Parental Assortative Mating and the Intergenerational Transmission of Human Capital*

Paul Bingley
Danish Center for Social Science Research (VIVE)

Lorenzo Cappellari
Università Cattolica Milano

Konstantinos Tatsiramos
University of Luxembourg and LISER

May 2020
Incomplete draft

Abstract
We study the contribution of parental similarity in schooling levels to the intergenerational transmission of educational attainment. We develop an empirical model for educational correlations within the family in which parental sorting only partly translates into intergenerational transmission, while transmission can either originate from both parents in conjunction and from each parent independently. We identify the model by exploiting variation in the gender mix of siblings Using parameter estimates we decompose parental assortative mating into its intra- and inter- generational parts. We show that about half of the variation in human capital generating factors that parents have in common translates into intergenerational persistence. We also show that the independent contribution of each parent to intergenerational transmission is negligible, with the bulk of transmission originating from both parents in conjunction. Similar results hold when we consider off-springs permanent incomes in place of their educational attainment. These results point towards strong complementarities of parental inputs in the production function of children human capital. Looking over time, we document a sizeable secular decline of parental assortative mating and we show that it mainly occurs through human capital generating factors that are not passed-on to the next generation, leaving intergenerational persistence unaffected.

Keywords: Assortative mating, intergenerational transmission
JEL Codes: I24, J62

*We gratefully acknowledge funding from the Danish Council for Independent Research (grant DFF-6109-00226) and the Università Cattolica D3.2 Strategic Research Grant "Evidence-based anti poverty policies".

1. Introduction

Understanding the mechanisms of intergenerational transmission is key for the design of policies aimed at promoting equal opportunities. Parental assortative mating is typically considered as one of the main channels though which intergenerational transmission takes place. According to this view, assortative mating is a sorting mechanism that generates segmentation in the distribution of parental human capital, with some households ending up with high levels of human capital of both parents, and other households being characterised by low levels of human capital for both parents (Becker 1973). Given that parental human capital is the engine of intergenerational transmission of advantage, parental sorting can act as an amplifier of intergenerational persistence.

There is a growing literature in economics on the relationship between assortative mating and intergenerational transmission. Most studies, however, concentrate on educational sorting in the off-springs generation, with the view that non-random sorting at destination slows down regression to the mean in the intergenerational transition (Chadwick and Solon, 2002; Ermisch et al., 2006; Guell et al. 2015; Holmlund, 2019). Conversely, very little is known about marital sorting in the parental generation, which can inform on the degree of complementarity of parental inputs in the human capital production function of their children. Recently, Handy (2016) estimates that in the US parental sorting only accounts for a quarter of the intergenerational elasticity of education, suggesting that most of intergenerational persistence takes the form of a one-to-one parent-child relationship that is independent of the other parent, consistent with substitutability of parental inputs.[1]

In this paper we contribute to the literature by developing an empirical model of educational correlations within the family that allows us estimating how much of intergenerational transmission comes from parental assortative mating. In our model, parental human capital has a common factor among the spouses (assortative mating) which may or may not be transferred to children. In addition, each parent may transmit human capital to the children independently from the spouse. Therefore,

[1] Bratsberg et al. (2018) consider the effect of assortative mating on the mean and variance of outcomes in the offspring generation, rather than the intergenerational transmission of outcomes.

intergenerational persistence can arise both from the joint contribution of the parents and from the independent contribution of each parent. We estimate the model using educational attainment of individuals within families reconstructed from the Danish population register, and we show that identification can be achieved by assuming that the sibling correlation stems entirely from intergenerational persistence. This assumption is corroborated by out-of-sample predictions of the sibling correlation generated by the model.

We show that about half of assortative mating translates into intergenerational transmission, while remaining parental similarities are not transmitted. In turn, about 80 percent of the intergenerational correlation comes from the joint contribution of the two parents, while independent transmission from each parent plays only minor role. This result points toward strong complementarities of parental inputs into the human capital production function of the children. When looking over time, we show that parental assortative mating has been declining since the 1960s, similarly to what has been happening in other Scandinavian countries and in the US (see. Bratsberg et al., 2018, Holmlund, 2019, and Eika et al., 2019), while intergenerational persistence has not. Using the model, we show that the decline of parental assortative mating has occurred only through the intragenerational component, while common parental traits that are transmitted intergenerationally have not changed much. This implies a change in the nature of assortative mating, that is, parental matches are increasingly formed over dimensions of human capital that are passed onto the next generation.

