# Black-White differences in intergenerational mobility in the US: evidence from heterogeneous sibling correlations*

Mauricio Bucca[1]

[1]European University Institute, Department of Political and Social Sciences
*mauricio.bucca@eui.eu

December 20, 2018

**Abstract**

In the literature on Black-White differences in social mobility there is a tension between accounts that report the existence of a "perverse openness" among African Americans and theories that predict convergence in the rates of social mobility of Black and Whites. I address this debate by asking three inter-related questions: (1) Does the sibling correlation in income differ for Blacks and Whites? (2) If so, is it because the effect of family background is different for the two groups or because Blacks and Whites produce a different extent of heterogeneity in economic achievement within the family? (3) Are Black-White differences in sibling correlation due to race-related factors? Can they be explained by the underlying socioeconomic characteristics of these two subpopulations? To answer these questions I develop Bayesian hierarchical linear models with "organized dispersion", which permit explicit modeling of heterogeneity in sibling correlations and the variance components that go into them. Using PSID data I find that Black-White differences in sibling correlation for men are mostly explained by the higher within-family heterogeneity of Black siblings as compared to White ones. In turn, such larger extend of heterogeneity is partly -but not entirely- driven by the comparatively poorer socioeconomic standing of Black families. For women, similarity in sibling correlation is due to parental income having an opposite effect on within-family variance for Blacks and Whites, a pattern that is hidden in the describe comparison. Results highlight the interaction between race and social background at producing differences in mobility for the two major racial groups in the US.

Early sociological research on social mobility documented a "perverse openness" among the African American population: a weak association between social origin and social destination that, coupled

*THIS IS A PRELIMINARY DRAFT. PLEASE DO NOT CITE OR DISTRIBUTE WITHOUT PERMISSION OF THE AUTHOR.

with the poor socioeconomic standings of Blacks, condemned subsequent generations of African Americans to a rapid regression towards a very low mean (Blau and Duncan, 1967). In other words, Black parents who had managed to attain a relatively advantageous social position could not effectively pass on those advantages to their offspring, while Black children born to disadvantaged families were likely to move upward, albeit towards a still-precarious material condition. This finding was then interpreted as demonstrating the pervasive effect of race and racial discrimination on achievement, capable of nullifying the stratifying power of social background evident in the White population (Blau and Duncan, 1967, p.208-209).

Just a decade later, on the heels of Civil Rights reforms, Wilson (1978) predicted that the socioeconomic advancement of African Americans would lead to convergence in the rates of intergenerational mobility of Blacks and Whites. In other words, as exclusionary practices abated, and more Blacks reached levels of material well-being similar to Whites, social origins would determine the economic opportunities of Blacks with the same strength that they did for Whites (Wilson, 1978). By the mid-1970s influential sociological work had reported consistent evidence in favor of Wilson's hypothesis. Using linear regression models, Featherman and Hauser (1976) found an increase in the effect of family background on the occupational achievement of Blacks between 1962 and 1973. Similarly, results from log-linear models showed a decline in intergenerational mobility among Black men (Hout, 1984), a transformation mostly driven by Black men from advantaged socioeconomic backgrounds benefiting from new opportunities unlocked by social reforms.

Whether or not convergence in the rates of intergenerational mobility has occurred, and which mechanisms might drive Black-White differences in mobility (or the lack thereof), remains poorly understood. Theoretical expectations are complicated by the simultaneous action of both equalizing and unequalizing forces for social mobility (Bloome and Western, 2011). On the one hand, cultural and institutional transformations aimed to fight race-based inequalities (e.g., expansion of legal equality, affirmative action policies, delegitimation of overt racism) might have ameliorated the "perverse" aspects of social openness among Blacks. On the other, the advancement of new forms of inequality that, although not based on race *per se*, disproportionately affect racial minorities might have offset these equalizing effects (e.g., the take-off on income inequality, school and neighborhood "re-segregation", mass incarceration)(Manduca, 2018).

Nevertheless, in recent years only a handful of studies have examined differences in the rates of intergenerational mobility among Blacks and Whites in the US (Bhattacharya and Mazumder, 2011). There are, however, some exceptions. In a series of studies, Conley and coauthors analyze sibling correlations in multiple socioeconomic outcomes for both racial groups (Conley and Glauber, 2005, 2008). Consistent with accounts of a "perverse openness", the general finding of these studies is that the resemblance in adult socioeconomic outcomes among Black siblings is much weaker than among Whites (Conley and Glauber, 2005), a result that also holds for siblings reared in disadvantaged family structures (Conley and Glauber, 2008). In contrast, Conley et al. (2007) found mixed results regarding Black-White differences in early childhood behavioral and cognitive outcomes.

More recently, Bloome and Western (2011) investigate whether the rise in income inequality in the last three decades has been accompanied by a decline in social mobility for Black and White men. By comparing intergenerational elasticities (IGE hereafter) between the cohort of men born in the late 1940s and the cohort born in the early 1960s, the authors show that educational mobility increased for Black men over this period, but income mobility declined for both racial groups. Moreover, income mobility declined faster among Blacks, reaching similar levels of intergenerational association. Unlike Conley and colleagues' results, these findings are consistent with the hypotheses of Black-White similarity in mobility rates.

Although these studies report results that, at first blush, are contradictory, it is possible that their measurement strategy capture different facets of the mobility process. While IGEs by definition analyze between-family variation (in fact, most IGE studies analyze parent-children dyads), the sibling correlation takes into account both between and within family heterogeneity. In fact, Conley and colleagues speculate that the observed Black-White difference in sibling correlation reflects differential heterogeneity within the family. Their motivating hypothesis is that socioeconomically disadvantaged families (of which Black families are used a proxy) produce greater disparities between siblings because, under resource constraints, parents might invest in children who are more likely to benefit from extra resources (i.e., the most endowed), at the expense of their siblings (Behrman et al., 1982; Conley and Glauber, 2005). By comparison, Bloome and Western (2011) examine factors that operate beyond the family to understand Black-White differences in mobility, particularly changes in the transmissibility of education and income, as well as changes in the returns to education across time and racial groups.

