{y^f z^m}) & \frac{\sigma_{z^m}}{\sigma_{y^f}} (\rho_{y^f z^m} - \rho_{z^f y^f} \rho_{z^f z^m}) \\ \frac{\sigma_{y^m}}{\sigma_{z^f}} (\rho_{z^f y^m} - \rho_{z^f y^f} \rho_{y^f y^m}) & \frac{\sigma_{y^m}}{\sigma_{y^f}} (\rho_{y^f y^m} - \rho_{z^f y^f} \rho_{z^f y^m}) \end{pmatrix} \tag{16}$$

and

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

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

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

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

and 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 $\rho_{y^m y^f}$.

Notice that the variance matrix of $(w^f_{t-1}, \varepsilon^f_{t-1})$ is given by

$$Var\begin{pmatrix} w^f_{t-1} \\ \varepsilon^f_{t-1} \end{pmatrix} = Var\begin{pmatrix} z^m_{t-1} \\ y^m_{t-1} \end{pmatrix} - \begin{pmatrix} r^f_{zz} & r^f_{zy} \\ r^f_{yz} & r^f_{yy} \end{pmatrix} Var\begin{pmatrix} z^f_{t-1} \\ y^f_{t-1} \end{pmatrix}\begin{pmatrix} r^f_{zz} & r^f_{zy} \\ r^f_{yz} & r^f_{yy} \end{pmatrix}'$$

We use these matching functions to write years of schooling, y^k_t, and social status, z^k_t, as a function of mother's years of schooling, y^f_{t-1}, and social status z^f_{t-1}. We write (14) as

$$\begin{aligned} z^k_t &= \gamma^k\left(\alpha^k_z z^m_{t-1} + (1-\alpha^k_z)z^f_{t-1}\right) + e^k_t \\ &= \gamma^k\left(\alpha^k_z\left(r^f_{zz}z^f_{t-1} + r^f_{zy}y^f_{t-1} + w^f_{t-1}\right) + (1-\alpha^k_z)z^f_{t-1}\right) + e^k_t \\ &= G^k_{zf}z^f_{t-1} + G^k_{yf}y^f_{t-1} + g^k_f\omega^f_{t-1} + e^k_t \end{aligned}$$

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 (13) as

$$\begin{aligned} y^k_t &= \beta^k\left(\alpha^k_y y^m_{t-1} + (1-\alpha^k_y)y^f_{t-1}\right) + z^k_t + x^k_t + u^k_t \\ &= \beta^k\left(\alpha^k_y\left(r^f_{yz}z^f_{t-1} + r^f_{yy}y^f_{t-1} + \varepsilon^f_{t-1}\right) + (1-\alpha^k_y)y^f_{t-1}\right) + z^k_t + x^k_t + u^k_t \\ &= \beta^k\left(\alpha^k_y\left(r^f_{yz}z^f_{t-1} + r^f_{yy}y^f_{t-1} + \varepsilon^f_{t-1}\right) + (1-\alpha^k_y)y^f_{t-1}\right) \\ & \quad G^k_{zf}z^f_{t-1} + G^k_{yf}y^f_{t-1} + g^k_f\omega^f_{t-1} + e^k_t + x^k_t + u^k_t \end{aligned}$$

$$y^k_t = B^k_{yf}y^f_{t-1} + B^k_{zf}z^f_{t-1} + b^k_f\varepsilon^f_{t-1} + g^k_f\omega^f_{t-1} + e^k_t + x^k_t + u^k_t$$

where

$$\begin{aligned} B^k_{yf} &= \beta^k\left(\alpha^k_y r^f_{yy} + (1-\alpha^k_y)\right) + G^k_{yf} \\ B^k_{zf} &= \beta^k\alpha^k_y r^f_{yz} + G^k_{zf} \\ b^k_f &= \beta^k\alpha^k_y \end{aligned}$$

Notice that ω^f_{t-1} and ε^f_{t-1} are related to ω^m_{t-1} and ε^m_{t-1}. We can write

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

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

$$= \left[I_2 - \begin{pmatrix} r_{zz}^f & r_{zy}^f \\ r_{yz}^f & r_{yy}^f \end{pmatrix} \begin{pmatrix} r_{zz}^m & r_{zy}^m \\ r_{yz}^m & r_{yy}^m \end{pmatrix} \right] \begin{pmatrix} z_{t-1}^m \\ y_{t-1}^m \end{pmatrix} + \begin{pmatrix} r_{zz}^f & r_{zy}^f \\ r_{yz}^f & r_{yy}^f \end{pmatrix} \begin{pmatrix} w_{t-1}^m \\ \varepsilon_{t-1}^m \end{pmatrix}$$

We assume that the second order moments of all variables are time invariant. We present below the formulas for the covariances in years of scholing for relatives of different degrees of kinship. The correlations are computed by dividing the covariances by σ_m^2, σ_f^2 or $\sigma_m \sigma_f$ depending on the gender.

We first compute the covariances between y_t^m and z_t^m

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

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

and

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

In the steady state we have $Cov(y_t^m, z_t^m) = Cov(y_{t-1}^m, z_{t-1}^m)$, then this equation implicitely defines $Cov(y_t^m, z_t^m)$ since r_{yy}^m, G_{ym}^m, r_{yz}^m and G_{zm}^m depend on $Cov(y_t^m, z_t^m)$.

