Intergenerational mobility or inequality: What does rank-mobility measure?

Mikkel Høst Gandil*

December 2019

Abstract

Rank-based measures of intergenerational mobility are generally justified by their invariance to changes in inequality. However, I show that whenever the source of inequality is uncorrelated to parent ranks, such as the cases of gender and birth-order, increasing equality leads to a fall in rank-mobility as measured by the rank-correlation. I develop an adjustment factor and show that US income mobility could have fallen by as much as 23 percent since 1970 due to increased gender equality. Without specifying a policy objective of interest, it is therefore unclear which conclusions to draw from differences in rank-based mobility across societies.

Keywords: Economic inequality, intergenerational mobility, mobility measurement, rank-correlation, gender inequality, wealth inequality

JEL Codes: C63,D31,E24,J16,J62

*Department of Economics, University of Oslo. Postboks 1095 Blindern 0317, Oslo, Norway. Telephone: +45 28736769. Mail: mikkel.gandil@econ.uio.no

This research was carried out at University of Copenhagen and University of Oslo with support from the Danish Innovation Fund and the Danish Council of the Labour Movement (Arbejderbevægelsens Erhvervsråd). I am grateful to Monique de Haan, Edwin Leuven, Paolo Giovanni Piacquadio, Christoph Schottmüller, Niels Johannesen, and Daniel Waldenström as well as participants at the Winter School on Inequality and Social Welfare Theory (2015) for suggestions and comments.

1 Introduction

Rank-based measures of intergenerational mobility have become popular in investigations of outcomes such as income, wealth, education, and health. One of the often-touted advantages of such measures is the ability to disentangle mobility from inequality. This in principle allows researchers to compare mobility across societies even when the societies differ in their levels of inequality.

In this paper, however, I show how intergenerational rank-mobility is increased by inequality which occurs due to random variation in attributes uncorrelated with parental heritage. The two prominent examples are gender and birth order. I illustrate the issue with the gender income gap and show that a society with high gender inequality will have *higher* rank-mobility (i.e. a lower rank-correlation between parent and child) than a society with low gender inequality. The intuition is as follows: If a society has no gender inequality, the gender of a child will not matter for the measure of mobility, ie. the association between the positions of parent and child in society. On the other hand, if a society has high gender inequality, the gender of the child becomes an important determinant of income. As the gender of a child is essentially random, and thus uncorrelated with parental rank, the sorting into genders will register as mobility from one generation to the next. Consequently, increasing gender equality leads to less mobility. The randomness of birth order presents similar issues; Societies in which bequests are shared equally among siblings will appear less mobile than societies where siblings inherit differently dependent on birth order. Therefore, it becomes unclear what to conclude from differences in rank-based mobility between societies when underlying inequality between genders or siblings cannot be held constant.

To asses the quantitative implications of changes in inequality on mobility, I develop an adjustment factor and calibrate it to US income data from 1970 to 2016. I show that, solely due to the narrowing gender income gap, the intergenerational rank-correlation based on individual incomes should have risen by 23 percent. In other words, without any change in relative mobility *within* genders, the aggregate mobility would appear to have fallen substantially due to increased gender equality. Based on this finding, a policy position could be that in order to enhance intergenerational mobility, societies should *increase* gender discrimination. This position is absurd and illustrates the issues that plague comparisons of rank-mobility whenever random inequality exists within families.

The present results hold for mobility concepts where the rank of a child is based on individual outcomes. Individual ranks are a natural choice for many outcomes such as education and health. However, a large part of the literature on income mobility has instead focused on the household as the unit of analysis, see Chetty et al. (2014a) for an example.[1] The concept of household income introduces the assortative matching of couples as an additional factor. Through simulations, I show that only in the case of perfect (and constant) assortative matching does the aggregate rank-mobility correspond to the within-gender mobility. Further, changes in gender equality over time will affect mobility measurements of sons and daughters differently. In other words, using households rather than individuals does not resolve the issue of how to interpret differences in rank-correlations, when the levels of inequality within families differ across societies.

The issue posed by inequality within families to the measurement of mobility comes down to a fundamental question: What is the relevant and correct comparison on which to judge whether a society is mobile? In the case of gender; is it fair to compare men and women, or should one compare men with men, and women with women? In the case of birth order: Should one classify discrimination based on the order of birth as a relevant source of mobility? There is no general an-

[1] Mobility measures are often calculated using solely sons. By ignoring women, gender inequality does not play a factor in these estimates. Chadwick et al. (2002) conjecture that the focus on sons might partly be due to unrealized sexism and the recognition that the entrance of women into the labor force represents one of the most salient and fundamental developments of the latter half of the twentieth century.

swer to these questions, but in the two cases above using only the outcomes of parents and children is not sufficient to answer the question of what constitutes a mobile society. If mobility measures are to inform policy, they should reflect the results of counterfactual policy choices. In other words, the measurement of mobility should relate to a clearly stated policy objective and the tools considered and not reflect statistical artifacts. To construct the relevant measure of mobility, one needs guidance from theory. Unfortunately, discussions of such factors are often missing from applied work on intergenerational mobility and, rather than forming relevant counterfactual outcomes, the availability of data seems a driving factor in deciding on the appropriate comparison groups.

Structure of the Paper Section 2 begins by informally introducing mobility concepts and the intuition for why increasing inequality may raise mobility. Section 3 presents the canonical conceptual framework for studying intergenerational mobility along with measures of mobility with a focus on the rank-correlation coefficient. I then introduce the issue of within-family inequality and show how the type of inequality affects the rank-correlation. The effect of gender inequality on the rank-correlation is quantified in section 4 and section 5 discusses possible generalizations of the result. Section 6 concludes.

2 Mobility and inequality

Most measures of mobility concern an association between parent and child outcomes, i.g. income, health or wealth. It is, however, well known that changes in inequality within generations directly affect such measures. This makes interpretation difficult; If two societies have different measures of mobility, how much of the difference can be attributed to different levels of inequality and how much to the transfer of privilege between generations? As an alternative, a branch of the mobility literature has moved towards a focus on *positions* in distributions, i.e. the ranks of the individuals. By construction, ranks are uniformly distributed.[2] In other words, regardless of the shape of distributions, the corresponding rank-distribution remains unchanged. The invariance of the rank-distribution makes it possible to separate out mobility from inequality by estimating the dependence structure between ranks rather than the actual outcomes. There exist several measures of mobility based on ranks, the most popular being transition matrices and the rank-correlation. The latter is simply the linear correlation coefficient of the parent and child rank. Because of the invariance of the distribution of ranks with respect to differences in inequality over time or between countries, the rank-correlation has been touted as superior over measured based on outcomes for measuring mobility in incomes (Chetty et al., 2014b) and has also increasingly been applied to the analysis of wealth mobility (Fagereng et al., 2018; Adermon et al., 2018) and health mobility (Wong et al., 2019; Fletcher and Jajtner, 2019).

Inequality reenters the picture The invariance of rank-based measures to inequality is however not robust. Some characteristics are attributed to individuals independently of parent rank. The two main examples are gender and birth order, which both have been shown to correlate with later income and wealth, see Blau and Kahn (1992) and Menchik (1980). This kind of randomness coupled with inequality will raise mobility (i.e. lower the rank-correlation). To illustrate, I construct an example with extreme gender income inequality. I draw pairs of parent and child rank for families with sons and daughters from the same distribution and plot the draws in Figures 1i and 1ii. By construction, parents and sons will have the same rank-correlation as parents and daughters. In other words, the within-gender rank-correlation is the same. However, due to gender inequal-

[2]Intuitively, when observations are ranked, they each receive a unique rank value. The density function will, therefore, be a constant for all values.

ity, a given rank will translate into lower incomes for women than for men. This process is illustrated in Figure 1iii where the income distributions of sons and daughters hardly overlap.

Normally, when researchers would estimate a rank-correlation they would do so across genders. As seen in Figure 1iv, this aggregate rank-correlation will be lower relative to the within-gender correlation in Figures 1i and 1ii. If gender inequality decreases over time so will measured mobility. Suppose the income distribution of daughters approaches that of sons as illustrated in Figure 1v. When gender inequality decreases, the randomness of gender becomes less consequential. Correspondingly, the link between the position of the parent to the child becomes more salient and the rank-correlation rises as illustrated in Figure 1vi. In the case of full gender equality the aggregate and within-gender rank-correlations will be equal.[3]

The implication of this example is that, while the rank-correlation may be invariant to changes in the overall shape of the income distribution, it is not invariant to changes in gender inequality. While the level of gender inequality in the example is extreme, the example shows a fundamental problem of using rank-correlations to compare mobility across societies or over time; More gender equality mechanically lowers intergenerational mobility. Correspondingly, a country with a large amount of gender discrimination will, all else equal, appear more mobile than a country with less discrimination.

A similar problem occurs with wealth mobility. Due to the randomness of birth order, if bequests are primarily given to the first-born child, the aggregate rank-correlation will be lower than if bequests are distributed evenly across siblings. Correspondingly, societies with primogeniture will appear mobile than countries with equal sharing of parental wealth among inheritors.

[3]Assuming that the within-gender rank-correlation is equal for sons and daughters.