The debate between accounts of a "perverse openness" and hypothesis of Black-White similarity in mobility rates represents an unresolved puzzle in the stratification literature. The same is the case for the long-standing tension between a racial or socioeconomic explanation of Black-White mobility differences. In this paper, I develop a comprehensive analytic framework that permits not only to compare sibling correlations across Black and White populations, but also to investigate the factors that drive potential differences in mobility across these groups. In addition, this framework situates the IGE within the more general structure of the sibling correlation, thus aiming to reconcile previous results reported using these two distinct approaches. In particular, the article asks three intertwined questions:

1. Does intergenerational transmission of income differ for Blacks and Whites?
2. If so, is it because the effect of family background on children's socioeconomic outcomes is different for the two subpopulations? (a between-family explanation) Alternatively, is it because Black and White families produce different degrees of heterogeneity in income among siblings? (a within-family explanation).
3. Can Black-White differences in sibling correlation be explained by the underlying socioeconomic characteristics of these two subpopulations?

# 1 Sibling correlation and comparisons across subpopulations

The primary goal of mobility research is to elucidate the extent to which social origins condition the life chances of individuals. In this spirit, a long tradition in social science research uses the sibling correlation as an overall measure of the role of family and community background on children achievement (Solon et al., 2000; Warren et al., 2002). The intuition behind this approach is that siblings, insofar they share genes and environment (e.g., family, school, neighborhood), would exhibit similar socioeconomic outcomes if genetic and environmental factors were consequential for the obtention of these outcomes. Hence, the more siblings resemble one another, the higher the influence attributable to family and community background. More specifically, as described in equation 2, the sibling correlation measures the proportion of variation in an observable trait that is attributable to the combined influence of these shared factors, independently of idiosyncratic characteristics of the individuals. Formally, let $y_{ijt}$ be a socioeconomic outcome (e.g. income) for the $ith$ sibling in family $j$, observed at year $t$. Assume that such outcome can be described as a linear additive function of three independent components:

$$y_{ijt} = a_j + \mu_{ij} + \nu_{ijt} \tag{1}$$

Here $a$ is a family component of individual income, which captures the combined effect of family and community background and is common to all children of the same family. $\mu$ and $\nu$, on the other hand, are idiosyncratic individual components. $\mu$ is an individual-specific permanent component that captures the long-term effects of individual characteristics on income. $\nu$ correspond to transitory deviations from individual permanent income, which reflects noise due to both temporary shocks and measurement error (Mazumder, 2011). Thus, it follows from this formulation that the correlation in permanent income for a randomly chosen pair of siblings can be expressed as:

$$\rho = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_\mu^2} \tag{2}$$

Although not properly a measure of social mobility, the sibling correlation is informative of how much of an individual's achievement is unrelated to her social origins. In this line, some authors claim that the sibling correlation provides an upper bound for the combined intergenerational influence of genetic and environmental factors, generally yielding better predictive power than IGEs (Solon et al., 2000). Against this strict interpretation, a well-acknowledged limitation of the sibling correlation is that it only captures the effect of factors that are shared within the family. However, siblings do not experience the exact same family and community nor do they share all genetic treats. Furthermore, a family might not exert the same influence on all children and siblings might actively try to differentiate from – or resemble – one another. For these reasons, a more conservative interpretation of the sibling correlation is that it measures the combined effects of shared environmental and genetic factors plus inter-sibling effects (Conley and Glauber, 2005).

Using sibling correlations to compare social mobility across subpopulations[1] entails additional challenges. Particularly relevant are the sources of variation that constitute the correlation, namely, between-family and within-family variance. *Ceteris paribus*, a higher between-family variance for one group will result in a larger correlation among siblings, while a higher within-family variance will have the opposite effect. Since these two variance components stem from distinct social processes, subpopulation comparisons should identify the sources of such discrepancy (or the lack thereof), and the mechanisms that might drive heterogeneity in these sources of variation. In other words, a thorough analysis of sibling correlations across groups should specify whether differences in sibling correlation arise from within and/or between family processes, and what factors might explain these differences and the absence of them -possibly due to countervailing effects.

For example, if a subpopulation exhibits comparatively higher within-family dispersion in a socioeconomic outcome, that might indicate that families in that subpopulation are less effective at securing a certain level of achievement for all their members. Theories on parental strategies for investment in children's human capital provide the main basis to understand how parents allocate resources within the family, the mechanisms that drive these strategies and the role they play at boosting or lessening heterogeneity in siblings' achievement. Indeed, previous studies that find differences in sibling correlation across subpopulations attribute such results to differences in within-family variation and speculate about the role of resource allocation at producing such heterogeneity (Conley and Glauber, 2005; Conley et al., 2007). Although this is a plausible explanation, it is not the only possible explanation.

Alternatively, differences in the sibling correlation across groups can arise from heterogeneity in between-family dispersion, pointing to different mechanisms of intergenerational transmission. Importantly, differences in between-family variance across subpopulations might indicate that there is a differential effect of families on children's outcomes (e.g. differential IGE) or that the distribution of relevant parental resources differs across subpopulations (see equation 4 in Section 2 for more details on this point).

An additional challenge when comparing sibling correlations has to do with the interpretation of differences across groups. Different levels of resemblance among siblings can arise from unobserved heterogeneity. That is, it is possible that the subpopulations differ in some unobserved dimension(s) that is correlated with the variance components incorporated in the sibling correlation, thus inducing differences across groups. In such cases, the meaning of the sibling correlation depends on the factor with which the variance components are correlated. For example, if a lower sibling correlation for one subpopulation is due to higher *idiosyncratic* individual heterogeneity – i.e., the critical assumption that makes the sibling correlation a social (im)mobility measure– it may indicate a less deterministic relationship between family background and children achievement. However, if larger within-family dispersion is due to correlation with an unobserved factor such as parental income, it would not indicate higher mobility but only another pathway through which social background affects social destination. To interpret the results of the sibling correlation as an indicator of social mobility it is, therefore, crucial to examine the drivers of heterogeneity both between and within families. Despite its relevance, these nuances tend to be overlooked in

[1] Throughout the article I will use the terms "subpopulations" and "groups" interchangeably.

studies that compare sibling correlations across groups, place or time.