Analogously we compute the covariances between y_t^f and z_t^f

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

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

and

$$
\begin{aligned}
Cov(y_t^f, z_t^f) &= \beta^f \left(\alpha_y^f r_{yy}^f + (1-\alpha_y^f)\right) G_{yf}^f \sigma_{y^f}^2 \\
&+\beta^f \left(\left(\alpha_y^f r_{yy}^f + (1-\alpha_y^f)\right) G_{zf}^f + \alpha_y^f r_{yz}^f G_{yf}^f\right) Cov(y_{t-1}^f, z_{t-1}^f) \\
&+(1 + b_f^f r_{yz}^f G_{zf}^f)\sigma_{z^f}^2 + b_f^f g_f^f Cov(\varepsilon_{t-1}^f, w_{t-1}^f)
\end{aligned}
$$

In the steady state we have $Cov(y_t^f, z_t^f) = Cov(y_{t-1}^f, z_{t-1}^f)$, then this equation implicitely defines $Cov(y_t^f, z_t^f)$ since r_{yy}^f, G_{yf}^f, r_{yz}^f and G_{zf}^f depend on $Cov(y_t^f, z_t^f)$.

Covariances

Husband and wife

$$
Cov_{h-w}(y_{t-1}^m, y_{t-1}^f) = Cov_{hw}(y_{t-1}^m, r_{yz}^m z_{t-1}^m + r_{yy}^m y_{t-1}^m + \varepsilon_{t-1}^m) = r_{yz}^m Cov(y_{t-1}^m, z_{t-1}^m) + r_{yy}^m \sigma_{y^m}^2
$$

Brothers

$$
\begin{aligned}
&Cov_b(y_t^{m,i}, y_t^{m,j}) = Cov(B_{ym}^m y_{t-1}^m + B_{zm}^m z_{t-1}^m + b_m^m \varepsilon_{t-1}^m + g_m^m \omega_{t-1}^m + x_t^m, B_{ym}^m y_{t-1}^m + B_{zm}^m z_{t-1}^m + b_m^m \varepsilon_{t-1}^m + g_m^m \omega_{t-1}^m + x_t^m) \\
&= \left(B_{ym}^m\right)^2 \sigma_{y^m}^2 + \left(B_{zm}^m\right)^2 \sigma_{z^m}^2 + 2B_{ym}^m B_{zm}^m Cov(y_{t-1}^m, z_{t-1}^m) + \left(b_m^m\right)^2 \sigma_{\varepsilon^m}^2 \\
&+ \left(g_m^m\right)^2 \sigma_{w^m}^2 + \sigma_{x^m}^2 + 2b_m^m g_m^m Cov(\varepsilon_{t-1}^m, \omega_{t-1}^m)
\end{aligned}
$$

alternatively we cam write it as a function of the mother

$$
\begin{aligned}
&Cov_b(y_t^{m,i}, y_t^{m,j}) = Cov(B_{yf}^m y_{t-1}^f + B_{zf}^m z_{t-1}^f + b_f^m \varepsilon_{t-1}^f + g_f^m \omega_{t-1}^f + x_t^m, B_{yf}^m y_{t-1}^f + B_{zf}^m z_{t-1}^f + b_f^m \varepsilon_{t-1}^f + g_f^m \omega_{t-1}^f + x_t^m) \\
&= \left(B_{yf}^m\right)^2 \sigma_{y^f}^2 + \left(B_{zf}^m\right)^2 \sigma_{z^f}^2 + 2B_{yf}^m B_{zf}^m Cov(y_{t-1}^f, z_{t-1}^f) + \left(b_f^m\right)^2 \sigma_{\varepsilon^f}^2 \\
&+ \left(g_f^m\right)^2 \sigma_{w^f}^2 + \sigma_{x^m}^2 + 2b_f^m g_f^m Cov(\varepsilon_{t-1}^f, \omega_{t-1}^f)
\end{aligned}
$$

Sisters

$$
\begin{aligned}
&Cov_s(y_t^{f,i}, y_t^{f,j}) = Cov(B_{ym}^f y_{t-1}^m + B_{zm}^f z_{t-1}^m + b_m^f \varepsilon_{t-1}^m + g_m^f \omega_{t-1}^m + x_t^f, B_{ym}^f y_{t-1}^m + B_{zm}^f z_{t-1}^m + b_m^f \varepsilon_{t-1}^m + g_m^f \omega_{t-1}^m + x_t^f) \\
&= \left(B_{ym}^f\right)^2 \sigma_{y^m}^2 + \left(B_{zm}^f\right)^2 \sigma_{z^m}^2 + 2B_{ym}^f B_{zm}^f Cov(y_{t-1}^m, z_{t-1}^m) + \left(b_m^f\right)^2 \sigma_{\varepsilon^m}^2 \\
&+ \left(g_m^f\right)^2 \sigma_{w^m}^2 + \sigma_{x^f}^2 + 2b_m^f g_m^f Cov(\varepsilon_{t-1}^m, \omega_{t-1}^m)
\end{aligned}
$$