Our results on the relevance of parental sorting in shaping intergenerational correlations somehow differ from the ones of Handy (2016) showing a limited role of parental educational correlations in explaining intergenerational persistence in the US. One key difference with our paper is that the empirical model of Handy (2016) implies an upper bound of 0.5 on the share of intergenerational transmission that can be imputed to parental sorting, whereas in our model that share is unrestricted. More generally, within the literature on intergenerational mobility, our paper contributes to the body of studies that jointly consider horizontal (intragenerational) and vertical (intergenerational) family ties. Collado et al. (2019) exploit extended horizontal families including kinships up to the fifth degree of separation to infer the characteristics of multi-generational transmission, pointing toward a predominant role of assortative mating especially in latent traits.[2] Our contribution in this context is to provide a framework for analysing the relationship between parental sorting and intergenerational persistence within nuclear families without the need of data on extended kinships.

Outside the intergenerational literature, there is growing interest in understanding the contribution of assortative mating to cross-sectional income inequality. Eika et al. (2019) show that in the US (and, with varying degrees, in other countries) assortative mating has been following heterogeneous trends depending on educational levels. Particularly, while the overall trend is mildly declining, looking at the bottom of the educational distribution shows sharp increases. Counterfactual decompositions of the income distribution show that these trends of assortative mating impact little on overall income inequality.[3] Relative to this literature, our contributions is to show that aggregate trends of assortative mating may mask changes in its nature, and that its contribution to long-run inequality may be persistent even in the face of an aggregate decline.

The rest of the paper is structured as follows. In the next section we describe the data source and provide some raw statistics on the extent of intra-family correlations in education. In Section 3 we outline the econometric model of educational correlations of family members, while Section 4 presents the estimation results. Section 5 contains a concluding discussion.

---

[2] Bingley and Cappellari (2019) exploit father-first son- second son triads to show that most of the sibling correlation of long run earnings and other outcomes has an intergenerational origin.
[3] A similarly limited effect of assortative mating on income inequality is reported for France by Fremeaux and Lefranc (2015).

2. Data description and raw patterns of educational correlations within the family

(A description of the registers and of the two key outcome variables: years of education and earnings)

We base our analysis on *family quartets* composed by two parents and their first two children. In principle data on siblings are not needed to estimate parental assortative mating or intergenerational correlations, but we show in the next section that observing two children per family is key for identifying the share of intergenerational correlation due to parental assortative mating.

We draw data on the full population of Danish families with parents born from 1920 (the first cohorts for which we have educational information) to 1979 with at least two children, a total of 639,516 families. We exclude families in which either parent was younger than 18 at the first birth, resulting in a sample of 624,883 families. Next, we select families whose first child is born in 1956 or later, that is the first year for which we can match children to parents (while children are still in compulsory school), and whose second child is born in 1988 or earlier, which is the last year for which we observe educational attainment at age 30 (our data were sourced in 2018). Selecting on first and second child year of birth returns a sample of 530,390 families. Next we drop families with sibling age spacing smaller than one year or larger than ten years, corresponding to the first and last percentile of the distribution of age spacing, resulting in a sample of 524,577 families. Finally, we drop families with first child born after 1978 to avoid artificially compressing sibling spacing on younger families, resulting in 477,953 families. We further drop families where at least one member has missing information on educational attainment, resulting in an estimating sample of 460,962.