In relation to the study of Black-White differences in sibling correlation, the caveats issued above imply that such differences can emerge from both within-family and between-family variance, each pointing to the action of different mechanisms. Moreover, Black-White differences may be due to race *per se* or to factors correlated with race, importantly the socioeconomic standing of families.

# 2 Heterogeneous Sibling Correlations

I develop an analytic framework that generalizes the sibling correlation in income to explicitly incorporate potential sources driving Black-White differences in intergenerational mobility. This approach has three main features: 1) it models Black-White differences in each of the variance components underlying the sibling correlation, 2) it decomposes Black-White differences into the fraction that is due to the effect of correlated observable factors (e.g. parental income) and Black-White gaps that remain after controlling for these factors, and 3) it embeds the intergenerational income elasticity as one of the potential drivers of Black-White differences in sibling correlation. It does so by treating each variance component as a stochastic function of race ($r$) and parental log-income ($y^p$), turning the sibling correlation into a function of two variance functions. Formally:

$$\rho_{ij} = \frac{\sigma_{a_j}^2(r_j, y_j^p)}{\sigma_{a_j}^2(r_j, y_j^p) + \sigma_{\mu_j}^2(r_j, y_j^p)} \tag{3}$$

Furthermore, the relationship between the sibling correlation and the intergenerational elasticity becomes clear if, following Solon et al. (2000), the family component of individual income (i.e., the average income of all children that belong to the same family) is expressed as the dependent variable in a standard intergenerational income mobility model. In this model income in the children's generations is described as a function of parental income plus an unexplained component that is orthogonal to the income of parents. Thus, as described in equation 5, by taking variances in both sides of equation 4, between-family variance becomes a function of three distinct factors: the intergenerational elasticity (squared), variance in parental log-income and variance in the residual error of this model. That is, *ceteris paribus*, Black-White differences in between-family variance can be due to (a) differences in the strength of association between children and parents' income, (b) differences in the extent of inequality in parental income and (c) differences in the extent to which factors uncorrelated with parental income affect a person's income. Moreover, because in my framework between-family variance is a function of race and parental income, it follows from equation 3 that each of the terms in the right-hand side of equation 5 is potentially a function of race and parental income.

$$a_j = \alpha + \beta_j y_j^p + \varepsilon_j \tag{4}$$

$$\sigma^2_{aj} = \beta^2_j \times \sigma^2_{y^p j} + \sigma^2_{\varepsilon j} \tag{5}$$

As such, this approach is well suited to answer the questions posited in this study. First, I can assess whether Black-White differences in social mobility reflect processes that take place within the family or between families by treating each variance component as a separate function of race. Second, by modeling the partial contribution of both race and parental income to each variance component I determine whether Black-White differences in social mobility reflect the socioeconomic characteristics of the two subpopulations or are, instead, due to other race-related factors. Finally, to resolve the conflicting diagnostic of a "perverse openness" among Black Americans and equality in the rates of social mobility of Whites and Blacks – as indicated by studies using the sibling correlation and the IGE, respectively – this framework integrates the IGE within the more general framework provided by heterogeneous sibling correlation by treating it as one component of the between-family variance function. Section 4 provides details on the procedure for parameter estimation.

# 3 Data and Measures

Data from the Panel Study of Income Dynamics (hereafter PSID) is used in this article. The PSID started in 1968 with a nationally representative sample of over 18,000 individuals living in 5,000 families in the United States. These individuals and their descendants were surveyed on a yearly basis until 1997 and biyearly since then. The PSID combines detailed information on kin relationships, rich sociodemographic data on both parents and children, and longitudinal records on income.

My original sample consists of 12,492 sons and daughters from families that were members of the original 1968 sample families (i.e., families in the Survey of Economic Opportunity or Survey Research Center samples) or moved into the sample. I restrict this sample to children of household heads who, as adults, became household heads or the legal spouse of a household head. I also restrict the sample to individuals born between 1952 and 1989. The 1952 restriction avoids over-representing children who left home after age 16. The 1989 restriction assures that the children's 2015 measures are observed when they are 25 at least.

My measure of individual income is labor income, which incorporates wages and salaries, the labor portion of farming and business income, and other sources of labor income (e.g., bonuses, overtime, tips, commissions). Even though the sample contains information about individuals across their life-course, the time span depends on their birth cohort. For example, I observe income from age 25 to 60 for the 1952 birth cohort, but only at age 25 for those born in 1989. This feature of the sample implies that, for younger birth cohorts, estimates of permanent income – the primary focus of this research– rely on fewer data points, with transitory income observed at earlier ages. For these reasons, my measure of children's income is the residual labor log-income after adjusting for life-cycle and general time effects, such as price-inflation and business-cycle (detail in Appendix 7.1. See (Solon et al., 1991) for a similar approach).

My measure of permanent parental income is a 10-year average, centered on parents' mid-career income. More specifically, I use the algorithm proposed by Mazumder (2014), which computes a running mean of the ten first income data point in the age range 35 to 50, starting at age 42. I apply this algorithm to mothers and fathers separately. Next, I define permanent parental income as the natural logarithm of sum of both parents "close-to-42" income (details in Appendix 7.1)

I determine sibship using the PSID's Family Identification Mapping System, which links children to their ancestors up to great-grandparents. The primary analyses rely on a sample of siblings, defined as individuals who have a common biological mother. By definition, the analysis of sibling correlations excludes individuals who do not have at least one sibling. Moreover, because of gender differences in terms of attachment to the labor market and earnings, I conduct the analyses separately for men and women. This implies that the analytic sample only contains individuals who have one or more same-sex siblings on the maternal side. In addition, these individuals must meet the conditions described above and have complete information on the variables included in the models (i.e., personal income in a given year, race and parental income). Finally, individual-year observations in which the person's income take on extremely low values (below the 0.5th percentile) are also excluded from the analyses. After applying these restrictions my analytic sample consists of 1,531 brothers nested into 651 mothers and 1,631 sisters nested into 686 mothers.