alternatively we cam write it as a function of the mother

$$
\begin{aligned}
&Cov_b(y_t^{f,i}, y_t^{f,j}) = Cov(B_{yf}^f y_{t-1}^f + B_{zf}^f z_{t-1}^f + b_f^f \varepsilon_{t-1}^f + g_f^f \omega_{t-1}^f + x_t^f, B_{yf}^f y_{t-1}^f + B_{zf}^f z_{t-1}^f + b_f^f \varepsilon_{t-1}^f + g_f^f \omega_{t-1}^f + x_t^f) \\
&= \left(B_{yf}^f\right)^2 \sigma_{y^f}^2 + \left(B_{zf}^f\right)^2 \sigma_{z^f}^2 + 2B_{yf}^f B_{zf}^f Cov(y_{t-1}^f, z_{t-1}^f) + \left(b_f^f\right)^2 \sigma_{\varepsilon^f}^2 \\
&+ \left(g_f^f\right)^2 \sigma_{w^f}^2 + \sigma_{x^f}^2 + 2b_f^f g_f^f Cov(\varepsilon_{t-1}^f, \omega_{t-1}^f)
\end{aligned}
$$

Brother-sister

$$
\begin{aligned}
&Cov_{b-s}(y_t^{m,i}, y_t^{f,j}) = Cov(B_{ym}^m y_{t-1}^m + B_{zm}^m z_{t-1}^m + b_m^m \varepsilon_{t-1}^m + g_m^m \omega_{t-1}^m + x_t^m, B_{ym}^f y_{t-1}^m + B_{zm}^f z_{t-1}^m + b_m^f \varepsilon_{t-1}^m + g_m^f \omega_{t-1}^m + x_t^f) \\
&= B_{ym}^m B_{ym}^f \sigma_{y^m}^2 + B_{zm}^m B_{zm}^f \sigma_{z^m}^2 + \left(B_{ym}^m B_{zm}^f + B_{ym}^f B_{zm}^m\right) Cov(y_{t-1}^m, z_{t-1}^m) + b_m^m b_m^f \sigma_{\varepsilon^m}^2 \\
&+ g_m^m g_m^f \sigma_{w^m}^2 + \sigma_{x^m, x^f} + \left(b_m^m g_m^f + b_m^f g_m^m\right) Cov(\varepsilon_{t-1}^m, \omega_{t-1}^m)
\end{aligned}
$$

alternatively we cam write it as a function of the mother

$$Cov_{b-s}(y_t^{m,i}, y_t^{f,j}) = Cov(B_{yf}^m y_{t-1}^f + B_{zf}^m z_{t-1}^f + b_f^m \varepsilon_{t-1}^f + g_f^m \omega_{t-1}^f + x_t^m, B_{yf}^f y_{t-1}^f + B_{zf}^f z_{t-1}^f + b_f^f \varepsilon_{t-1}^f + g_f^f \omega_{t-1}^f + x_t^f)$$
$$= B_{yf}^m B_{yf}^f \sigma_{y^f}^2 + B_{zf}^m B_{zf}^f \sigma_{z^f}^2 + \left(B_{yf}^m B_{zf}^f + B_{yf}^f B_{zf}^m\right) Cov(y_{t-1}^f, z_{t-1}^f) + b_f^m b_f^f \sigma_{\varepsilon^f}^2$$
$$+ g_f^m g_f^f \sigma_{w^f}^2 + \sigma_{x^m, x^f} + \left(b_f^m g_f^f + b_f^f g_f^m\right) Cov(\varepsilon_{t-1}^f, \omega_{t-1}^f)$$

Male cousins (fathers are brothers)

We first compute the following covariances for their fathers (who are brothers)

$$\begin{aligned} Cov_b(z_{t-1}^{m,i}, z_{t-1}^{m,j}) &= Cov(G_{zm}^m z_{t-2}^m + G_{ym}^m y_{t-2}^m + g_m^m \omega_{t-2}^m, G_{zm}^m z_{t-2}^m + G_{ym}^m y_{t-2}^m + g_m^m \omega_{t-2}^m) \\ &= (G_{zm}^m)^2 \sigma_{z^m}^2 + \left(G_{ym}^m\right)^2 \sigma_{y^m}^2 + 2G_{zm}^m G_{ym}^m Cov(y_{t-2}^m, z_{t-2}^m) + (g_m^m)^2 \sigma_{w^m}^2 \end{aligned}$$

and

$$\begin{aligned} Cov_b(y_{t-1}^{m,i}, z_{t-1}^{m,j}) &= Cov(B_{ym}^m y_{t-2}^m + B_{zm}^m z_{t-2}^m + b_m^m \varepsilon_{t-2}^m + g_m^m \omega_{t-2}^m, G_{zm}^m z_{t-2}^m + G_{ym}^m y_{t-2}^m + g_m^m \omega_{t-2}^m) \\ &= \left(B_{ym}^m G_{zm}^m + B_{zm}^m G_{ym}^m\right) Cov(y_{t-2}^m, z_{t-2}^m) + B_{ym}^m G_{ym}^m \sigma_{y^m}^2 \\ &\quad + B_{zm}^m G_{zm}^m \sigma_{z^m}^2 + b_m^m g_m^m Cov(\varepsilon_{t-2}^m, w_{t-2}^m) + (g_m^m)^2 \sigma_{w^m}^2 \end{aligned}$$