Besides educational outcomes, we are also interested in studying permanent incomes, which we can investigate by drawing information from tax records, more specifically labour income that is available for the years 1980-2018. Following many papers in the literature, we define permanent income as average log income in the 30-40 age bracket. To provide a meaningful proxy of permanent income, we require individuals to have at least 5 valid observations on income in the 30-40 age bracket. Averaging multiple observations on current incomes reduces the impact of transitory fluctuations, while considering the 30-40 age range limits life-cycle bias.. Using income data in the

30-40 age range precludes the possibility of deriving permanent incomes for the vast majority of parents' birth cohorts. We therefore focus on the intergenerational link between parental education and off-springs permanent incomes. Moreover, we exclude from estimation families whose children are too young to reach age 40 in the available income data, that is we exclude families whose second child is born after 1978. To allow each family to have a maximum sibling spacing of 10 years, we also exclude families whose first child is born after 1968. In this way we obtain a sample of 276,057 families, which is further reduced to 219,443 families after dropping families where at least one child has fewer than 5 valid observations on income in the 30-40 age bracket.[4] Table 1 describes key features of the estimating samples. In the larger sample with 460,962 families and 1,843,848 individuals, the average year of birth is around 1940 for fathers and 1943 for mothers, while it is around 1967 for first born children and 1970 for second born children. The restricted sample where we have children's permanent income (219,443 families and 877,772 individuals) is approximately four years older as a consequence of imposing that the younger child is at least 40 years old as of 2018 (the last year of income data) to calculate her permanent income. There is no big difference in educational attainment in the two samples, with fathers attaining on average 11 years, mothers 10 years and each child 13 years.

In Table 2 we provide a summary of raw intra-family correlations in years of education and permanent income. We residualize years of education on year of birth dummies and log income on year by year of birth dummies to remove common trends in education or income that may induce spurious patterns of correlation. Correlations are estimated by birth cohorts after grouping individuals in 3 years brackets by year of births, and we exclude from estimation cells with fewer than 100 cases. Estimating correlations by cohorts will enable us considering secular changes. We also compute bootstrapped standard errors of these correlation coefficients, that will be used later in the paper to adjust inference.

[4] Estimating the educational model on the restricted sample produces results that are very similar to the ones obtained in the larger sample.

Table 2 reports in Panel A educational correlation coefficients averaged over cohorts and splitted by family types, that we define according to the gender mix of the two children. The educational correlation among spouses is substantial, around 0.47 and stable across family types. This is our basic raw measure of parental assortative mating, and its estimate is in line with previous research. Intergenerational correlations are substantially lower at about 0.27, and do not vary much both across family types when mothers are considered, while there is some evidence that intergenerational transmission of education from fathers is stronger in all-boys families. Comparing these figures with the ones for parental assortative mating suggests that not all factors that make spouses similar among themselves are transmitted to the children, because otherwise intergenerational and spousal correlations would mirror each other. The one correlation that is more sensitive to the type of family is the sibling correlation. In particular, it is larger in families with same gender children that in those with mixed gender children, possibly reflecting a greater exposure to shared community effects for siblings that are of the same gender. Panel B of Table 2 repeats the exercise but using permanent incomes as children outcomes, and computing correlations on the restricted sample. We find larger educational correlation among spouses, which, given that this sample is older, reflects a declining secular trend in assortative mating the we document in subsequent paragraphs. Considering incomes as children outcome there is more evidence of a gendered intergenerational transmission. Also, there is a larger gap in siblings correlations between same-gender and mixed gender siblings.

In Figure 1 we take advantage of the cohort-wise structure of the estimated raw correlations and plot them against the cohort of birth. We perform this exercise pooling correlations across family types. Figure 1 shows that parental assortative mating in human capital has been steadily declining for the cohorts of parents born between the mid-1920s and mid-1950s. This pattern is consistent with findings from the literature (see e.g. Eika et al. 2019; Bratsberg et al 2018; Holmlund 2019) that typically point to a secular trend of falling assortativeness in spouses' outcomes in a variety of countries, which is generally connected with educational expansions that affected those cohorts.