Table 1 compares descriptive statistics across the original sample and the analytic sample. Both samples are similar in terms of children's and parent's income averages, but differ in some important regards: in the analytic sample, the proportion of black individuals is lower than in the original sample (33% and 38%, respectively), and the same is observed with the proportion of black individual-year observation compared to whites (30% and 38%, respectively). Since sibship size is higher among black families than among whites, the higher attrition in the black sub-sample is most likely related to a higher rate of missingness in the income variables. Descriptive statistics also reveal an earlier timing of fertility for black mothers compared to their white counterparts. This difference is, however, similar in the original and the analytic sample. Table 6 in Appendix provides descriptive statistics for the analytic sample of brothers and sisters separately.

Table 1: Descriptive statistics

| | Full Sample | | | | Analytic Sample | | | |
|---|---|---|---|---|---|---|---|---|
| | Black | | White | | Black | | White | |
| Children's log-income | 10.01 | (1.01) | 10.41 | (1.02) | 9.98 | (0.96) | 10.45 | (0.96) |
| Parent's log-income near 42 | 10.31 | (0.79) | 11.00 | (0.77) | 10.13 | (0.79) | 10.94 | (0.76) |
| More than 2 siblings | 0.61 | (0.49) | 0.50 | (0.50) | 1.00 | (0.00) | 1.00 | (0.00) |
| More than 2 brothers | 0.29 | (0.45) | 0.23 | (0.42) | 0.59 | (0.49) | 0.66 | (0.47) |
| More than 2 sisters | 0.37 | (0.48) | 0.23 | (0.42) | 0.77 | (0.42) | 0.58 | (0.49) |
| Sibling age spread | 5.33 | (6.95) | 3.77 | (6.32) | 8.69 | (4.80) | 7.34 | (4.27) |
| Sibship size | 2.63 | (1.89) | 2.03 | (1.36) | 4.19 | (1.82) | 3.53 | (1.40) |
| Mother's age at first birth | 21.67 | (5.04) | 23.84 | (4.80) | 20.93 | (4.12) | 22.79 | (3.85) |
| Mother's age at birth | 24.43 | (5.19) | 26.15 | (4.74) | 24.33 | (4.58) | 25.53 | (3.92) |
| Number of person-years | 217,813 | | 349,713 | | 9,174 | | 20,231 | |
| Number of persons | 4,783 | | 7,709 | | 1,065 | | 2,097 | |
| Number of mothers | 2,753 | | 5,165 | | 379 | | 812 | |

# 4 Parameter estimation

This study implements heterogeneous sibling correlations as a mean to assess Black-White differences in intergenerational mobility and identify its driving sources. The analysis develops in two stages: First, I replicate findings from previous literature on sibling correlations for Black and White Americans. To estimate these correlations, I use hierarchical linear models with random intercepts via Restricted Maximum Likelihood estimation (RMLE). RMLE provides unbiased estimates of variance components and is, for this reason, the conventional approach in the sibling correlation literature (Björklund and Jantti, 2007) (details in Appendix 7.3).

Previous studies have investigated group differences in sibling correlation by estimating them separately for subpopulations of interest (e.g., (Conley and Glauber, 2005; Conley et al., 2007; Conley and Glauber, 2008; Erola and Jalovaara, 2016)). Such approach is appropriate for mere description but is insufficient if one wants to understand the sources of differences across groups. First, conducting analyses by subsample compromises statistical power. Second, investigating the partial contribution of different factors to differences in the sibling correlation (e.g., race and parental income) is often unfeasible because, in practice, only a limited number of partitions are possible (the "curse of dimensionality"). This limitation is particularly severe in the case of continuous predictors. Finally, raw group comparisons might mask countervailing trends. These might arise from the different behavior of between-family and within-family variances across groups, and from differences in the effects of covariates on variance components.

Second, to cope with these limitations, the second core stage of the analysis introduces a novel modeling approach to the study of heterogeneity in sibling correlations. This approach builds on the standard statistical framework underlying sibling correlations but, as described in section 2, treats the variance components not as fixed parameters but as random variables that can be modeled as functions of covariates – in this case, race and parental income. Furthermore, it embeds the intergenerational income elasticity within the sibling correlation by decomposing between-family variance into three components: (a) the association between children and parents' income or IGE (squared), (b) the variance in parental income and (c) the variance in factors that affect a person's income but are uncorrelated with parental income. Each of these components is modeled as functions of race and parental income.

Parameter estimation is a non-trivial aspect of this modeling approach, as standard models are not equipped for estimating hierarchical linear models with "organized dispersion" (or structured uncertainty). Standard statistical packages for mixed effects models do not permit modeling heteroskedastic variance components (e.g. lmer4 in R or xtmixed in Stata). Variance-function regression (Western, 2009) only allows modeling the residual variance in a single-level regression setting but not in a multilevel one.

Hence, in order to estimate the parameters of interest, I implement a Bayesian hierarchical linear model with "organized dispersion". This approach permits simultaneous modeling of all variance components in a mixed-effects model context[2]. A Bayesian approach is particularly useful in this case because it facilitates estimation and inference for a complex and flexible model structure.

All models were fit using the Hamiltonian Monte Carlo algorithm via the Bayesian modeling software Stan (Stan Development Team, 2016; Sorensen and Vasishth, 2015). Details on the model structure and implementation in Appendix 7.3. In addition, in order to compare results to those yielded by a frequentist estimation approach, I re-estimate all models using recently developed extensions of Hierarchical Generalized Linear Models (HGLM hereafter), which also permit simultaneous modeling of all sources of dispersion for mixed-effects models (Rönnegård et al., 2010). Results yielded by HGLM models are used as a robustness check, but they are interpreted with caution (especially inference) because the H-likelihood theory underlying these models remains controversial in the statistics community (Meng, 2009). Finally, to prevent the possibility that differences in income attrition and fertility described in the previous section drive part of the results, these models are fit using inverse weights that adjust for the estimated probability of being in the analytic sample. These probabilities are estimated separately for the samples of bothers and siblings and incorporate race, parental income and their interaction as predictors (see details in Appendix 7.3.3).