The covariance for male cousins whose fathers are brothers is

$$\begin{aligned} Cov_{mc\_fb}(y_t^{m,i}, y_t^{m,j}) &= Cov(B_{ym}^m y_{t-1}^{m,i} + B_{zm}^m z_{t-1}^{m,i}, B_{ym}^m y_{t-1}^{m,j} + B_{zm}^m z_{t-1}^{m,j}) \\ &= (B_{ym}^m)^2 Cov_b(y_{t-1}^{m,i}, y_{t-1}^{m,j}) + 2B_{ym}^m B_{zm}^m Cov_b(y_{t-1}^{m,i}, z_{t-1}^{m,j}) \\ &\quad + (B_{zm}^m)^2 Cov_b(z_{t-1}^{m,i}, z_{t-1}^{m,j}) \end{aligned}$$

Male cousins (mothers are sisters)

We first compute the following covariances for their mothers (who are sisters)

$$\begin{aligned} Cov_s(z_{t-1}^{f,i}, z_{t-1}^{f,j}) &= Cov(G_{zm}^f z_{t-2}^m + G_{ym}^f y_{t-2}^m + g_m^f \omega_{t-2}^m, G_{zm}^f z_{t-2}^m + G_{ym}^f y_{t-2}^m + g_m^f \omega_{t-2}^m) \\ &= (G_{zm}^f)^2 \sigma_{z^m}^2 + \left(G_{ym}^f\right)^2 \sigma_{y^m}^2 + 2G_{zm}^f G_{ym}^f Cov(y_{t-2}^m, z_{t-2}^m) + (g_m^f)^2 \sigma_{w^m}^2 \end{aligned}$$

and

$$\begin{aligned} Cov_s(y_{t-1}^{f,i}, z_{t-1}^{f,j}) &= Cov(B_{ym}^f y_{t-2}^m + B_{zm}^f z_{t-2}^m + b_m^f \varepsilon_{t-2}^m + g_m^f \omega_{t-2}^m, G_{zm}^f z_{t-2}^m + G_{ym}^f y_{t-2}^m + g_m^f \omega_{t-2}^m) \\ &= \left(B_{ym}^f G_{zm}^f + B_{zm}^f G_{ym}^f\right) Cov(y_{t-2}^m, z_{t-2}^m) + B_{ym}^f G_{ym}^f \sigma_{y^m}^2 \\ &\quad + B_{zm}^f G_{zm}^f \sigma_{z^m}^2 + b_m^f g_m^f Cov(\varepsilon_{t-2}^m, w_{t-2}^m) + (g_m^f)^2 \sigma_{w^m}^2 \end{aligned}$$

The covariance for male cousins whose mothers are sisters is

$$\begin{aligned} Cov_{mc\_ms}(y_t^{m,i}, y_t^{m,j}) &= \left(B_{yf}^m\right)^2 Cov_s(y_{t-1}^{f,i}, y_{t-1}^{f,j}) + 2B_{yf}^m B_{zf}^m Cov_s(y_{t-1}^{f,i}, z_{t-1}^{f,j}) \\ &\quad + \left(B_{zf}^m\right)^2 Cov_s(z_{t-1}^{f,i}, z_{t-1}^{f,j}) \end{aligned}$$

Male cousins (father and mother are brother and sister)

We first compute the following covariances for their father and mother (who are brother and sister)

$$\begin{aligned} Cov_{b-s}(z_{t-1}^{m,i}, z_{t-1}^{f,j}) &= Cov(G_{zm}^{m} z_{t-2}^{m} + G_{ym}^{m} y_{t-2}^{m} + g_{m}^{m}\omega_{t-2}^{m}, G_{zm}^{f} z_{t-2}^{m} + G_{ym}^{f} y_{t-2}^{m} + g_{m}^{f}\omega_{t-2}^{m}) \\ &= G_{zm}^{m} G_{zm}^{f} \sigma_{z^m}^2 + G_{ym}^{m} G_{ym}^{f} \sigma_{y^m}^2 + (G_{zm}^{m} G_{ym}^{f} + G_{zm}^{f} G_{ym}^{m}) Cov(y_{t-2}^{m}, z_{t-2}^{m}) + g_{m}^{m} g_{m}^{f} \sigma_{w^m}^2 \end{aligned}$$

and

$$\begin{aligned} Cov_{b-s}(y_{t-1}^{m,i}, z_{t-1}^{f,j}) &= Cov(B_{ym}^{m} y_{t-2}^{m} + B_{zm}^{m} z_{t-2}^{m} + b_{m}^{m}\varepsilon_{t-2}^{m} + g_{m}^{m}\omega_{t-2}^{m}, G_{zm}^{f} z_{t-2}^{m} + G_{ym}^{f} y_{t-2}^{m} + g_{m}^{f}\omega_{t-2}^{m}) \\ &= \left(B_{ym}^{m} G_{zm}^{f} + B_{zm}^{m} G_{ym}^{f}\right) Cov(y_{t-2}^{m}, z_{t-2}^{m}) + B_{ym}^{m} G_{ym}^{f} \sigma_{y^m}^2 \\ &\quad + B_{zm}^{m} G_{zm}^{f} \sigma_{z^m}^2 + b_{m}^{m} g_{m}^{f} Cov(\varepsilon_{t-2}^{m}, w_{t-2}^{m}) + g_{m}^{m} g_{m}^{f} \sigma_{w^m}^2 \end{aligned}$$