There is also a declining trend of intergenerational transmission, but the pace of decline is much lower compared with assortative mating: while the parental correlation is twice as large as the parent-child correlation for parents born in the mid-1920s, for parents born 30 years later the two correlations are of about the same size. That declining assortative mating of the parents is only mildly reflected in intergenerational mobility is consistent with two alternative scenarios. It may be that the *joint contribution* of parents to the human capital of the children has been declining but this decline has been compensated by the *individual contribution* of each parent and her ability to transmit human capital to the children. For example, parenting style might have been shifting to a one-to-one parent-child relationship. Alternatively, the decline of assortative mating reflects a decline on commonality of traits that are not relevant for intergenerational transmission. This would point to a more efficient matching process in the marriage market, in which traits that are uninfluential for the human capital of children are less and less valued. Disentangling between the two scenarios is not possible based only on the contrast between raw parental assortative mating and intergenerational transmission, and the model that we present in the next Section offers a way for identifying the relevant explanation.

3. An econometric model of educational correlations among family members

In this section we present a model for the educational attainment of family members which enables us disentangling the contribution of parental assortative mating to the intergenerational transmission of human capital, while allowing each parent to also affect the human capital of the off-springs independently from his or her spouse.[5] Families are composed of four members: father (*F*), mother (*M*), first child (C_1) and second child (C_2). Let y_{ij} denote the years of education for person *i* in family *j*, and let $H(i)$ denote the role of i in the family, that is $H(i) \in \{M, F, C_1, C_2\}$. If *i* is a parent, we write her years of education as:

[5] We formulate the model referring to educational attainment as the outcome throughout, but we also apply it to other relevant economic outcomes such as the permanent incomes of the children.

$$y_{ij} = a_{ij} + \gamma_{Aj} + \gamma_{H(i)j} + \mu_j, \quad H(i) = F, M \qquad (1)$$

In (1) educational attainment is factored into the sum of orthogonal components. a_{ij} represents sources of human capital that are specific of person i and are not in common with other family members, like idiosyncratic ability or luck. The γ terms denote human capital-generating factors that are passed across generations. We distinguish two such factors. The first, γ_A is shared between parents and represents the contribution of assortative mating to intergenerational transmission. This term captures the existence of intergenerationally transferable skills that are relevant for the acquisition of education and that are in common between the parents. For example, both parents are proficient in maths and jointly transfer their proficiency to the off-springs. Another example is human capital generating genes, as long as there exists assortative mating on genes. The second factor of intergenerational transmission, $\gamma_{H(i)}$ is instead parent-specific and is not shared among the spouses. This is what each parent transmits to children independently from the spouse. For example, only one parent is proficient in foreign languages, so any transmission of that particular skill to the off-springs will only come from this parent. We may also think of it as capturing the genetic component of human capital that is not shared among the parents. Finally, μ_j represents those determinants of human capital that are shared among the parents but are not passed on to their children, i.e. a purely intragenerational component of parental assortative mating. For example, parents may share some human capital-hampering trait (say smoking) that they do not want to pass to their children.

In the children generation, human capital is determined by idiosyncratic factors, by all factors that are received from the prior generation, and by other intra-generational influences that siblings share independently from the parents, such as school or community effects orthogonal to the family; therefore, when i is a child, education is given by the following sum of orthogonal components:

$$y_{ij} = a_{ij} + \gamma_{Aj} + \gamma_{Mj} + \gamma_{Fj} + \theta_j, \quad H(i) = C_1, C_2 \qquad (2)$$

By assuming orthogonality across the factors of Equation (2), and in particular between each of the γ terms and θ, we are effectively ruling out the sorting of families across schools and communities. Bingley, Cappellari and Tatsiramos (2020) show how sorting can be identified in models of the sibling correlation by exploiting the timing of families mobility across communities; their results indicate that ignoring sorting leads to overestimate community-related effects. Therefore, our results about the relevance of intra-generational influences within the sibling correlation will have to be seen as upper bounds.