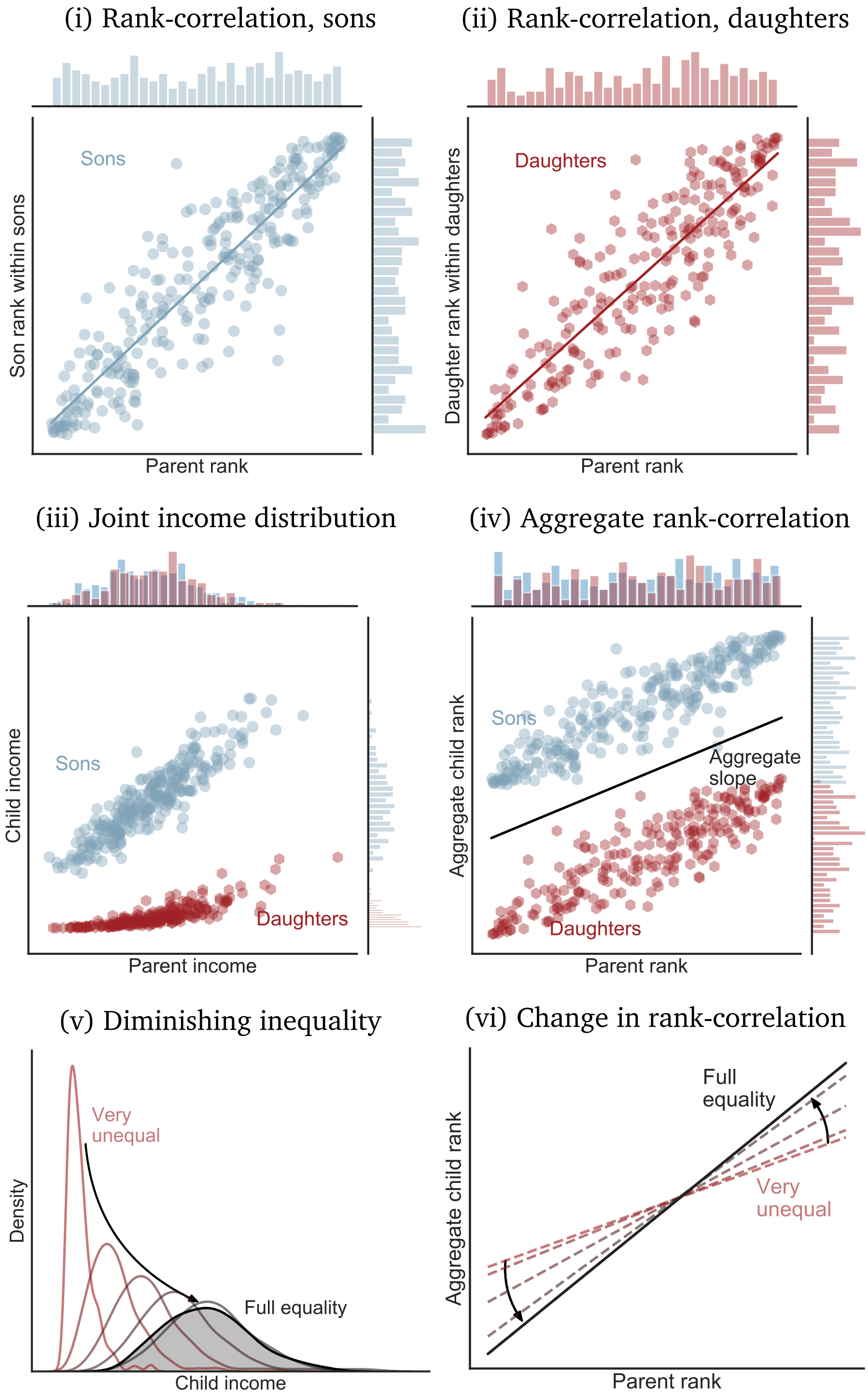


Figure 1: Example: From within to aggregate rank-correlations

Figures 1i and 1ii show a simulated association of parent and child ranks for sons and daughters respectively, drawn from the same distribution and therefore have equal within-gender rank-correlation. Figure 1iii shows the translation from rank to income. Due to gender inequality, daughters tend to rank lower than sons in the aggregate distribution as seen in Figure 1iv resulting in a lower aggregate rank-correlation. When gender inequality decreases as in Figure 1v the rank-correlation increases as seen by the the slope of the linear fit of parent-child ranks in Figure 1vi.

3 Derivation of the link between inequality and rank-correlations

While the previous section provides the intuition and the most important takeaways, a more formal structure is needed to prove the result. The basic framework to conceptualize inequality and mobility follows Jäntti and Jenkins (2015) closely.

Let a family be defined by a parent and a child outcome, X_i and Y_i respectively. Most measures of intergenerational mobility can be thought of as describing the joint distribution of (X_i, Y_i). Denote this bivariate joint distribution as $H(x, y)$ with the corresponding marginal distributions, $F(x)$ and $G(y)$. In this framework, the marginal distributions are the outcome distributions of the two generations and all the regular measures of inequality within generations, such as the Gini coefficient, can be calculated from F and G. IGE is calculated as the linear correlation of the (log-transformed) joint distribution, H.[4]

Denote the ranks of parent and child as U_i and V_i respectively. The rank-correlation is then given as:

$$\rho = \frac{Cov(U_i, V_i)}{Var(U_i)} = \frac{Cov(U_i, V_i)}{\sigma_U \sigma_V} \tag{1}$$

Assuming the outcome distributions are continuous, a rank can be calculated as the value of the cumulative distribution function evaluated at the level of outcome, $U_i = F(X_i)$ and $V_i = G(Y_i)$. A link between the intergenerational dependence structure of outcomes and ranks is provided by Sklar's theorem. The theorem states that any joint multivariate distribution can be described by the marginal distributions and a joint distribution describing the dependence of the ranks.[5] The latter distribution is called a copula and is defined on the unit square.

[4] This framework corresponds to Figure 1iii where $F(x)$ and $G(y)$ are represented by the horizontal and vertical histograms respectively.

[5] The theorem was introduced in Sklar (1959). For a proper proof see Nelsen (2006). Through-

Formally, $H(x,y) = C(F(x),G(y)) = C(u,v)$. The formulation of the copula shows the irrelevance of marginal distributions. Additionally, Sklar's theorem shows how to simulate a joint distribution by way of a copula and marginal distributions. I exploit this in the empirical investigation in section 4.

Introducing inequality within families In order to investigate the importance of inequality within families the conventional framework needs to be amended. In what follows I frame the issue in terms of gender inequality but birth order may anytime be substituted for gender.

In what follows, I assign the superscript m to families with a male child and f to families with a female child. I assume a data-generating process where families draw a parent and a child rank and a gender of the child. The gender is assigned by the random variable S_i, such that child i is male if $S_i = 1$. Gender is assumed to be orthogonal to parental outcomes.[6] The share of males in a cohort is given by $E[S_i] = \mu$ which I assume is stable at 0.5. Families are, therefore, defined by the tuple (U_i, V_i, S_i). The gender of a child assigns him or her to separate outcome distributions through the quantile functions, ${G^m}^{-1}(v)$ and ${G^f}^{-1}(v)$:

$$Y_i = S_i {G^m}^{-1}(V_i) + (1 - S_i){G^f}^{-1}(V_i) \tag{2}$$

In this simplified setup, gender is only important insofar as the gender-specific outcome distributions differ, $G^f \neq G^m$. The aggregate outcome distribution of the

out, I implicitly assume continuous monotonic outcome distributions. This assumption may be empirically problematic as a mass point at zero is common. I return to this point in section 4.

[6]There is evidence that the gender of a child might not be completely uncorrelated to parent outcomes, such as income, through differing mortality rates between male and female fetuses and the mother's circumstances and lifestyle. See Orzack et al. (2015) for an analysis of the prenatal gender ratio. Furthermore, there is evidence that the gender of the child might alter parent behavior and, therefore, maybe income. Lundberg et al. (2007) present evidence that father involvement and fragility of families may be affected by the gender of the child. Nonetheless, the assumption of orthogonality between gender and parent outcome is maintained throughout the rest of the analysis.

child generation is then given by

$$G(y) = \mu G^m(y) + (1-\mu)G^f(y) \tag{3}$$

Let V_i^A be the *aggregate* rank of the child, i.e. the outcome rank of the child when she is compared to every other child regardless of gender. Using (2) and (3), the *aggregate* rank can be expressed as a function of the *within gender* rank and the gender indicator:

$$\begin{aligned} V_i^A = \Lambda(V_i, S_i) =& \mu \left[G^m \left(S_i G^{m^{-1}}(V_i) + (1-S_i) G^{f^{-1}}(V_i) \right) \right] \\ &+ (1-\mu) \left[G^f \left(S_i G^{m^{-1}}(V_i) + (1-S_i) G^{f^{-1}}(V_i) \right) \right] \end{aligned} \tag{4}$$

Full gender equality implies that the aggregate rank is equal to the within-gender rank, $G^m = G^f \Rightarrow V_i = V_i^A$. However, whenever men and women have different outcome distributions the two rank can differ.