$$\begin{aligned} Cov_{b-s}(y_{t-1}^{f,i}, z_{t-1}^{m,j}) &= Cov(B_{ym}^{f} y_{t-2}^{m} + B_{zm}^{f} z_{t-2}^{m} + b_{m}^{f}\varepsilon_{t-2}^{m} + g_{m}^{f}\omega_{t-2}^{m}, G_{zm}^{m} z_{t-2}^{m} + G_{ym}^{m} y_{t-2}^{m} + g_{m}^{m}\omega_{t-2}^{m}) \\ &= \left(B_{ym}^{f} G_{zm}^{m} + B_{zm}^{f} G_{ym}^{m}\right) Cov(y_{t-2}^{m}, z_{t-2}^{m}) + B_{ym}^{f} G_{ym}^{m} \sigma_{y^m}^2 \\ &\quad + B_{zm}^{f} G_{zm}^{m} \sigma_{z^m}^2 + b_{m}^{f} g_{m}^{m} Cov(\varepsilon_{t-2}^{m}, w_{t-2}^{m}) + g_{m}^{f} g_{m}^{m} \sigma_{w^m}^2 \end{aligned}$$

The covariance for male cousins whose father and mother are brother and sister is

$$\begin{aligned} Cov_{mc\_fb-ms}(y_t^{m,i}, y_t^{m,j}) &= Cov(B_{ym}^{m} y_{t-1}^{m,i} + B_{zm}^{m} z_{t-1}^{m,i}, B_{yf}^{m} y_{t-1}^{f,j} + B_{zf}^{m} z_{t-1}^{f,j}) \\ &= B_{ym}^{m} B_{yf}^{m} Cov_{b-s}(y_{t-1}^{m,i}, y_{t-1}^{f,j}) + B_{ym}^{m} B_{zf}^{m} Cov_{b-s}(y_{t-1}^{m,i}, z_{t-1}^{f,j}) \\ &\quad + B_{yf}^{m} B_{zm}^{m} Cov_{b-s}(y_{t-1}^{f,i}, z_{t-1}^{m,j}) + B_{zm}^{m} B_{zf}^{m} Cov_{b-s}(z_{t-1}^{m,i}, z_{t-1}^{f,j}) \end{aligned}$$

Female cousins (fathers are brothers)

$$\begin{aligned} Cov_{fc\_fb}(y_t^{f,i}, y_t^{f,j}) &= (B_{ym}^{f})^2 Cov_b(y_{t-1}^{m,i}, y_{t-1}^{m,j}) + 2 B_{ym}^{f} B_{zm}^{f} Cov_b(y_{t-1}^{m,i}, z_{t-1}^{m,j}) \\ &\quad + (B_{zm}^{f})^2 Cov_b(z_{t-1}^{m,i}, z_{t-1}^{m,j}) \end{aligned}$$

Female cousins (mothers are sisters)

$$\begin{aligned} Cov_{fc\_ms}(y_t^{f,i}, y_t^{f,j}) &= \left(B_{yf}^{f}\right)^2 Cov_s(y_{t-1}^{f,i}, y_{t-1}^{f,j}) + 2 B_{yf}^{f} B_{zf}^{f} Cov_s(y_{t-1}^{f,i}, z_{t-1}^{f,j}) \\ &\quad + \left(B_{zf}^{f}\right)^2 Cov_s(z_{t-1}^{f,i}, z_{t-1}^{f,j}) \end{aligned}$$

Female cousins (father and mother are brother and sister)

$$\begin{aligned} Cov_{fc\_fb-ms}(y_t^{f,i}, y_t^{f,j}) &= B_{ym}^{f} B_{yf}^{f} Cov_{b-s}(y_{t-1}^{m,i}, y_{t-1}^{f,j}) + B_{ym}^{f} B_{zf}^{f} Cov_{b-s}(y_{t-1}^{m,i}, z_{t-1}^{f,j}) \\ &\quad + B_{yf}^{f} B_{zm}^{f} Cov_{b-s}(y_{t-1}^{f,i}, z_{t-1}^{m,j}) + B_{zm}^{f} B_{zf}^{f} Cov_{b-s}(z_{t-1}^{m,i}, z_{t-1}^{f,j}) \end{aligned}$$

Male-female cousins (fathers are brothers)

$$\begin{aligned} Cov_{m-fc\_fb}(y_t^{m,i}, y_t^{f,j}) &= B_{ym}^{m} B_{ym}^{f} Cov_b(y_{t-1}^{m,i}, y_{t-1}^{m,j}) + \left(B_{ym}^{m} B_{zm}^{f} + B_{ym}^{f} B_{zm}^{m}\right) Cov_b(y_{t-1}^{m,i}, z_{t-1}^{m,j}) \\ &\quad + B_{zm}^{m} B_{zm}^{f} Cov_b(z_{t-1}^{m,i}, z_{t-1}^{m,j}) \end{aligned}$$

Male-female cousins (mothers are sisters)

$$
\begin{aligned}
Cov_{m-fc\_ms}(y_t^{m,i}, y_t^{f,j}) &= B_{yf}^m B_{yf}^f Cov_s(y_{t-1}^{f,i}, y_{t-1}^{f,j}) + \left(B_{yf}^m B_{zf}^f + B_{yf}^f B_{zf}^m\right) Cov_s(y_{t-1}^{f,i}, z_{t-1}^{f,j}) \\
&\quad + B_{zf}^m B_{zf}^f Cov_s(z_{t-1}^{f,i}, z_{t-1}^{f,j})
\end{aligned}
$$