We standardize educational attainment to have zero mean and unit variance, such as estimated variance components are interpretable as (components of) correlation coefficients, and derive moment restrictions for the correlation structure of the outcomes across family members, that we will map into empirical moments for estimating the parameters. Starting from the parents, their correlation of educational attainment is our measure of parental assortative mating and is given by:

$$\rho_{MF} = \sigma_{\gamma A}^2 + \sigma_{\mu}^2 \qquad (3)$$

that is, the sum of the variances (denoted with σ^2) of the two factors generating assortative mating, those that are transmitted to the children ($\sigma_{\gamma A}^2$), and this that are not (σ_{μ}^2); instead, parent-specific determinants of human capital (idiosyncratic or intergenerationally transmissible) do not contribute to assortative mating.

Looking across generations, the intergenerational correlation of education is given by:

$$\rho_{PC_k} = \sigma_{\gamma A}^2 + \sigma_{\gamma P}^2, \quad k = 1,2; \; P = M, F \qquad (4)$$

The sibling correlation, based on the model assumptions, is given by:

$$\rho_{C_1 C_2} = \sigma^2_{\gamma A} + \sigma^2_{\gamma M} + \sigma^2_{\gamma F} + \sigma^2_{\theta} \qquad (5)$$

The expression in (5) shows that the sibling correlation depends on *total* intergenerational transmission, i.e. what a child receives from *both* parents, and on the intragenerational factors that siblings share independently from the parents.

Note that despite this being a system of six equations (two intragenerational equations –(3) and (5) —and four intergenerational equations –(4)) in five parameters ($\sigma^2_{\gamma A}, \sigma^2_{\gamma M}, \sigma^2_{\gamma F}, \sigma^2_{\mu}, \sigma^2_{\theta}$), it is not identified because the intergenerational equations of each parent do not vary with birth order, eliminating two degrees of freedom.

We therefore need additional information for identification. To this end, let's consider differencing outcomes between parents and children. For example, the difference between a mother (denoted *i*) and one of her children (denoted *i*') is given by:

$$y_{\tilde{M}j} \equiv y_{ij} - y_{i'j} = a_{ij} - a_{i'j} - \gamma_{Fj} + \mu_j - \theta_j \qquad (6)$$

that is, such difference will depend on idiosyncratic factors, the father-specific contribution to intergenerational transmission and the intragenerational terms μ and θ. A similar result will hold for the difference in outcome between the father (that we denote *l*) and the other child of the family (*l*'):

$$y_{\tilde{F}j} \equiv y_{lj} - y_{l'j} = a_{lj} - a_{l'j} - \gamma_{Mj} + \mu_j - \theta_j \quad . \qquad (6')$$

Crucially, the idiosyncratic factors entering (6') are different from the ones entering (6), and each parent-specific contribution to intergenerational transmission is orthogonal to the spouse's one, such as the correlation between parental outcomes taken in difference from children outcomes identifies the sum of intragenerational terms:

$$\rho_{\tilde{M}\tilde{F}} = \sigma_{\mu}^{2} + \sigma_{\theta}^{2} \qquad (7)$$

Taking differences of parents and offsprings outcomes is useful as it adds one moment condition. This shows the value of having data on two children even if siblings correlations are not the main interest of the analysis: both children are required for differencing out parent data because each parent need to be related to a different child in order to eliminate all idiosyncratic components when taking second moments.

The inclusion of Equation (7) still leaves the system under identified with five equations and six parameters. To solve the indeterminacy, we exploit the gender composition of the siblings and make the assumption that for mixed-gender sibling couples there is no shared influence on top of what is received from the parents, that is we assume

$$\theta_j = 0 \ \ if\ g_{ij} \neq g_{i'j};\ \ H(i) = C_1,\ H(i') = \ C_2. \qquad (8).$$

where *g* is a gender indicator.

The rationale for this assumption is that, particularly at young ages, interactions independent from the family environment occur predominantly among same-gender peers (COULD WE QUOTE SOME PAPER HERE?). The assumption is untestable, but the evidence of lower educational correlations for mixed gender siblings compared with same gender ones shown in Section 2 provides some corroboration, suggesting that our estimated correlation in siblings shared influences for same gender siblings provide will provide an upper bound on the values of he parameter for mixed gender siblings, helping assessing the plausibility of assumption (8). Using assumption (8), Equation (7) identifies the intragenerational component of parental assortative mating (σ_{μ}^{2}) and solves the identification problem.