Because both V_i and V_i^A are uniformly distributed, the aggregate rank-correlation can be fully described by the covariance $Cov(U_i, V_i^A) = Cov(U_i, \Lambda(V_i, S_i))$. Due to the law of total covariance, the covariance between parent and child rank can be decomposed into a within-gender and a between-gender component:

$$\begin{aligned} Cov(U, V^a) =& Cov(U_i, \Lambda(V, S)) \\ =& \underbrace{E\left[Cov(U, \Lambda(V,S)|S)\right]}_{\text{within-group cov.}} + \underbrace{Cov\left(E\left[U|S\right], E\left[\Lambda(V,S)|S\right]\right)}_{\text{between-group cov.}=0} \end{aligned} \tag{5}$$

As gender (or birth order) by assumption is uncorrelated to parental outcomes, the last term on the right-hand side of Equation (5) equals zero. Accordingly, only the within-gender covariance matters in this case. Importantly, this is not the case for other types of group inequality, such as race, where children inherit group affiliation from their parents. In such cases, the between-group covariance will most likely be positive. The main conclusion that the aggregate rank-correlation will be lower for larger levels of inequality, therefore, hinges on the randomness *within families*. I return to this in Section 5.1.

To fully describe the functional form of $Cov(U, V^a)$ would require assuming specific distributions. Instead, I show in the appendix that the first-order Taylor-approximation of the covariance around a given rank, $\hat{v}$ can be expressed as:

$$Cov(U_i, V_i^A) \approx A(\hat{v}) \times Cov(U_i, V_i), \tag{6}$$

$$A(v) = \left\{ \frac{1}{2} + \frac{1}{4} \left[\lambda^m(v) + \lambda^f(v) \right] \right\} \tag{7}$$

$$\lambda^m(v) = \frac{g^f(G^{m-1}(v))}{g^m(G^{m-1}(v))}, \tag{8}$$

$$\lambda^f(v) = \frac{g^m(G^{f-1}(v))}{g^f(G^{f-1}(v))}, \tag{9}$$

where I assume equal arrival probability of sons and daughters.

The aggregate covariance is thus, to an approximation, a linear function of the withing-gender covariance, where the relationship is governed by an adjustment term, $A(v)$. The adjustment term is, in turn, a function of two likelihood ratios, λ^m and λ^f evaluated at different outcomes. For a given rank, v, $\lambda^m(v)$ is a likelihood ratio evaluated at the outcome corresponding to that rank in the male outcome distribution. Conversely, $\lambda^f(v)$ is the reciprocal likelihood ratio evaluated at the outcome corresponding to the rank in the female distribution.

As densities can only take positive values, it follows that the adjustment term is never negative. In other words, the total covariance will always have the same sign as the within-gender correlation. Under full gender equality it follows that $\lambda^m = \lambda^f = 1 \Rightarrow A(\hat{v}) = 1$; when there is no gender inequality, the within-gender covariance equals the total covariance. The magnitude of the adjustment term in the absence of gender equality remains to be determined.

Evaluation of adjustment factor In order to investigate the size of the adjustment term, I need to make assumptions concerning the outcome distributions. I assume that the likelihood ratio, $\frac{g^m(y)}{g^f(y)}$, increases monotonically. Intuitively, this means that the higher the level of the outcome, the larger is the ratio of men to women at that. It follows that the male distribution stochastically dominates the female distribution, $G^m(y) < G^f(y)$ and that densities cross only once. This is a natural assumption when the outcome is income.[7] The density functions are displayed in the middle of Figure 2.

An example of such a situation can be seen in Figure 2 with two artificial distributions: one for women (in red) and one for men (in blue). The top plot displays the two cumulative distribution functions (CDF). The distributions are chosen such that they have monotonic likelihood ratios.

Denote the point of crossing of the two densities by y^* where $g^m(y^*) = g^f(y^*)$. The single crossing implies the following:

$$\frac{g^m(y)}{g^f(y)} < 1 \text{ if } y < y^* \tag{10}$$

$$\frac{g^m(y)}{g^f(y)} > 1 \text{ if } y > y^* \tag{11}$$

In other words, the male density is lower than the female density when the outcome is below the crossing point and higher when the outcome is above. Using this observation we can bound the interval, where $\lambda^m(v)$ and $\lambda^f(v)$ are *both* smaller than one:

$$\lambda^m(v) < 1 \Leftrightarrow \frac{g^f\left(G^{m^{-1}}(v)\right)}{g^m\left(G^{m^{-1}}(v)\right)} < 1 \Leftrightarrow G^{m^{-1}}(v) < y^* \Leftrightarrow v < G^m(y^*) \tag{12}$$

$$\lambda^f(v) < 1 \Leftrightarrow \frac{g^m\left(G^{f^{-1}}(v)\right)}{g^f\left(G^{f^{-1}}(v)\right)} < 1 \Leftrightarrow G^{f^{-1}}(v) > y^* \Leftrightarrow v > G^m(y^*), \tag{13}$$

[7] In the empirical analysis in section 4 I show that the assumption of monotonic likelihood ratios in incomes is reasonable in a US context.

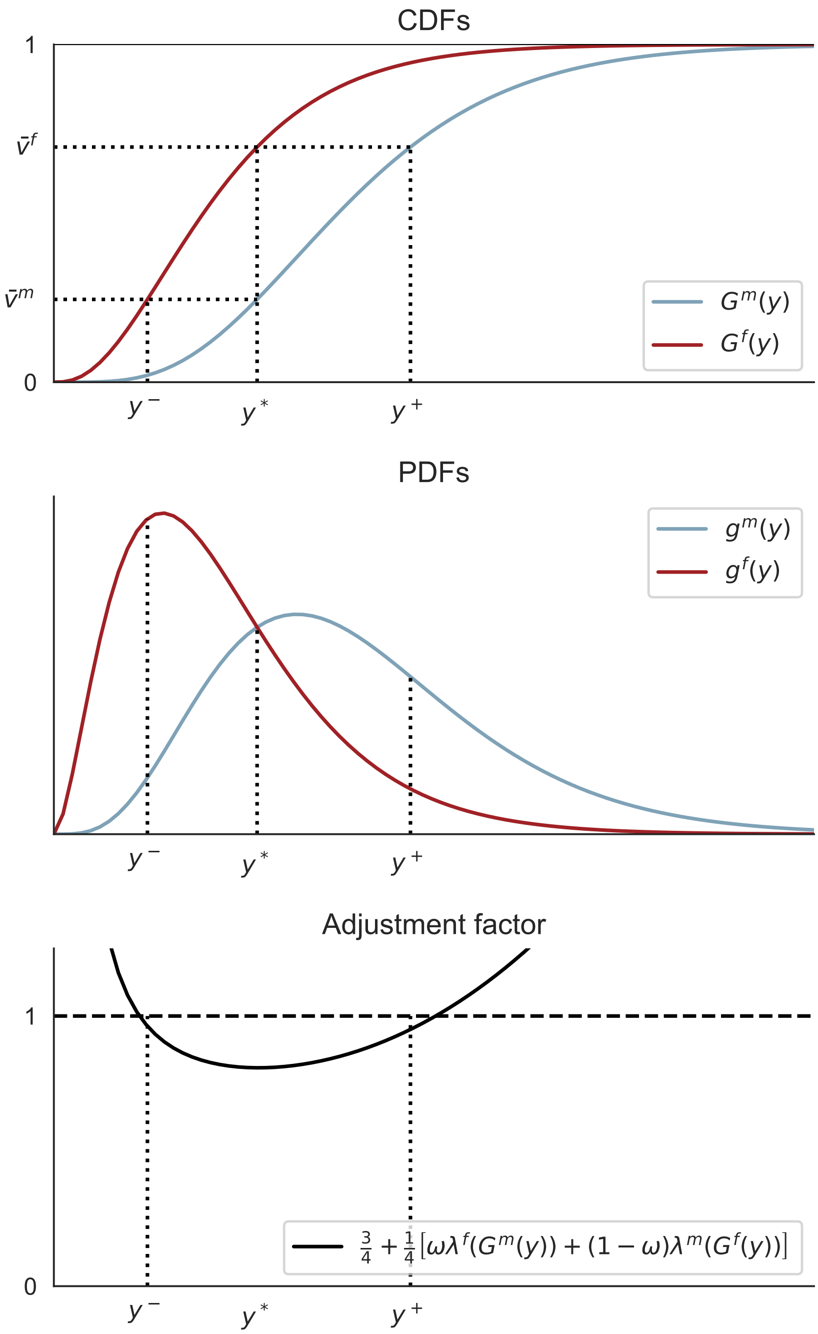


Figure 2: Range where aggregate covariance $<$ within-gender covariance

The figure depicts artificial distributions which exhibit monotonic likelihood ratios and $\frac{g^f(v)}{g^m(v)}$ is decreasing. The top figure displays the cumulative distribution functions (CDFs). From these functions one can identify $\bar{v}^m = G^m(y^*)$ and $\bar{v}^f = G^f(y^*)$ where the likelihood ratios are evaluated, where y^* is the crossing of densities. The ranks are transformed back to outcomes, $y^- = {G^f}^{-1}(\bar{v}^m)$ and $y^+ = {G^m}^{-1}(\bar{v}^f)$. Evaluating the covariance at any outcome level in the interval (y^-, y^+) the aggregate rank-correlation is smaller than the within gender correlation. With the distributions specified in this example, y^- and y^+ correspond to the 13th and the 82nd percentile in the aggregate outcome distribution.

where the second inequality follows from the single crossing property. Therefore, the adjustment term is less than one for $\hat{v} \in \left[G^m(y^*), G^f(y^*)\right]$.