Male-female cousins (father of the male is brother of the mother of the female)

$$
\begin{aligned}
Cov_{m-fc\_fb-ms}(y_t^{m,i}, y_t^{f,j}) &= B_{ym}^m B_{yf}^f Cov_{b-s}(y_{t-1}^{m,i}, y_{t-1}^{f,j}) + B_{ym}^m B_{zf}^f Cov_{b-s}(y_{t-1}^{m,i}, z_{t-1}^{f,j}) \\
&\quad + B_{yf}^f B_{zm}^m Cov_{b-s}(y_{t-1}^{f,i}, z_{t-1}^{m,j}) + B_{zm}^m B_{zf}^f Cov_{b-s}(z_{t-1}^{m,i}, z_{t-1}^{f,j})
\end{aligned}
$$

Male-female cousins (mother of the male is sister of the father of the female)

$$
\begin{aligned}
Cov_{m-fc\_ms-fb}(y_t^{m,i}, y_t^{f,j}) &= B_{yf}^m B_{ym}^f Cov_{b-s}(y_{t-1}^{m,i}, y_{t-1}^{f,j}) + B_{ym}^f B_{zf}^m Cov_{b-s}(y_{t-1}^{m,i}, z_{t-1}^{f,j}) \\
&\quad + B_{yf}^m B_{zm}^f Cov_{b-s}(y_{t-1}^{f,i}, z_{t-1}^{m,j}) + B_{zf}^m B_{zm}^f Cov_{b-s}(z_{t-1}^{m,i}, z_{t-1}^{f,j})
\end{aligned}
$$

Son-Father

$$
\begin{aligned}
Cov_{sf}(y_t^m, y_{t-1}^m) &= Cov(B_{ym}^m y_{t-1}^m + B_{zm}^m z_{t-1}^m + b_m^m \varepsilon_{t-1}^m + g_m^m \omega_{t-1}^m, y_{t-1}^m) \\
&= B_{ym}^m \sigma_{y^m}^2 + B_{zm}^m Cov(y_{t-1}^m, z_{t-1}^m)
\end{aligned}
$$

alternatively we can write it as a function of the mother

$$
\begin{aligned}
Cov_{sf}(y_t^m, y_{t-1}^m) &= Cov(B_{yf}^m y_{t-1}^f + B_{zf}^m z_{t-1}^f + b_f^m \varepsilon_{t-1}^f + g_f^m \omega_{t-1}^f, r_{yz}^f z_{t-1}^f + r_{yy}^f y_{t-1}^f + \varepsilon_{t-1}^f) \\
&= B_{yf}^m r_{yy}^f \sigma_{y^f}^2 + (B_{yf}^m r_{yz}^f + B_{zf}^m r_{yy}^f) Cov(y_{t-1}^f, z_{t-1}^f) \\
&\quad + B_{zf}^m r_{yz}^f \sigma_{z^f}^2 + b_f^m \sigma_{\varepsilon^f}^2 + g_f^m Cov(\omega_{t-1}^f, \varepsilon_{t-1}^f)
\end{aligned}
$$

Son-Mother

$$
\begin{aligned}
Cov_{sm}(y_t^m, y_{t-1}^f) &= Cov(B_{yf}^m y_{t-1}^f + B_{zf}^m z_{t-1}^f + b_f^m \varepsilon_{t-1}^f + g_f^m \omega_{t-1}^f, y_{t-1}^f) \\
&= B_{yf}^m \sigma_{y^f}^2 + B_{zf}^m Cov(y_{t-1}^f, z_{t-1}^f)
\end{aligned}
$$

alternatively we can write it as a function of the father

$$
\begin{aligned}
Cov_{sm}(y_t^m, y_{t-1}^f) &= Cov(B_{ym}^m y_{t-1}^m + B_{zm}^m z_{t-1}^m + b_m^m \varepsilon_{t-1}^m + g_m^m \omega_{t-1}^m, r_{yz}^m z_{t-1}^m + r_{yy}^m y_{t-1}^m + \varepsilon_{t-1}^m) \\
&= B_{ym}^m r_{yy}^m \sigma_{y^m}^2 + (B_{ym}^m r_{yz}^m + B_{zm}^m r_{yy}^m) Cov(y_{t-1}^m, z_{t-1}^m) \\
&\quad + B_{zm}^m r_{yz}^m \sigma_{z^m}^2 + b_m^m \sigma_{\varepsilon^m}^2 + g_m^m Cov(\omega_{t-1}^m, \varepsilon_{t-1}^m)
\end{aligned}
$$

Daughter-Father

$$
\begin{aligned}
Cov_{df}(y_t^f, y_{t-1}^m) &= Cov(B_{ym}^f y_{t-1}^m + B_{zm}^f z_{t-1}^m + b_m^f \varepsilon_{t-1}^m + g_m^f \omega_{t-1}^m, y_{t-1}^m) \\
&= B_{ym}^f \sigma_{y^m}^2 + B_{zm}^f Cov(y_{t-1}^m, z_{t-1}^m)
\end{aligned}
$$

alternatively we can write it as a function of the mother

$$\begin{aligned}
Cov_{df}(y_t^f, y_{t-1}^m) &= Cov(B_{yf}^f y_{t-1}^f + B_{zf}^f z_{t-1}^f + b_f^f \varepsilon_{t-1}^f + g_f^f \omega_{t-1}^f, r_{yz}^f z_{t-1}^f + r_{yy}^f y_{t-1}^f + \varepsilon_{t-1}^f) \\
&= B_{yf}^f r_{yy}^f \sigma_{y^f}^2 + (B_{yf}^f r_{yz}^f + B_{zf}^f r_{yy}^f) Cov(y_{t-1}^f, z_{t-1}^f) \\
&\quad + B_{zf}^f r_{yz}^f \sigma_{z^f}^2 + b_f^f \sigma_{\varepsilon^f}^2 + g_f^f Cov(\omega_{t-1}^f, \varepsilon_{t-1}^f)
\end{aligned}$$