This modeling approach is well suited for the purpose of the present study. From a statistical point of view, this approach allows me to estimate Black-White differences in sibling correlation by treating race as an explanatory variable in variance models, thus avoiding the necessity of partitioning the sample and permitting a direct test of statistical significance.

[2]The model effectively recovered the "true" parameters in a Monte Carlo simulation (results upon request).

# 5 Findings

## 5.1 Analyses by sub-sample

Tables 2 and 3 report estimates of sibling correlations, as well as the variance components, for White and Black men and women. These results are based on the conventional the strategy of analyzing sub-samples. I report results from RMLEs from standard hierarchical linear models and Bayesian estimates from Bayesian hierarchical models. Point estimates and inference are highly comparable with either estimation approach[3] (HGLM estimates are in table 7 in the Appendix).

Table 2: Sibling correlation in permanent log-income for men

| | Black | | White | |
|---|---|---|---|---|
| | RMLE Estimate | Bayesian Estimate | RMLE Estimate | Bayesian Estimate |
| $\sigma_a^2$ | 0.16 (0.09-0.23) | 0.16 (0.10-0.23) | 0.13 (0.11-0.16) | 0.14 (0.11-0.17) |
| $\sigma_\mu^2$ | 0.51 (0.45-0.59) | 0.51 (0.44-0.59) | 0.19 (0.17-0.22) | 0.20 (0.17-0.22) |
| $\sigma_\nu^2$ | 0.48 (0.46-0.49) | 0.48 (0.46-0.49) | 0.26 (0.26-0.27) | 0.26 (0.26-0.27) |
| $\rho$ | 0.24 (0.14-0.32) | 0.24 (0.15-0.33) | 0.41 (0.34-0.47) | 0.41 (0.34-0.48) |

Table 3: Sibling correlation in permanent log-income for women

| | Black | | White | |
|---|---|---|---|---|
| | RMLE Estimate | Bayesian Estimate | RMLE Estimate | Bayesian Estimate |
| $\sigma_a^2$ | 0.22 (0.16-0.28) | 0.23 (0.16-0.30) | 0.18 (0.13-0.24) | 0.18 (0.13-0.25) |
| $\sigma_\mu^2$ | 0.45 (0.40-0.50) | 0.46 (0.40-0.52) | 0.48 (0.42-0.54) | 0.48 (0.42-0.55) |
| $\sigma_\nu^2$ | 0.53 (0.51-0.54) | 0.53 (0.51-0.54) | 0.52 (0.51-0.54) | 0.52 (0.51-0.54) |
| $\rho$ | 0.33 (0.26-0.40) | 0.33 (0.25-0.41) | 0.28 (0.20-0.35) | 0.28 (0.20-0.36) |

The estimated sibling correlation for White men is consistent with previous results for the United States, which fall in the 0.3 to 0.5 range with typical values around 0.4 (Solon, 1992; Björklund and Kjellström, 2002; Conley and Glauber, 2005; Mazumder, 2008). This is not surprising given that the vast majority of these studies are limited to the study of correlations in brothers income, using samples of mostly White individuals. Moreover, the few studies that focus on sister correlation report that, as I find here for White women, the correlation in income among sisters is substantially lower than among bothers, with estimates typically around 0.35 (Mazumder, 2008; Black and Devereux, 2011). A common explanation for gender differences in intergenerational mobility is women's weaker attachment to the labor market compared to men. Given men's average higher earnings and prevailing gender norms in the division household work and child-rearing, women typically have discontinuous labor trajectories, work fewer hours and earn lower salaries than men.

[3] In parentheses, I report 95% confidence intervals for RMLE and 95% credible intervals for Bayesian estimates. These intervals are estimated by the corresponding quantiles of the simulation distribution (or the posterior distribution in the Bayesian case) of each parameter. The reported Bayesian point estimate corresponds to the posterior mean of each parameter.

As a result, estimates of women's long-term income are noisier, inducing attenuation bias in IGEs and sister correlations. This argument is consistent with the finding that the sibling correlation is somewhat larger for Black sisters compared to their White counterparts. Black women are more likely than White women to be head of a single parent household, and therefore their labor supply is less elastic than that of White women.

In line with previous research, I also find that sibling correlations in income are substantively larger for White men compared to Black men, while a minor difference is observed among women. Moreover, differences for men arise from the comparatively higher within-family income variance among Black families compared to Whites, a result that is consistent with Conley and Glauber (2005)'s theorizing regarding the source of Black-White differences in the sibling correlation. By contrast, Black-White comparability in sisters correlation results from similarity in all variance component. Figure **??** plots the posterior distribution of the sibling correlation for Black and White men and women, illustrating the findings described above. Tables 8 to 11 in the Appendix reports comparable findings using race as a sole predictor in the variance models.

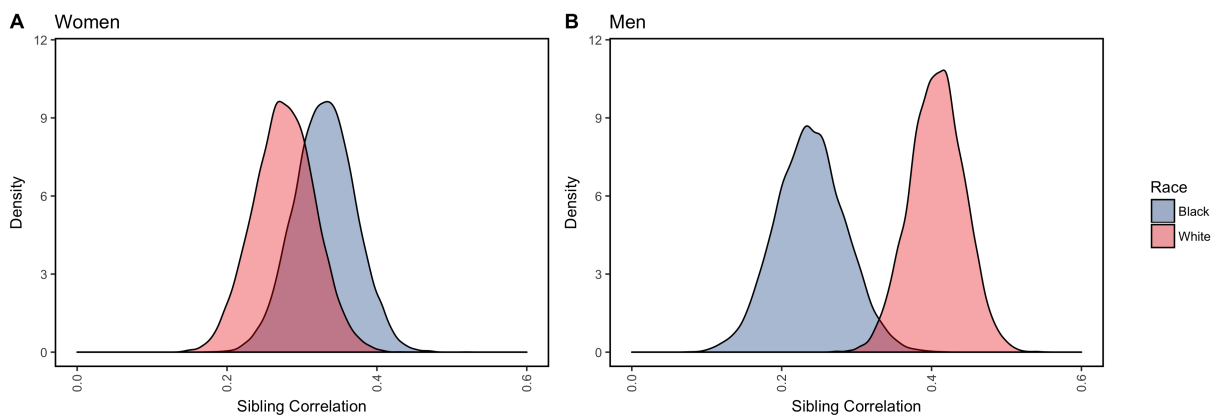


Figure 1: Posterior distribution of sibling correlation in long-term income

The meaning of these findings for social mobility critically depends on whether within-family variance reflects idiosyncratic heterogeneity among siblings or if, instead, such heterogeneity is structured along socioeconomic lines, thus informing of yet another pathway through which social background affects social destination. Similarly, the absence of Black-White differences in between-family variance does not necessarily indicate that the process of intergenerational transmission is the same for both groups, as similarity may mask countervailing effects in its different components (see equation 4). The next sections present findings from models that directly test these possibilities.