The interval can be read off Figure 2 in the following way. Find the crossing point, y^*, in the middle figure. Now evaluate this value in the CDFs in the top figure to find the corresponding bounding ranks, $\bar{v}^m = G^m(y^*)$ and $\bar{v}^f = G^f(y^*)$. These bounds can, in turn, be transferred back into outcome levels, denoted by y^- and y^+. The adjustment term is less than one for every outcome level within these bounds.

A calculation of the adjustment term is performed in the bottom plot of Figure 2. Note that these bounds are sufficient but not necessary conditions for an adjustment factor smaller than one. The adjustment term can be smaller in cases where either λ^m or λ^f are larger than one, as long as the other ratio is sufficiently small.

Evaluation at a single outcome level The approximated relationship in Equation (6) is evaluated at a *rank* level. However, if we want to approximate the adjustment factor at an *outcome* level, this may correspond to two different ranks: one for women and one for men. The two ranks give rise to two different adjustment factors. In order to collapse the factors into a simple evaluation, I suggest taking the density-weighted mean, where the densities are evaluated at an outcome level in the aggregate distribution.

Denote $\bar{y}$ as the outcome of evaluation and $\omega = \frac{g^m(\bar{y})}{g^m(\bar{y})+g^f(\bar{y})}$ as the share of males at $\bar{y}$. Now define the two ranks corresponding to $\bar{y}$ for men as $\bar{v}^m$ and for women $\bar{v}^f$. We can now calculate the weighted mean of the adjustment term:

$$\begin{aligned} &\omega A(\bar{v}^m) + (1-\omega)A(\bar{v}^f) \\ =&\frac{3}{4} + \frac{1}{4}\left(\omega\lambda^f(\bar{v}^m) + (1-\omega)\lambda^m(\bar{v}^f)\right), \end{aligned} \tag{14}$$

where I have assumed equal overall shares of men and women.[8] An example of such an evaluation at the mean of the outcome distributions is displayed in Appendix Figure 1. In the example, the adjustment factor is less than one for all outcomes but the ones in the extreme tails of the distribution.

4 Gender equality and US income mobility

To assess the empirical relevance, I calibrate the developed adjustment factors to empirical income distributions and investigate the importance of gender inequality. Firstly, I describe the data and estimation techniques. I then turn to the calibration of the approximated weights developed in the previous section. Lastly, I briefly compare the calibrated weights to results obtained from copula-simulations.

4.1 Estimation of income distributions

I estimate the income distributions on data provided by IPUMS-USA (Ruggles et al. 2017), from the largest sample available in each year from 1970 to 2016. I restrict attention to individuals between the ages of 30 to 39. The variable of interest is *total individual income* (*totinc*). Only individuals with positive earn-

[8]The more general case is given by:

$$\begin{aligned}
&\omega A(\bar{v}^m) + (1-\omega)A(\bar{v}^f) \\
=&\mu^2 + (1-\mu)^2 + \mu(1-\mu)\left(\omega\left[\lambda^m(\bar{v}^m) + \lambda^f(\bar{v}^m)\right] + (1-\omega)\left[\lambda^m(\bar{v}^f) + \lambda^f(v^f)\right]\right) \\
=&\mu^2 + (1-\mu)^2 + \mu(1-\mu)\left(1 + \omega\lambda^f(\bar{v}^m) + (1-\omega)\lambda^m(\bar{v}^f)\right).
\end{aligned}$$

When I assume overall equal shares of males and females in the economy, $\mu = 1/2$ this term collapses to 14. Taking the density-weighted mean is a convenient choice as it simplifies the terms considerably. Other approaches are also possible. As long as the outcome level of evaluation is such that the corresponding male rank translates into an outcome for females below the crossing point for the densities, then $\lambda^f(v^m) < 1$. Likewise, if the corresponding female rank corresponds to an outcome for men above the crossing point, then $\lambda^m(v^f) < 1$. This implies that other weighted means where these conditions are fulfilled also will evaluate to a mean adjustment factor less than one.

ings are kept in the data.[9] The sample is weighted and I draw artificial samples from the original sample according to these weights. I then apply Kernel-Density-Estimation (KDE) to obtain densities, cumulative distributions and quantile functions for each gender in each year. I smooth the estimated functions in order to interpolate values for incomes not present in the data.[10]

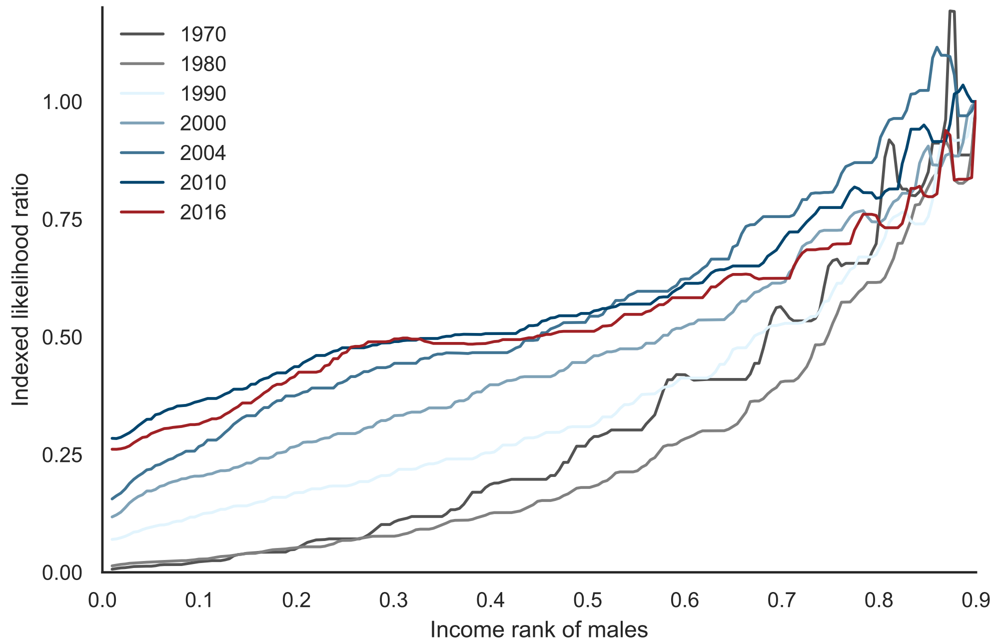
4.2 Results

Figure 3 shows the likelihood ratios and quantile ratios of the estimated income distribution of men over women. Disregarding small deviations, the likelihood-ratios in each year appear to be increasing monotonically as seen in Figure 3i. This implies that the income distributions of males stochastically dominate the distributions of females in each year. This is reflected in Figure 3ii where the quantile ratios are consistently above 1. However the ratios have approached one from above throughout the period since 1970. This implies that while stochastic dominance is maintained, men and women are considerably less unequal today than in the seventies.

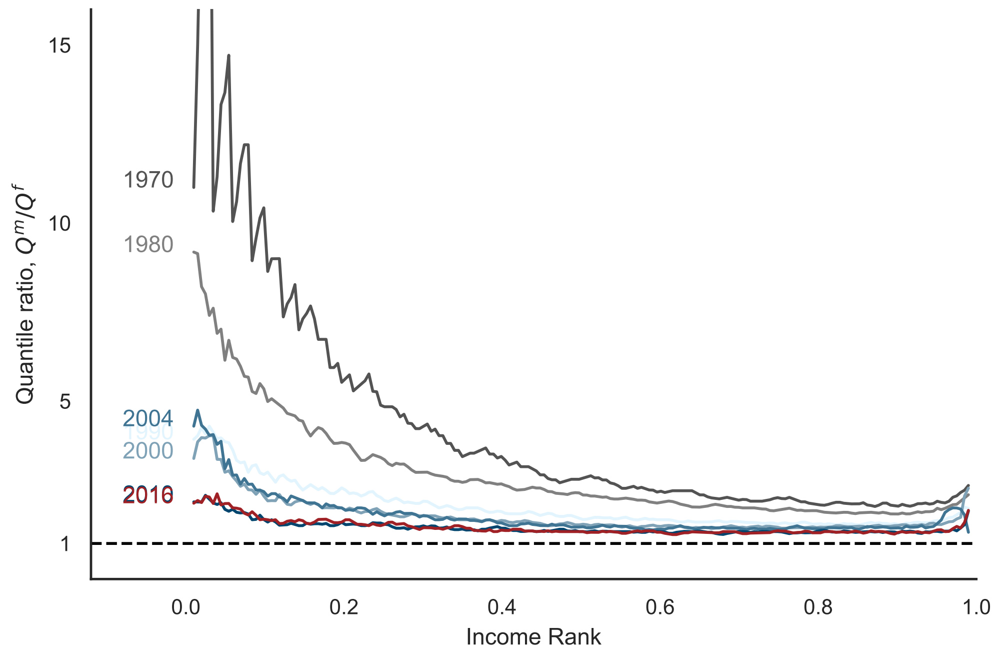
Using the estimated densities, I calibrate the adjustment factor, developed in section 3. The density-weighted mean adjustment factor evaluated at the aggregate mean and median income in each year is shown in Figure 4. The general

[9] This is not an innocuous restriction as it affects individuals along the external margin. The quantitative results become more dramatic is gender-inequality increases by including individuals with no income. However including zero-incomes also breaks with the assumption of smooth distributions. As the purpose of this section is to serve as an illustration, I find the drawbacks of exclusion of zero-incomes acceptable.