We estimate the model by Minimum Distance and match empirical educational correlations with their counterparts predicted by the model on the basis of model parameters, using boostrapped standard errors of the correlations to weight the minimisation problem. Our baseline estimations consider years of education as the outcome of interest for all family members; we also estimate a version of the model in which offspring years of education are substituted with their permanent incomes. Let $k \in \{M, F, C_1, C_2, \widetilde{M}, \tilde{F}\}$ be the index for the outcomes of family members, including parent-child differences of outcomes, and let $r_{kk'}$ be the correlation of the outcomes for members k and k'. The Minimum Distance estimator is the weighted least squares estimator of the following regression

$$
\begin{aligned}
r_{kk'} = \sigma_{\gamma A}^{2}[I(k=M)I(k'=F) + I(k=P)I(k'=C) + I(k=C_1)I(k'=C_2)] \\
+ \sigma_{\mu}^{2}\left[I(k=M)I(k'=F) + I\big(k=\widetilde{M}\big)I\big(k'=\tilde{F}\big)\right] \\
+ \sigma_{\gamma M}^{2}[I(k=M)I(k'=C) + I(k=C_1)I(k'=C_2)] \\
+ \sigma_{\gamma F}^{2}[I(k=F)I(k'=C) + I(k=C_1)I(k'=C_2)] \\
+ \sigma_{\theta}^{2}\left[I(k=C_1)I(k'=C_2)I(g_{C_1}=g_{C_2}) + I\big(k=\widetilde{M}\big)I\big(k'=\tilde{F}\big)\right] + u_{kk'}
\end{aligned}
$$

where P=M, F, C=C1, C2 and the weights are given by $(var(r_{kk'}))^{-1}$.

4. Results

We report Minimum Distance estimates for the model of educational correlations in Table 3. Estimated parameters are interpretable as components of correlation coefficients. In Column (1) we report estimates for the baseline specification laid out in the previous section. Results show that, on average across birth cohorts and family types, the most of the factors that parents share are transmitted intergenerationally, but there is as well a non-negligible component of parental assortative mating which remains within the generation of the parents, the ratio between non-transmitted factors and total parental assortative mating being equal to about 43 percent. Parental specific contributions to intergenerartional transmission appear of a modest size, about a tenth of the joint parental contribution, and without much difference between mothers and fathers. Thus, about 80 percent of intergenerational correlation in education comes from the joint contribution of the parents, pointing towards strong complementarities of parental inputs in producing the human capital of children. Finally, sibling correlations independent of the parents appear limited, accounting for about 8 percent of the overall sibling correlation. All these estimates are derived under the assumption that the independent sibling correlation is null for mixed gender sibling, and the estimates from same gender sibling (that can be thought of as upper bounds for the mixed gender one) suggest that e assumption of zero independent correlation for mixed gender siblings is not too strong.[6]

Finding that independent parental contributions to intergenerational transmission are of second order compared with joint contributions may miss part of the parental effect as long as each parent's ability to influence children achievement depends on the family environment. For example, it may be that fathers have greater impacts on children outcomes for boys while the opposite holds for mothers. To investigate whether gender differences of siblings matter for the extent of each parent's independent influences, we extend the model by allowing all the intergenerational parameters to change with gender mix of the siblings. Results, reported in the second column of Table 3, show that indeed parent-child gender interaction maters for intergenerational transmission, each parent's independent contribution to intergenerational correlations being about a third larger when both siblings have the same gender of the parent. Still, this heterogeneity is not sufficient to modify the conclusion from

[6] To further corroborate the assumption, we re-estimated the model without using data on mixed gender siblings, and then we used parameter estimates to predict the overall sibling correlation for mixed gender siblings under the assumption of null independent correlation. We obtained an estimate of 0.30 (se0.0053) that is rather close to the raw correlation of Table 2 (0.31)

Column (1) that joint transmission from both parents is the dominant factor of intergenerational correlations of education.