## 5.2 Findings from the heterogeneous sibling correlations model

My core empirical contribution, and the subject of this section, are results from hierarchical models that allow for organized dispersion (or structured variances) in all levels. These models treat between and within family variance (the components of the sibling correlation) as a function of race and parental income. I further decompose between-family variance into the IGE, variance in parental income and variance in other factors that affect children's income but are uncorrelated with the income of parents. Figures ?? and ?? summarize the findings of this analysis for men and women, respectively. Tables 4 and 5 summarize the estimates underlying these figures.

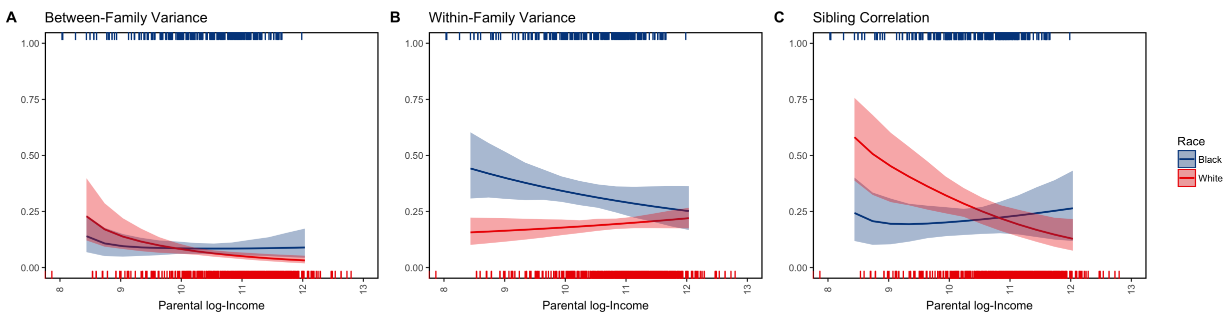


Figure 2: Posterior distribution of predicted variances and sibling correlation in long-term income for Men

Findings for men confirm the lower correlation in income among Black brothers compared to Whites – Panel C in figure ?? – a result that is only driven by differences in within-family variance across the two groups. Specifically, I find an interaction between race and parental income, where within-family variance is generally larger for Blacks compared to White brothers, but the gap declines as parental income increases. This closing Black-White gap in within-family variance with income is the result of different patterns for Black and Whites. For Blacks, I find a negative relationship between parental income and within-family variance. For Whites, within-family variance remains low and relatively stable along the entire socioeconomic spectrum of parents. In other words, while Black brothers of high-income parents are more alike in terms of adult income than Black brothers born to low-income parents, no such a difference is found among White brothers of different economic origins. Only for the children of very high-income parents I find no Black-White difference in within-family variance. These findings show no substantive difference in between-family variance by race or parental income.

Taken together, these patterns aggregate to produce a sizeable Black-White gap in sibling correlation. The correlation is low (about 0.2) and remains at the same level for Black brothers of all economic backgrounds. By contrast, the sibling correlation in income among White brothers is larger than for Blacks (about 0.6 at its highest value) but declines as parental income increases. There is no gap between White and Black children of high-income parents.

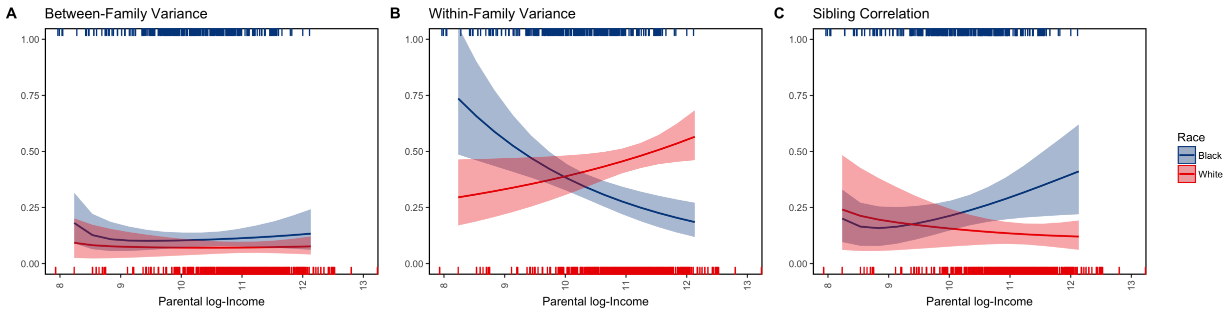


Figure 3: Posterior distribution of predicted variances and sibling correlation in long-term income for Women

Findings for women show that the similarity in sibling correlation among White and Black sisters follows from similarity in the extent and structure of between-family variance, combined with the opposite effect of parental income on within-family variation: while income dispersion among Black sisters sharply declines as parental income increases, within-family among White sisters increases with parental income. These opposite trends offset each other in the aggregate, producing a similar sibling correlations for most White and Black sisters. The exception is women born to high-income parents, where Black sisters show a higher sibling correlation than comparable White sisters. However, because high-income parents represent a relatively small fraction of the population – especially among Black families – this difference has little impact on the aggregated Black-White comparison.