[10] When estimating empirical distributions, there are a number of issues such as top-coding, weighting, and sampling error. As the calculations are made for an illustrative purpose, I ignore most of these things. That means that I do not regard my quantitative results as definitive, although highly suggestive of magnitudes. The full code used to generate the results is available on my website. The program is written in Python and extensively documented. One can therefore easily change assumptions to investigate the stability of results. Note, that I use income in a single year. It is recognized across the literature that this is problematic as it may be a poor proxy for lifetime income, see Chen et al. (2017) and Guvenen et al. (2017) for recent discussions. However, Chetty et al. (2014b) find that the rank-correlation in the US is insensitive to the number of years over which to average income levels. The estimated distributions are only used to calculate ratios in the same year. Hence, there is no need to deflate the distributions, as the deflation cancels out.



(i) Likelihood ratios



(ii) Quantile ratios

Figure 3: Estimated distributions

Figure 3ii shows the male quantile distribution divided by the female, and will thus take the value 1 when the quantiles are equal. I omit ratios at the very bottom, as these are very volatile due to division by values close to zero. It is evident that first-order stochastic dominance is maintained for all years. Figure 3i shows the density of males divided by the density of women, i.e. likelihood ratios. For comparison, the ratios are indexed at the 90th percentile of the male income distribution in a given year and evaluated at incomes corresponding to male ranks.

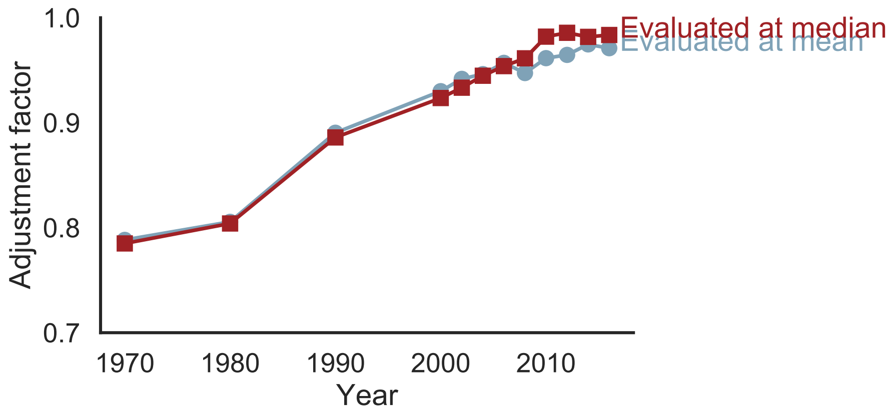


Figure 4: Calibrated adjustment factor

The figure shows the calibrated density weighted adjustment factors according to Equation (14) evaluated at the aggregate mean and median. The weights are not sensitive to choice of the income level at which to evaluate the adjustment factors. For the value of the adjustment factor for other income levels I refer to Appendix Figure 3.

trend confirms the intuition that the increased gender equality should in an of itself increase the total rank-correlation. The adjustment factor rises from 0.79 to 0.97 between 1970 and 2016 as seen in Figure 4. Assuming constant intergenerational mobility within gender, the aggregate mobility would have *fallen* by 23 percent solely due to greater gender equality.[11] In other words, the greater gender equality achieved over the last half of the twentieth century has decreased mobility considerably if measured by the aggregate rank-correlation.[12]

4.3 Copula simulation of adjustment factors

The adjustment factor is developed using a Taylor approximation. Another way to quantify the importance of gender inequality is through copula simulations. This exercise serves to validate the results above. With the estimated the income distributions and assuming a functional form of the copula, I can simulate the effect of changes in gender inequality while holding the within-gender rank-correlation

[11]Calculated as 0.97/0.79-1=0.23.

[12]In appendix Figure 3 I calibrate the adjustment factor for incomes corresponding to male ranks between the first and the 99th percentile. In practice, it matters little at which point in the income distribution the adjustment factor is calculated, as long as it is not in the extreme tails.

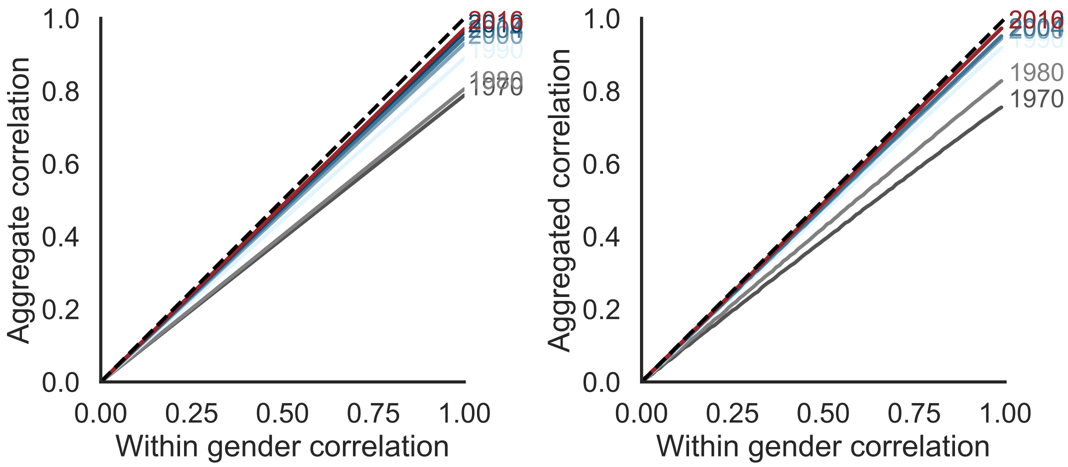
constant.

For simplicity, I make use of Gaussian copulas, as they only contain one parameter in need of calibration. This parameter maps one-to-one to the rank-correlation.[13] Figure 5 compares the results from calibrating the adjustment factor and the copula simulations. The linear relationship in Figure 5i is mechanical, as I plot the within gender rank on the x-axis and the same correlation multiplied by the calibrated adjustment factor on the y-axis. However, I have not imposed the linear relationship in Figure 5ii, where I plot the mean aggregate rank-correlations from the simulations as a function of the imposed within-gender rank-correlation. The figure shows that the linear relationship between the within-gender and total rank-correlation is a good approximation of the true relationship with Gaussian copulas.[14]

Figure 5 also presents another way to interpret the importance of gender inequality for the measurement of mobility. A given rank-correlation can map into many different societies. An aggregate rank-correlation of 0.4 can thus map into a within-gender rank-correlation between 0.4 today and 0.5 in 1970. This can be seen by following a horizontal line from 0.4 on the y-axis in 5i. Furthermore, due to the proportional relationship, the larger the underlying rank-correlation, the larger is the absolute difference between the aggregate and within-gender

[13]The Gaussian copula has been controversial as it cannot describe tail dependence. This is relevant in mobility studies, as one often observes lower mobility in tails of the distribution. In other words, children of high very income parents are much more likely to become very high earners themselves, see Chetty et al. (2014b) and Boserup et al. (2014) for examples of such a pattern in the US, Canada, and Denmark. However, as I do not concern myself with the copula itself, but rather the link between within-gender and aggregate mobility, I find these issues to be of minor importance in the present setting.

[14]The simulation exercise runs as follows: For a range of rank-correlations between 0 and 1, I draw ranks for children and parents from the Gaussian copula corresponding to that rank-correlation along with a random binary gender indicator. I then map the ranks into incomes according to the gender dummy and the estimated gendered quantile functions. The total rank-correlation is then calculated as the correlation between the ranked child income and parent rank. I repeat the exercise a hundred times and take the mean of the resulting hundred rank-correlation. The simulation protocol is described in Appendix C and code is available online. Appendix B shows that the approximation performs well when compared directly to the simulations using Gaussian copulas. Although the relationship is not exactly one, they follow each other closely.



(i) Calibrated adjustment factor (ii) Simulated ratio

Figure 5: Calibrated and simulation association between within-gender and aggregate rank-correlations.

Figure 5i shows the relationship between within-gender and aggregate rank-correlation with calibrated density-weighted adjustment factor as described in Equation (14) evaluated at the aggregate mean. The aggregate gender correlation is simply the adjustment factor are multiplied by the within gender rank-correlation. The dashed black line represents full gender equality. Figure 5ii shows the simulated aggregate rank-correlation as a function of a known within-gender rank-correlation. The simulation assumes a Gaussian copula and follows the protocol described in appendix C. No linear relationship has been imposed in the simulation. Given the copula, the relationship seems to be almost perfectly linear. For a direct comparison between the calibration and simulation see Appendix Figure 2.

correlations.

5 Income concepts and dimensions of inequality

Lastly, I briefly discuss issues of mobility measurement which relate to the findings of this paper.