Daughter-Mother

$$\begin{aligned}
Cov_{dm}(y_t^f, y_{t-1}^f) &= Cov(B_{yf}^f y_{t-1}^f + B_{zf}^f z_{t-1}^f + b_f^f \varepsilon_{t-1}^f + g_f^f \omega_{t-1}^f, y_{t-1}^f) \\
&= B_{yf}^f \sigma_{y^f}^2 + B_{zf}^f Cov(y_{t-1}^f, z_{t-1}^f)
\end{aligned}$$

alternatively we can write it as a function of the father

$$\begin{aligned}
Cov_{dm}(y_t^f, y_{t-1}^f) &= Cov(B_{ym}^f y_{t-1}^m + B_{zm}^f z_{t-1}^m + b_m^f \varepsilon_{t-1}^m + g_m^f \omega_{t-1}^m, r_{yz}^m z_{t-1}^m + r_{yy}^m y_{t-1}^m + \varepsilon_{t-1}^m) \\
&= B_{ym}^f r_{yy}^m \sigma_{y^m}^2 + (B_{ym}^f r_{yz}^m + B_{zm}^f r_{yy}^m) Cov(y_{t-1}^m, z_{t-1}^m) \\
&\quad + B_{zm}^f r_{yz}^m \sigma_{z^m}^2 + b_m^f \sigma_{\varepsilon^m}^2 + g_m^f Cov(\omega_{t-1}^m, \varepsilon_{t-1}^m)
\end{aligned}$$

Nephew and uncle (brother of the father)

$$Cov_{neph-u\_bf}(y_t^{m,i}, y_{t-1}^{m,j}) = B_{ym}^m Cov_b(y_{t-1}^{m,i}, y_{t-1}^{m,j}) + B_{zm}^m Cov_b(y_{t-1}^{m,i}, z_{t-1}^{m,j})$$

Nephew and uncle (brother of the mother)

$$Cov_{neph-u\_bm}(y_t^{m,i}, y_{t-1}^{m,j}) = B_{yf}^m Cov_{b-s}(y_{t-1}^{f,i}, y_{t-1}^{m,j}) + B_{zf}^m Cov_{b-s}(y_{t-1}^{m,i}, z_{t-1}^{f,j})$$

Nephew and aunt (sister of the father)

$$Cov_{neph-au\_sf}(y_t^{m,i}, y_{t-1}^{f,j}) = B_{ym}^m Cov_{b-s}(y_{t-1}^{m,i}, y_{t-1}^{f,j}) + B_{zm}^m Cov_{b-s}(y_{t-1}^{f,i}, z_{t-1}^{m,j})$$

Nephew and aunt (sister of the mother)

$$Cov_{neph-au\_sm}(y_t^{m,i}, y_{t-1}^{f,j}) = B_{yf}^m Cov_s(y_{t-1}^{f,i}, y_{t-1}^{f,j}) + B_{zf}^m Cov_s(y_{t-1}^{f,i}, z_{t-1}^{f,j})$$

Nice and uncle (brother of the father)

$$Cov_{nice-u\_bf}(y_t^{f,i}, y_{t-1}^{m,j}) = B_{ym}^f Cov_b(y_{t-1}^{m,i}, y_{t-1}^{m,j}) + B_{zm}^f Cov_b(y_{t-1}^{m,i}, z_{t-1}^{m,j})$$

Nice and uncle (brother of the mother)

$$Cov_{nice-u\_bm}(y_t^{f,i}, y_{t-1}^{m,j}) = B_{yf}^f Cov_{b-s}(y_{t-1}^{f,i}, y_{t-1}^{m,j}) + B_{zf}^f Cov_{b-s}(y_{t-1}^{m,i}, z_{t-1}^{f,j})$$

Nice and aunt (sister of the father)

$$Cov_{nice-au\_sf}(y_t^{f,i}, y_{t-1}^{f,j}) = B_{ym}^f Cov_{b-s}(y_{t-1}^{m,i}, y_{t-1}^{f,j}) + B_{zm}^f Cov_{b-s}(y_{t-1}^{f,i}, z_{t-1}^{m,j})$$

Nice and aunt (sister of the mother)

$$Cov_{nice-au\_sm}(y_t^{f,i}, y_{t-1}^{f,j}) = B_{yf}^f Cov_s(y_{t-1}^{f,i}, y_{t-1}^{f,j}) + B_{zf}^f Cov_{b-s}(y_{t-1}^{f,i}, z_{t-1}^{f,j})$$

Grandson-Grandfather (father-line)