These results provide partial support for the claim that Black-White differences in the process of intergenerational transmission reflect the socioeconomic standing of Black and White families. The fact that within-family income variance for Black brothers and sisters sharply declines as parental income increases suggests that part of the observed Black-White difference in sibling correlation is due to the fact that Black children are more likely than White children to belong to families located at the lower end of the income distribution, families whose offspring present larger dispersion in income as adults. Similarly, Black sisters born to affluent parents present a higher sibling correlation than comparable White sisters but this gap is concealed in the aggregate due to the small fraction of families – especially Black families – at the upper end of the income distribution.

The negative relationship between parental income and within-family variance found for Black men and women is consistent with the hypothesis that disadvantaged families engage in practices that reinforce initial disparities among their offspring, while better-off parents incur compensatory investments that blur out such disparities. This hypothesis is, however, challenged by the null association between parental income and within-family variance found for White men, and the positive relationship found for White women. Furthermore, the persistence of a Black-White gap among families of similar income levels indicates that socioeconomic differences can not solely explain the racial difference in sibling correlations. As such, understanding the sources of this Black-White gap constitutes a crucial endeavor for future research on intergenerational transmission. It remains unclear, in particular, if the comparatively large within-family heterogeneity among Blacks is an indication of individual idiosyncrasies or if it, instead, arises from unobserved structural sources that make it more difficult for Black families to affect the economic fate of their offspring.

## 5.3 Decomposition of between-family variance

Findings indicate that between-family variance, the other element of the sibling correlation, cannot explain Black-White differences in siblings correlation because this variance component does not vary by race or parental income – Panel A in figures **??** and **??**. Nevertheless, the absence of differences might be due from countervailing effects in the components of between-family variance. As described in equation 5, the between-family variance can be expressed in terms of the elements in a standard intergenerational income mobility model. In particular, it can be shown that between-family variance depends on three factors: the intergenerational elasticity ($\beta$), the variance in parental log-income ($\sigma^2_{yp}$) and the variance of the residual error ($\sigma^2_{\varepsilon}$). Variation in between-family variance can thus arise from any of these sources, each pointing to different mechanisms of intergenerational transmission. Furthermore, because in my framework between-family variance is treated as a function of race and parental income, each of these terms is potentially a function of race and parental income[4].

Table 4: Bayesian Variance model for permanent log-income (Men)

| | $\beta$ | log-$\sigma^2_{y^p}$ | log-$\sigma^2_e$ | log-$\sigma^2_\mu$ | log-$\sigma^2_\nu$ |
|---|---|---|---|---|---|
| Intercept | 0.24 (0.03) | 29.74 ( 4.94) | -1.94 ( 2.59) | 0.54 ( 1.20) | 1.31 ( 0.31) |
| White | -0.01 (0.01) | 2.80 ( 5.35) | 4.94 ( 3.17) | -3.40 ( 1.56) | -1.79 ( 0.38) |
| Parents's logY | | -3.58 ( 0.58) | -0.06 ( 0.25) | -0.16 ( 0.12) | -0.22 ( 0.03) |
| White * Parents's logY | | -0.35 ( 0.63) | -0.49 ( 0.30) | 0.27 ( 0.15) | 0.15 ( 0.04) |

Table 5: Bayesian Variance model for permanent log-income (Women)

| | $\beta$ | log-$\sigma^2_{y^p}$ | log-$\sigma^2_e$ | log-$\sigma^2_\mu$ | log-$\sigma^2_\nu$ |
|---|---|---|---|---|---|
| Intercept | 0.26 (0.04) | 28.48 ( 4.40) | -3.18 ( 2.29) | 2.59 ( 1.16) | -0.40 ( 0.29) |
| White | -0.02 (0.01) | 0.70 ( 6.90) | 0.10 ( 3.01) | -5.31 ( 1.60) | 0.10 ( 0.36) |
| Parents's logY | | -3.45 ( 0.52) | 0.09 ( 0.22) | -0.36 ( 0.11) | -0.03 ( 0.03) |
| White * Parents's logY | | -0.64 ( 2.31) | -0.05 ( 0.28) | 0.53 ( 0.15) | -0.01 ( 0.03) |

The first contributor to between-family variance is the IGE, that is, the conditional expectation of children's income given their parents' income. I test whether the IGE differs for Blacks and Whites and whether the relationship between children's and parents' income is strictly linear, or it varies with parental income. Figure **??** illustrates the results of this analysis by race and gender. Each

[4] I also estimate these models in "reduced-form", that is, without imposing a structural form to between-family variance. Instead, I model between-family variance directly, as a function of race, parental income and their interaction. Tables 12 and 12 report Bayesian estimates of these models, and tables 13 and 13 show HGLM estimates.

plot represents the association between the family component of individual income and parental income, where the straight line corresponds to the estimated intergenerational elasticity, and the dashed line is a cubic spline that allows for a nonlinear relationship. These results yield intergenerational elasticities of about 0.25 and suggest that they do not vary across Black and White men, or across Black and White women. Similarly, the spline describes a nearly linear relationship in the regions of high density but departs from linearity elsewhere.