5.1 Individual or household income

This analysis has exclusively focused on individuals as the unit of analysis. As such, the results apply directly to papers using ranks based on individual outcomes to compare societies, such as Landersø and Heckman (2017); Adermon et al. (2018). Other studies have instead used household incomes to rank indi-

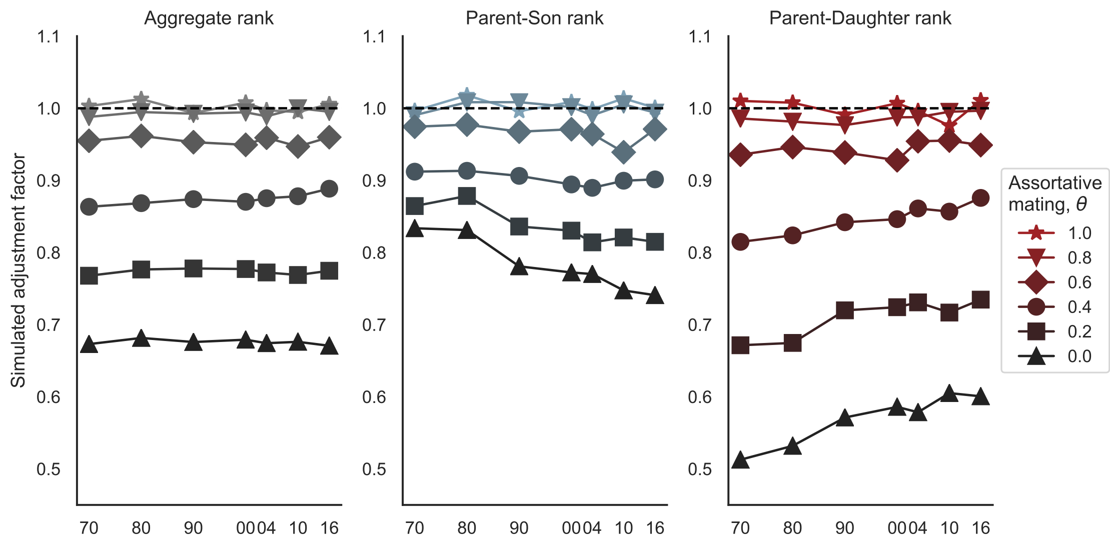


Figure 6: Assortative matching and gender equality

The figure shows the interaction of assortative matching and gender inequality for measured rank-correlations. Within-gender rank-correlations are held constant at different levels through copula simulations and incomes are retrieved via the estimated quantile distributions described in Section 4.2. Assortative matching is simulated by sorting the male income vector by rank, v and sorting the female income vector by $\theta v + (1-\theta)x$ where $x \sim \mathcal{U}(0,1)$ and $\theta \in [0,1]$. θ is the degree of assortative matching and is evaluated for a range of positive values. The simulation assumes that all children form couples and their simulated individual incomes are summed to form household income. The simulated adjustment factor is calculated as the ratio of within-gender rank-correlations and the aggregate rank-correlation on household income for the specified group.

viduals. In this context men and women in couples share the same rank in the distribution. This introduces a new parameter in the framework: the degree of assortative matching: The shape of the distribution of household income is now affected by the degree to which men with high (low) will tend to match with women with high (low) incomes.

I construct a simple simulation exercise where I group men and women according to a degree of assortative matching. I then calculate the implied rank-correlations using household incomes rather than individual incomes. Figure 6 shows that the degree of assortative matching is orders of magnitude more important for measures based on household income than changes in gender equality. However, the interaction between assortative matching and gender inequality can lead to opposing findings.

If couples form according to perfect assortative matching, the shapes of the distributions of men and women do not matter for the relationship between the within-gender rank-correlation and the aggregate correlation using household income. Intuitively, the summation of two perfectly correlated vectors of incomes does not affect the ranking of individuals. This finding holds regardless of measuring correlations for sons or daughters or across genders. However, this is a knife-edge case. In the absence of perfect matching, the adjustment factor between aggregate and within-gender rank-correlations shift downwards. In other words, non-perfect sorting works the same way as gender inequality. An increase of sorting into couples will register as a fall in mobility, even if within-gender rank-correlation and gender inequality is held constant.

For the aggregate rank-correlation, changes in gender inequality over time are not reflected in the correlations as the coupling of incomes for a given degree of assortative matching evens out the changes in the gender income gap. This can explain the stable rank-correlations documented by Chetty et al. (2014b). However, when splitting the population into subgroups, income distributions enter the fray again. Holding the degree of assortative matching fixed, the closing of the gender income gap will register as a *decrease* in the rank-correlation for sons and an *increase* in the correlation for daughters. In other words, due to greater gender equality, two opposing narratives of the development in intergenerational mobility emerge depending on the choice of population.

The two different narratives can be explained by the increase in the importance of women's income for household income. Under non-perfect assortative matching, the increase in income of women adds "noise" to the rank-correlation for sons and increases the "signal" for the correlation for daughters. Hence rank-mobility for sons increases and rank-mobility for women decreases. The simulations, therefore, show that using household income is not sufficient to elevate concerns of the effect of group inequality on rank-correlations.[15]

Wealth inequality, bequests, and birth order As previously discussed, birth order is a random process much akin to gender. As such, the rank-correlation will tend to be lower, the larger is the systematic difference in earnings or wealth between older and younger siblings. This is especially important in the context of wealth inequality and inheritance. Societies differ in the extent to which the oldest child inherits the estate of deceased patters and this has long been recognized to affect inequality, see Menchik (1980). This, however, also directly affects the rank-mobility. In a context where each estate has two descendants, the rank-correlation will be maximized when they each receive an equal share of the bequests. For any other division, the rank-correlation will be smaller. Thus legislative or normative changes towards a more equal sharing of estates among siblings will tend to mechanically lower mobility in society.

Racial inequality An important question is whether insights of this paper translate to other forms of group inequality, the most present being race. All types of families have sons and daughters and siblings will necessarily differ in the order of birth. This allows for the interpretation of gender and birth order as a completely random sorting mechanism uncorrelated to parent rank. This randomness is not present in case of racial or ethnic inequality. In this case, assignment to a group, i.e. race, will correlate with parent income. Thus, the aggregate rank-correlation can be *higher* than the within-race rank-correlation. This could be the case if the races are so unequal that their income distributions have little overlap *in both generations*. This can be seen directly from the decomposition of the parent-child covariance into its within- and between-group components in Equation (5). With

[15]If assortative matching changes over time, the implications for mobility will be even more significant for mobility measures. Chadwick et al. (2002) discuss these issues in the context of estimating IGEs, and Eika et al. (2018) document recent trends in assortative matching. The effect of a rise in assortative matching can be seen in Figure 6 by changing curves.

race, the between-race covariance will be positive if race-inequality is persistent from one generation to the next. In other words, nor in the case of race inequality, does the rank-correlation allow us to ignore group inequality, but the effect of the adjustment factor between within and aggregate mobility might be larger or smaller than 1.

6 Conclusion

The potential for confusing increased wealth or gender equality with decreased mobility shows how the interpretation of rank-based mobility measures is less obvious than it might initially appear. Relative mobility measures are an important part of the toolbox for investigating mobility and inequality. But too little focus has been afforded the importance of inequality for positional mobility measures; It is not correct to claim that the use of rank-mobility allows the researcher to disregard inequality. No one correct mobility measure exists, nor should there, but this paper has illustrated the perils of comparing rank-mobility in societies over time and across space without theoretical guidance on who to compare with whom.

References

Adermon, Adrian, Mikael Lindahl, and Daniel Waldenström, "Intergenerational wealth mobility and the role of inheritance: Evidence from multiple generations," *The Economic Journal*, 2018, *128* (612), F482–F513.

Blau, Francine D and Lawrence M Kahn, "The gender earnings gap: learning from international comparisons," *The American Economic Review*, 1992, *82* (2), 533–538.

Boserup, Simon Halphen, Wojciech Kopczuk, and Claus Thustrup Kreiner, "Stability and persistence of intergenerational wealth formation: Evidence from Danish wealth records of three generations," *Unpublished Working Paper. University of Copenhagen*, 2014.

Chadwick, Laura, Gary Solon, and Laura Chadwick, "Intergenerational Income Mobility among Daughters," *The American Economic Review*, mar 2002, *92* (1), 335–344.

Chen, Wen-Hao, Yuri Ostrovsky, and Patrizio Piraino, "Lifecycle variation, errors-in-variables bias and nonlinearities in intergenerational income transmission: new evidence from Canada," *Labour Economics*, 2017, *44*, 1–12.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States - Online Appendix," *The Quarterly Journal of Economics*, 2014, (June), qju022.

_ , _ , _ , _ , and Nicholas Turner, "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility †," *American Economic Review*, may 2014, *104* (5), 141–147.

Eika, Lasse, Magne Mogstad, and Basit Zafar, "Educational assortative mating and household income inequality," 2018.

Fagereng, Andreas, Magne Mogstad, and Marte Ronning, "Why do wealthy parents have wealthy children?," *University of Chicago, Becker Friedman Institute for Economics Working Paper*, 2018, (2019-22).

Fletcher, Jason and Katie M. Jajtner, "Intergenerational Health Mobility: Magnitudes and Importance of Schools and Place," *NBER Working Paper Series*, 2019, *No. 26442*.

Guvenen, Fatih, Greg Kaplan, Jae Song, and Justin Weidner, "Lifetime incomes in the United States over six decades," Technical Report, National Bureau of Economic Research 2017.