We have that

$$\begin{aligned} Cov_{sf}(z^m_{t-1}, y^m_{t-2}) &= Cov(G^m_{zm} z^m_{t-2} + G^m_{ym} y^m_{t-2} + g^m_m \omega^m_{t-2}, y^m_{t-2}) \\ &= G^m_{zm} Cov(z^m_{t-2}, y^m_{t-2}) + G^m_{ym} \sigma^2_{y^m} \end{aligned}$$

and

$$\begin{aligned} Cov_{gs-gf}(y^m_t, y^m_{t-2}) &= Cov(B^m_{ym} y^m_{t-1} + B^m_{zm} z^m_{t-1} + b^m_m \varepsilon^m_{t-1} + g^m_m \omega^m_{t-1}, y^m_{t-2}) \\ &= B^m_{ym} Cov_{sf}(y^m_{t-1}, y^m_{t-2}) + B^m_{zm} Cov_{sf}(z^m_{t-1}, y^m_{t-2}) \end{aligned}$$

Great-grandson-Great-grandfather (father-line)
We have that

$$\begin{aligned} Cov_{gs-gf}(z^m_{t-1}, y^m_{t-3}) &= Cov(G^m_{zm} z^m_{t-2} + G^m_{ym} y^m_{t-2} + g^m_m \omega^m_{t-2}, y^m_{t-3}) \\ &= G^m_{zm} Cov_{sf}(z^m_{t-2}, y^m_{t-3}) + G^m_{ym} Cov_{sf}(y^m_{t-2}, y^m_{t-3}) \end{aligned}$$

and

$$\begin{aligned} Cov_{ggs-ggf}(y^m_t, y^m_{t-3}) &= Cov(B^m_{ym} y^m_{t-1} + B^m_{zm} z^m_{t-1} + b^m_m \varepsilon^m_{t-1} + g^m_m \omega^m_{t-1}, y^m_{t-3}) \\ &= B^m_{ym} Cov_{gs-gf}(y^m_{t-1}, y^m_{t-3}) + B^m_{zm} Cov_{gs-gf}(z^m_{t-1}, y^m_{t-3}) \end{aligned}$$

Great-great-grandson-Great-great-grandfather (father-line)
We have that

$$\begin{aligned} Cov_{ggs-ggf}(z^m_{t-1}, y^m_{t-4}) &= Cov(G^m_{zm} z^m_{t-2} + G^m_{ym} y^m_{t-2} + g^m_m \omega^m_{t-2}, y^m_{t-4}) \\ &= G^m_{zm} Cov_{gsgf}(z^m_{t-2}, y^m_{t-4}) + G^m_{ym} Cov_{gsgf}(y^m_{t-2}, y^m_{t-4}) \end{aligned}$$

and

$$\begin{aligned} Cov_{gggs-gggf}(y^m_t, y^m_{t-4}) &= Cov(B^m_{ym} y^m_{t-1} + B^m_{zm} z^m_{t-1} + b^m_m \varepsilon^m_{t-1} + g^m_m \omega^m_{t-1}, y^m_{t-4}) \\ &= B^m_{ym} Cov_{ggs-ggf}(y^m_{t-1}, y^m_{t-3}) + B^m_{zm} Cov_{ggs-ggf}(z^m_{t-1}, y^m_{t-4}) \end{aligned}$$

Variance decomposition

We have that

$$y^k_t = \beta^k \tilde{y}^k_{t-1} + z^k_t + x^k_t + u^k_t$$

Then

$$\sigma^2_{y^k} = (\beta^k)^2 \sigma^2_{\tilde{y}^k} + \sigma^2_{z^k} + \beta^k Cov(\tilde{y}^k_{t-1}, z^k_t) + \sigma^2_{x^k} + \sigma^2_{u^k}$$

- $\sigma^2_{u^k}$ is obtained as a residual.
- $\beta^k, \sigma^2_{x^k}$ and $\sigma^2_{z^k}$ are directly estimated.
- $\sigma^2_{\tilde{y}^k}$

$$\sigma^2_{\tilde{y}^k} = \left(\alpha^k_y\right)^2 \sigma^2_{y^m} + \left(1 - \alpha^k_y\right)^2 \sigma^2_{y^f} + \alpha^k_y (1 - \alpha^k_y) \rho_{y^m y^f} \sigma_{y^m} \sigma_{y^f}$$

and we use the estimates of α^k_y and the empirical values for σ_{y^m}, σ_{y^f} and $\rho_{y^m y^f}$

- $Cov(\widetilde{y}^k_{t-1}, z^k_t)$

$$\begin{aligned} Cov(\widetilde{y}^k_{t-1}, z^k_t) &= \left(\alpha^k_y + (1-\alpha^k_y) r^k_{yy}\right) G^k_{ym} \sigma^2_{y^m} \\ &+ \left(\left[\alpha^k_y + (1-\alpha^k_y) r^k_{yy}\right] G^k_{zm} + (1-\alpha^k_y) r^k_{yz} G^k_{ym}\right) Cov(y^m_{t-1}, z^m_{t-1}) \\ &+ (1-\alpha^k_y) r^k_{yz} G^k_{zm} \sigma^2_{z^m} + (1-\alpha^k_y) g^k_m Cov(\varepsilon^k_{t-1}, w^k_{t-1}) \end{aligned}$$

and we use the estimates of $\alpha^k_y, r^k_{yy}, r^k_{yz}, G^k_{ym}, G^k_{zm}, g^k_m, \sigma^2_{z^m}, Cov(\varepsilon^m_t, w^m_t)$ and $Cov(y^m_t, z^m_t)$.