Figure 1 showed a marked decline in parental assortative mating between the cohorts born in the mid-1920s and those born in the mid-1950s, while the contemporaneous evolution of intergenerational correlations was also declining, albeit at a more modest pace. There we argued that the heterogeneous trends in assortative mating and intergenerational transmission are compatible with two alternative scenarios, one in which there is a decline of joint intergenerational transmission from both parents compensated by an increase in parent-specific intergenerational transmission, and another where it is only the non-transmissible component of assortative mating that declines over time. We now exploit the cohort structure of the empirical moments to estimate these differential trends in the components of assortative mating and intergenerational transmission. Results are reported in Column (3) of Table 3. The estimates show that the sharp decline of assortative mating occurred mainly through the component that is not-transmitted intergenerationally, while the intergenerational component of assortative mating declined at a more modest pace. Trends of parent-specific components of intergenerational transmission are negligible, while there is a significant increasing trend of sibling correlations independent of the parents, which point towards an increased relevance of community factors that siblings share.

In Table 4 we report results obtained from the model in which children outcomes are given by their permanent incomes in the 30-40 age bracket, while maintaining years of schooling as the outcome for the parents. Results from the baseline model differ in various respects from their counterparts in the fully educational model. First of all, there is now a large gap between the intergenerational and intergenerational components of assortative mating, which is to be expected given that in this model with mixed outcomes the first component depends on interpersonal correlations on the same outcome (years of schooling between the parents) while the second stems from interpersonal correlations of non-homogeneous outcomes (years of schooling for the parents and permanent incomes for the children). Perhaps less expected is the change of balance between the three components of intergenerational transmission (all estimated based on non-homogeneous outcomes between parents and children). The evidence points to an equal contribution of each (joint, mother-specific, father-specific), which is in some contrast with the evidence of a predominance of joint transmission emerging from the purely educational model. This suggests a greater relevance of parent-specific factor in affecting children income compared with education. Results from other specifications using children permanent income as outcome tend to go in the directions already highlighted by the purely educational model, namely stronger parent-specific transmission when the parent and the children are of the same gender, and a marked declining secular trend of the non-transmitted component of assortative mating.

5. Discussion and conclusion

TO BE WRITTEN

References

Becker, Gary S., 1973, A Theory of Marriage: Part I, The Journal of Political Economy, Vol. 81, No. 4, pp. 813-846

Bratsberg Bernt, Simen Markussen, Oddbjørn Raaum, Knut Røed, Ole Røgeberg (2018), Trends in Assortative Mating and Offspring Outcomes, IZA DP No. 11753

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

Bingley, Paul, Lorenzo Cappellari and Konstantinos Tatsiramos, 2017. "Family, Community and Life-Cycle Earnings: Evidence from Siblings and Youth Peers," CESifo Working Paper Series 6743,

Chadwick, Laura, and Gary Solon. 2002. "Intergenerational Income Mobility Among Daughters ." American Economic Review, 92 (1): 335-344.

Collado M. Dolores, Ignacio Ortuno-Ortin and Jan Stuhler, 2019 , Estimating Intergenerational and Assortative Processes in Extended Family Data, mimeo

Eika Lasse, Magne Mogstad and Basit Zafar, 2019, Educational Assortative Mating and Household Income Inequality, Journal of Political Economy, vol. 127, no. 6

Ermisch, J., M. Francesconi, and T. Siedler (2006): "Intergenerational Mobility and Marital Sorting," The Economic Journal, 116(513), pp. 659679

Frémeaux, Nicolas and Arnaud Lefranc (2017): Assortative Mating and Earnings Inequality in France, IZA DP No. 11084

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

Handy, Christopher, 2016 "Assortative Mating and Intergenerational Persistence of Schooling and Earnings", paper presented at the SOLE Meetings

Holmlund Helena (2019) How much does marital sorting contribute to intergenerational socio-economic persistence? JHR

Figure 1: Parental assortative mating and intergenerational correlations in education