Jäntti, Markus and Stephen P. Jenkins, "Income Mobility," in "Handbook of Income Distribution," 1 ed., Elsevier B.V., 2015, pp. 807–935.

Landersø, Rasmus and James J Heckman, "The Scandinavian fantasy: The sources of intergenerational mobility in Denmark and the US," *The Scandinavian journal of economics*, 2017, *119* (1), 178–230.

Lundberg, Shelly., Sara. McLanahan, and Elaina. Rose, "Child Gender and Father Involvement in Fragile Families," *Demography*, 2007, *44* (1), 79–92.

Menchik, Paul L, "Primogeniture, equal sharing, and the US distribution of wealth," *The Quarterly Journal of Economics*, 1980, *94* (2), 299–316.

Meyer, Christian, "The Bivariate Normal Copula," *Communications in Statistics - Theory and Methods*, 2013, *42* (13), 2402–2422.

Nelsen, Robert B., *An Introduction to Copulas* Springer Series in Statistics, New York, NY: Springer New York, 2006.

Orzack, Steven Hecht, J. William Stubblefield, Viatcheslav R. Akmaev, Pere Colls, Santiago Munné, Thomas Scholl, David Steinsaltz, and James E. Zuckerman, "The human sex ratio from conception to birth," *Proceedings of the National Academy of Sciences*, 2015, *112* (16), E2102–E2111.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek, "Integrated Public Use Microdata Series: Version 7.0 [dataset]," 2017.

Sklar, M, "Fonctions de repartition an dimensions et leurs marges," *Publ. inst. statist. univ. Paris*, 1959, *8*, 229–231.

Wong, Ashley, Bhashkar Mazumder, Timothy J Halliday et al., "The Intergenerational Transmission of Health in the United States: A Latent Variables Analysis," Technical Report 2019.

A Taylor approximation of covariance

As in the main text, I assume that a family can be described by 3 variables:

$$(U_i, V_i, S_i),$$

where U_i and V_i are parent and child rank and S_i is an indicator variable which evaluates to one if the child is male. For simplicity, I assume that the covariance between parent and child rank is not a function of gender, $E(V_i|U_i, S_i) = E(V_i|U_i)$. This is easily changed, although the resulting equations will be less neat. Furthermore, I assume that parent income does not depend on the gender of the child, $E(U_i|S_i) = E(U_i)$. The income distributions are gender-specific and taken for granted. The translation from child rank to child income is described by the following relation:

$$y(v, s) = sG^{m^{-1}}(v) + (1 - s)G^{f^{-1}}(v),$$

where G^m and G^f are the cumulative income distribution functions of male and female children respectively. I assume these to be invertible, thereby excluding mass-points, including zero income. The *aggregate* income distribution of children is then a mixture distribution given by $G(y) = \mu G^m(y) + (1 - \mu)G^f(y)$, where $\mu = E[S_i]$.

Given these definition, the aggregate rank, v^A can now be described as a function of the within-gender rank, v. Using the *probability integral transform*:

$$v^A(v, s) = G(y(v, s)) = \mu G^m(y(v, s)) + (1 - \mu)G^f(y(v, s)) \tag{15}$$

Define $v^A = \Lambda(v, s)$. All results in this section will relate to the functional properties of $\Lambda(v, s)$.

We want to approximate the covariance between parent rank and child aggregate rank, $Cov(U, V^a)$. Due to the law of total covariance:

$$\begin{aligned} Cov(U, V^a) =& Cov(U_i, \Lambda(V, S)) \\ =& \underbrace{E\left[Cov(U, \Lambda(V, S)|S)\right]}_{\text{within gender}} + \underbrace{Cov\left(E\left[U|S\right], E\left[\Lambda(V, S)|S\right]\right)}_{\text{between gender=0}}. \end{aligned} \tag{16}$$

The decomposition shows that we can handle the description of the covariance in two separate parts. The first term is the mean of the covariance between parent rank and child aggregate rank *conditional on gender*. The second term is the covariance between means of parent and child rank for families of sons and daughters. Because I have assumed that $E(U_i|S_i) = E(U_i)$, the latter term evaluates to zero and we only need to focus of the within-gender covariance.

Because S is a binary variable, it follows that (16) can be written as:

$$E\left[Cov(U, \Lambda(V, S)|S)\right] = \mu \underbrace{Cov(U, \Lambda(V, 1))}_{(i)} + (1-\mu) \underbrace{Cov(U, \Lambda(V, 0))}_{(ii)}. \tag{17}$$

As will become clear, it is easier to handle gender-specific versions of $\Lambda(v, s)$. Therefore define:

$$\Lambda^m(v) = \Lambda(v, 1) = \mu v + (1-\mu) G^f\left(G^{m^{-1}}(v)\right) \tag{18}$$

$$\Lambda^f(v) = \Lambda(v, 0) = \mu G^m\left(G^{f^{-1}}(v)\right) + (1-\mu) v \tag{19}$$

We first focus on (i) in (17):

$$Cov(U, \Lambda(V, 1)) = E\left[U\Lambda(V, 1)\right] - E\left[U\right] E\left[\Lambda(V, 1)\right] \tag{20}$$

Insert (18):

$$
\begin{aligned}
Cov(U, \Lambda(V,1)) =& E\left[U\left(\mu V + (1-\mu)G^f\left(G^{m-1}(V)\right)\right)\right] \\
& - E\left[U\right] E\left[\mu V + (1-\mu)G^f\left(G^{m-1}(V)\right)\right] \\
=& \mu \underbrace{Cov(U,V)}_{\text{Within-gender covariance}} + (1-\mu)\underbrace{Cov\left(U, G^f\left(G^{m-1}(V)\right)\right)}_{\text{“Adjusted covariance”}}
\end{aligned} \tag{21}
$$

While the within-gender covariance is known, we cannot yet describe the “adjusted covariance”. We, therefore, reformulate the second term:

$$
\begin{aligned}
& Cov\left(U, G^f\left(G^{m-1}(V)\right)\right) = \\
& E\left\{(U - E[U])\left(G^f\left(G^{m-1}(V)\right) - E\left[G^f\left(G^{m-1}(V)\right)\right]\right)\right\}
\end{aligned} \tag{22}
$$

Now perform a first order Taylor expansion of $G^f\left(G^{m-1}(V)\right)$ around $\hat{v}$:

$$
G^f\left(G^{m-1}(v)\right) \approx G^f\left(G^{m-1}(\hat{v})\right) + \underbrace{\left.\frac{g^f\left(G^{m-1}(v)\right)}{g^m\left(G^{m-1}(v)\right)}\right|_{\hat{v}}}_{\lambda^m(\hat{v})} (v - \hat{v}) \tag{23}
$$

Reinserting equation (23) into equation (22):

$$
\begin{aligned}
& Cov\left(U, G^f\left(G^{m-1}(V)\right)\right) \approx \\
& \lambda^m(\hat{v}) E\left\{(U - E[U])\left([(V-\hat{v}) - E[V-\hat{v}]]\right)\right\}
\end{aligned} \tag{24}
$$

Reinserting equation (24) into equation (21):

$$
\begin{aligned}
Cov(U, \Lambda(V,1)) \approx& \mu Cov(U,V) \\
& + (1-\mu)\lambda^m(\hat{v}) E\left\{(U - E[U])\left([(V-\hat{v}) - E[V-\hat{v}]]\right)\right\} \\
=& \mu Cov(U,V) + (1-\mu)\lambda^m(\hat{v}) Cov(U, V-\hat{v}) \\
=& \left\{\mu + (1-\mu)\lambda^m(\hat{v})\right\} Cov(U,V)
\end{aligned}
$$

We perform the exact same operation on (ii) in equation (17) which yields:

$$Cov(U, \Lambda(V, 0)) \approx \left\{\mu\lambda^f(\hat{v}) + (1-\mu)\right\} Cov(U, V)$$

where $\lambda^f(v) = \frac{g^m\left(G^{f^{-1}}(v)\right)}{g^f\left(G^{f^{-1}}(v)\right)}$. Reinserting into equation (17) and inserting into equation (16) yields:

$$Cov(U, V^a) \approx \left\{\mu^2 + (1-\mu)^2 + \mu(1-\mu)\left[\lambda^m(\hat{v}) + \lambda^f(\hat{v})\right]\right\} Cov(U, V) \tag{25}$$

Assuming equal share of men and women, $\mu = 1/2$, the expression simplifies to:

$$Cov(U, V^a) \approx \left\{\frac{1}{2} + \frac{1}{4}\left[\lambda^m(\hat{v}) + \lambda^f(\hat{v})\right]\right\} Cov(U, V),$$

which is identical to equation (6) in the main text.

B Additional figures

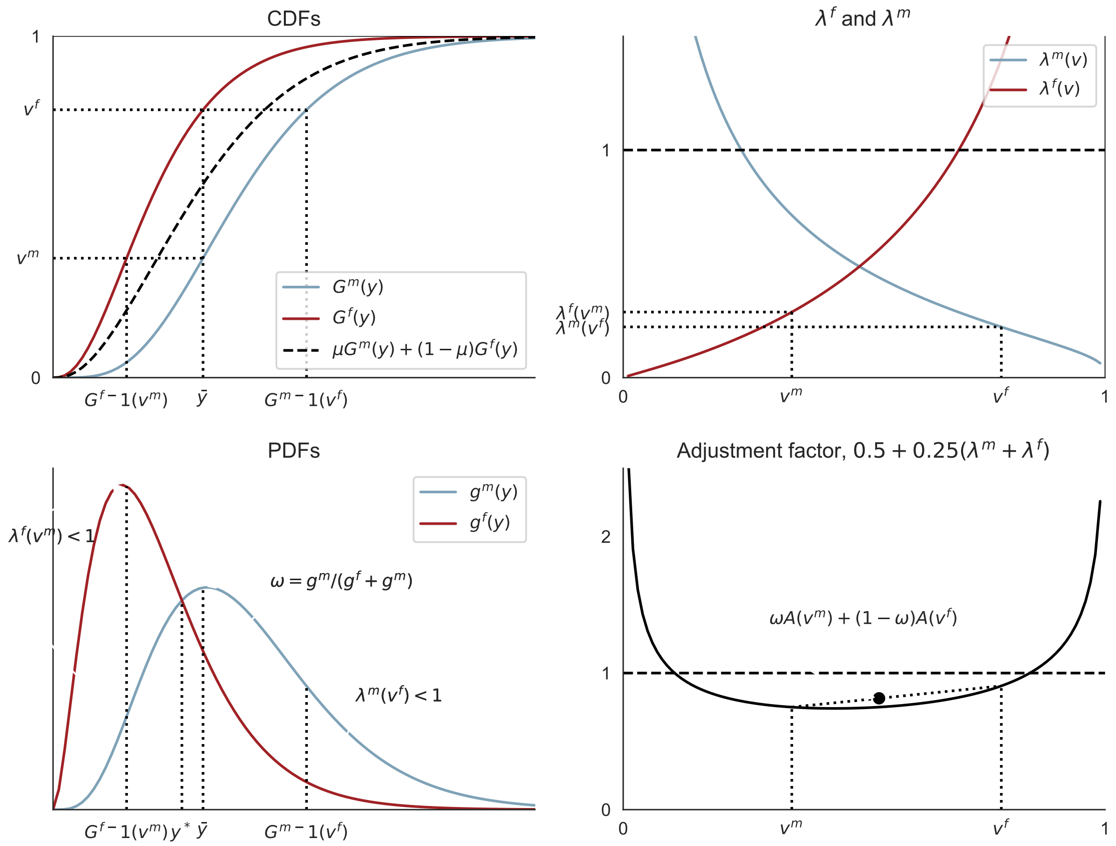


Figure 1: Example of evaluation of the adjustment factor at the mean income

The figure depicts artificial distributions which exhibit monotonic likelihood ratios. In this example the adjustment factor is evaluated at the mean of the aggregate distribution, $\bar{y}$. The top left figure displays the cumulative distribution functions (CDFs). From these functions one can identify $v^m = G^m(\bar{y})$ and $v^f = G^f(E(y))$ where the likelihood ratios are evaluated. In this example ${G^f}^{-1}(v^m) < y^* < {G^f}^{-1}(v^f)$, where y^* is the crossing of densities. This implies that both $\lambda^f(v^m)$ and $\lambda^m(v^f)$ are below one. This can be read off the top right figure, which plots the values for λ^f and λ^m for all values of the rank, v. The adjustment factor for all values of v is shown in the bottom right figure, where the density weighted mean between the two adjustment factors are shown as a black circle.

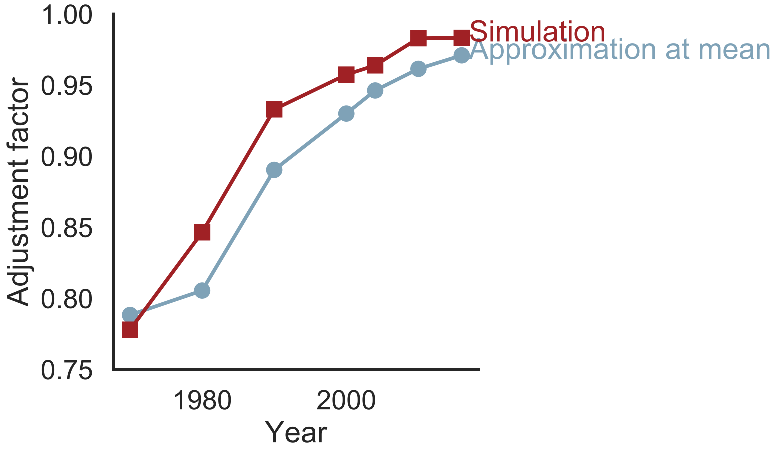


Figure 2: Calibrated and simulated adjustment factor

The figure shows the calibrated adjustment factors according to Equation (14) in blue. It is evaluated at the aggregate mean. In red I have taken the average ratio between the total and the within-gender rank correlation across all values of the with-gender rank correlation. As seen in Figure 2 the relationship between the two rank correlations is approximately linear and thus have an almost constant ratio.

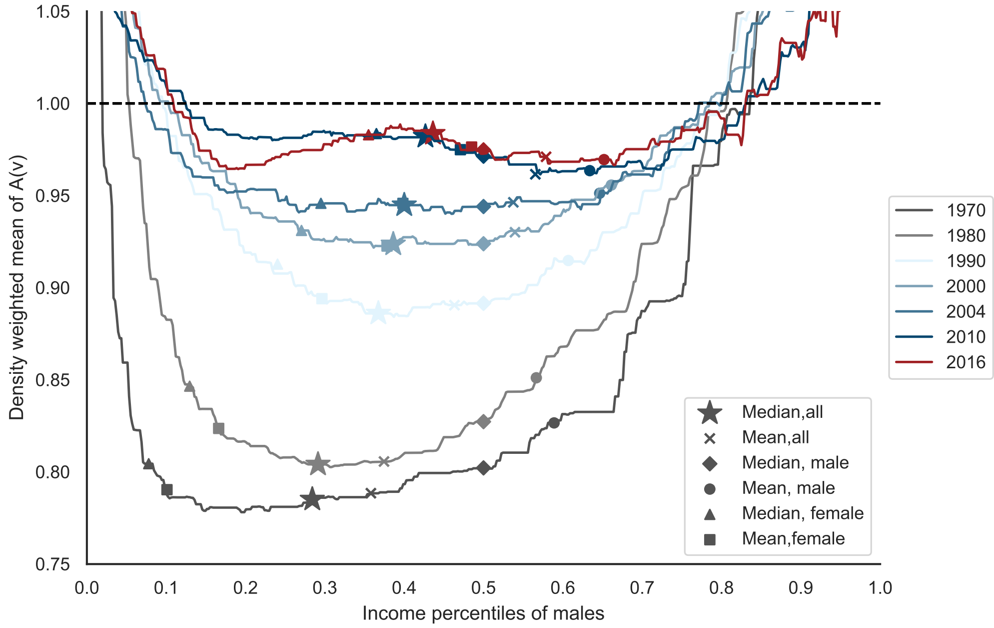


Figure 3: Evaluation of the adjustment factor at different levels of income

The figure shows the importance of the choice of evaluation point. For each year I calculate the density weighted adjustment factor at an income corresponding to the rank on the x-axis for males. The ranks serve as a normalisation of the income distributions, such that one can compare across years. The very flat lines indicate that it matters little where the adjustment factor is approximated, less it not be in the tails. The figure also displays the adjustment factors for means and medians for the two genders and for the total income distribution. These are represented by the symbols on the line placed at the corresponding ranks in the male income distribution. By definition, the median of males are always located at 0.5 on the x-axis.

C Simulation protocol

The data is generated from a Gaussian copula, which only take one parameter which maps 1:1 with the rank correlation. Marginal distributions for parents and children have to continuous and invertible. No gender is attributed to parents and thus it is assumed that the marginal distribution of parents is uniform, $F_p(x) = x$. As described in the text children are divided into to groups m and f having different marginal distributions, G^m and G^f respectively. Let G be the income distribution containing both group m and f. The protocol for simulating is as follows:

For a given ρ repeat R times:

- Draw two vectors K_p and K_c from a bivariate normal distribution[16], $K_p, K_c \sim N\left([0,0], \begin{bmatrix} 1 & \rho_G \\ \rho_G & 1 \end{bmatrix}\right)$
- Apply the cdf of the standard normal distribution to K_p and K_c to obtain U and V: $U = \Phi(K_p), V = \Phi(K_c)$
- Depending on gender:
  - If no gender:
    - Apply quantile functions of the marginal distributions to ranks to obtain income variables:
      - For parents: $Y_p = F_p^{-1}(U)$
      - For children: $Y_c = G_c^{-1}(V)$
  - If gender:
    - Draw a binary variable, S with probability μ.

[16]In the Gaussian copula there is a 1:1-mapping between the correlation in the Gaussian copula, ρ, and the rank correlation coefficient,ρ_s:$\rho_G = 2\sin(\rho\frac{\pi}{6})$. See Meyer (2013) for details.

* Apply quantile functions of the marginal distributions to ranks to obtain income variables.
    - For parents: $Y_p = F_p^{-1}(U)$
    - For children: $Y_c = S \times G^{m^{-1}}(V) + (1 - S) \times G^{f^{-1}}(V)$

Finally take the mean of the desired statistic over the R realizations.