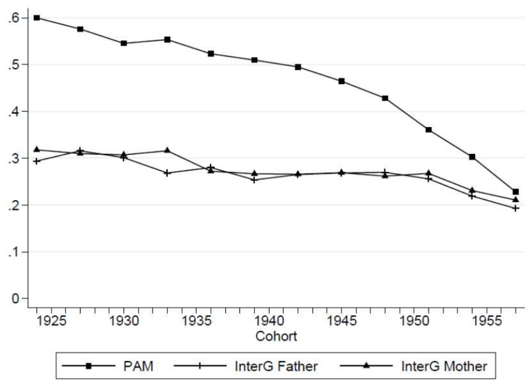

Table 1: Descriptive statistics

	Education sample (N=)				Permanent Income sample (N=)					
	Year of birth		Years of education		Year of birth		Years of education		Permanent Income	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
M	1942.99	7.12	10.90	3.17	1938.85	5.48	10.17	3.17		
F	1940.18	7.69	11.62	3.27	1935.84	6.11	11.18	3.33		
C1	1966.66	6.30	13.39	2.33	1962.31	3.69	13.17	2.27	12.51	0.66
C2	1970.07	6.58	13.36	2.30	1965.70	4.10	13.15	2.24	12.52	0.64

Table 2: Raw correlations in years of education among family members, by gender composition of the children

index	BB	GG	MX	Total
	Full sample			
PAM	0.47	0.47	0.47	0.47
SIBS	0.35	0.35	0.31	0.33
IGM	0.27	0.27	0.28	0.27
IGF	0.29	0.26	0.27	0.27
Total	0.32	0.31	0.32	0.32
	PI sapmle			
PAM	0.49	0.51	0.50	0.50
SIBS	0.16	0.15	0.10	0.13
IGM	0.08	0.09	0.09	0.09
IGF	0.09	0.08	0.09	0.09
Total	0.18	0.18	0.18	0.18

Table 3: Parameter estimates of education model (Number of persons, Number of moments)

igpam	0.2589	0.0038			0.2518	0.0052	0.2952	0.0057
igm	0.0232	0.0038			0.0268	0.0042	0.0163	0.0047
igf	0.0288	0.0038			0.0324	0.0042	0.0074	0.0051
rpam	0.1937	0.0032	0.1959	0.0034	0.1937	0.0032	0.3829	0.0050
rs	0.0282	0.0042	0.0239	0.0048	0.0282	0.0042	0.0207	0.0030
igpam_mx			0.2614	0.0045				
igpam_bb			0.2565	0.0062				
igpam_gg			0.2475	0.0065				
igm_mx			0.0204	0.0045				
igm_bb			0.0204	0.0068				
igm_gg			0.0414	0.0071				
igf_mx			0.0255	0.0045				
igf_bb			0.0443	0.0068				
igf_gg			0.0277	0.0071				
igpam1					0.0072	0.0035		
t_igpam							-0.0015	0.0003
t_igm							-0.0001	0.0002
t_igf							0.0004	0.0002
t_rpam							-0.0088	0.0002
t_rs							0.0024	0.0004

Table 4: Parameter estimates of education-permanent income model (Number of persons, Number of moments)

igpam	0.0468	0.0040			0.0356	0.0057	0.0566	0.0063
igm	0.0428	0.0042			0.0486	0.0047	0.0265	0.0053
igf	0.0425	0.0042			0.0480	0.0047	0.0385	0.0060
rpam	0.4477	0.0031	0.4500	0.0033	0.4477	0.0031	0.6186	0.0048
rs	0.0140	0.0043	0.0095	0.0046	0.0140	0.0043	0.0109	0.0029
igpam_mx			0.0486	0.0048				
igpam_bb			0.0457	0.0065				
igpam_gg			0.0334	0.0071				
igm_mx			0.0384	0.0052				
igm_bb			0.0454	0.0075				
igm_gg			0.0605	0.0079				
igf_mx			0.0376	0.0052				
igf_bb			0.0538	0.0075				
igf_gg			0.0506	0.0079				
igpam1					0.0112	0.0041		
t_rpam							-0.0097	0.0003
t_rs							0.0032	0.0008
t_igpam							-0.0009	0.0004
t_igm							0.0014	0.0003
t_igf							0.0006	0.0004