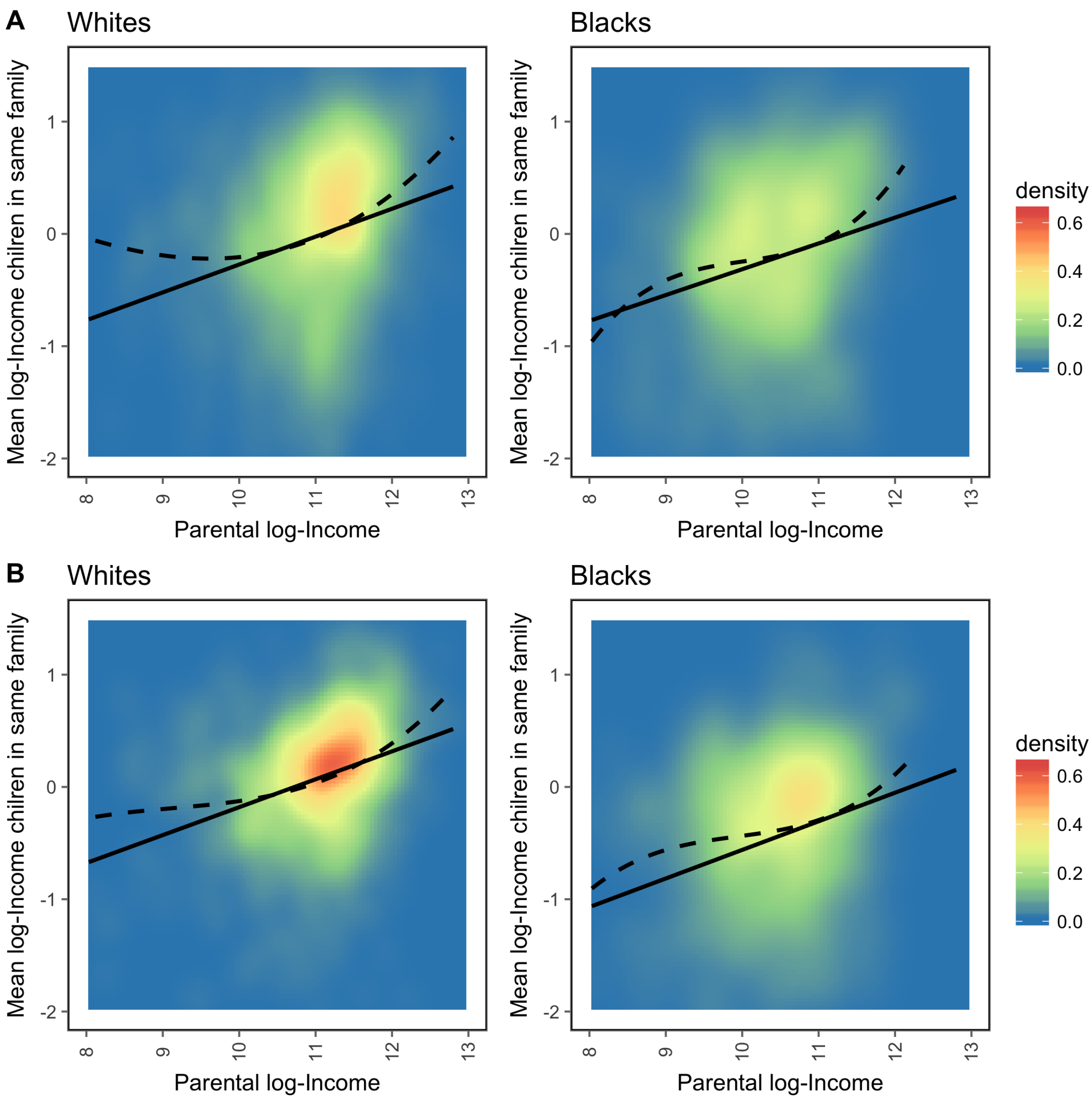


Figure 4: Association between children and parents log-Income. **A:** Women, **B:** Men. Straight line corresponds to the OLS regression line and dashed line is a cubic spline. Areas in red and yellow indicate areas of the joint distribution where most cases are concentrated.

A second contributor to between-family variance is the income variance in the parental generation. This term expresses the fact that, for any given (positive) level of association between parents' and children's income, the higher the income inequality in the parental generation, the higher inequality in the children's generation. I test whether variation in parental income differs by race and by parental income itself (heteroskedastic variance). Figure ?? summarizes the findings of this analysis by race and gender. The results indicate that inequality in parental income is higher at the bottom of the income distribution and, in that region, is somewhat larger among Black parents than White ones. Variance is lower elsewhere and does not vary by race or parental income level.

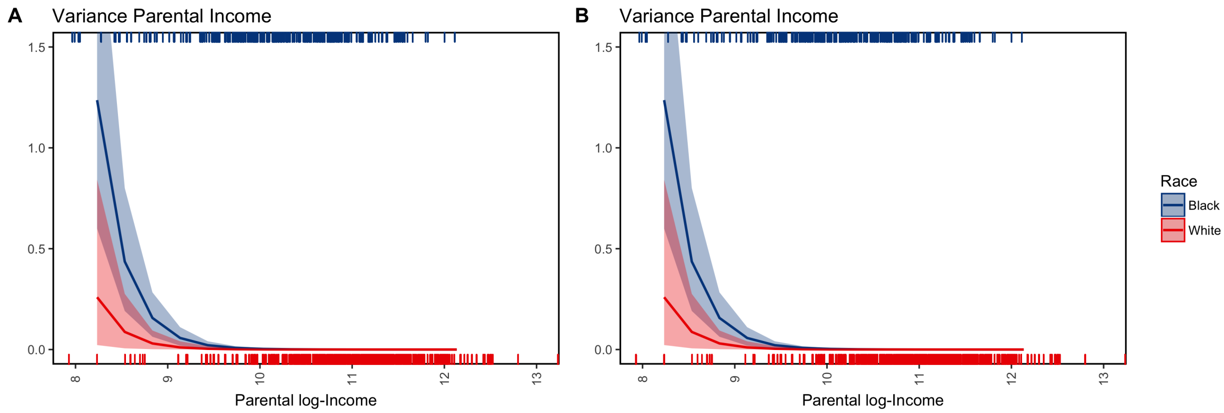


Figure 5: Variance in parents' income as a function of race and parents' income. **A:** Women, **B:** Men.

The third component of between-family variance in income is the residual variance in intergenerational income mobility model. This term captures the contribution of all factors that affect the family component of individual income but are uncorrelated with parental income, be these unobserved predictors of children's income or idiosyncratic characteristics. It is important to stress that in the latter case such idiosyncrasies are not attributes of individuals (captured as within-family variance) but family-level idiosyncratic characteristics. In any case, the magnitude of residual variation indicates the extent to which parental income does not predict the family component of individual income. To evaluate its contribution to heterogeneity in between-family variance, I test whether residual variance differs by race and by parental income. Figure **??** plots the results of this analysis by race and gender. These findings suggest that parental income has a comparable predictive power on the income of Black and White men of all socioeconomic origins. For White women but not for Black women, the residual variance decreases as parental income increases. No clear racial gap is found at any level of parental income.

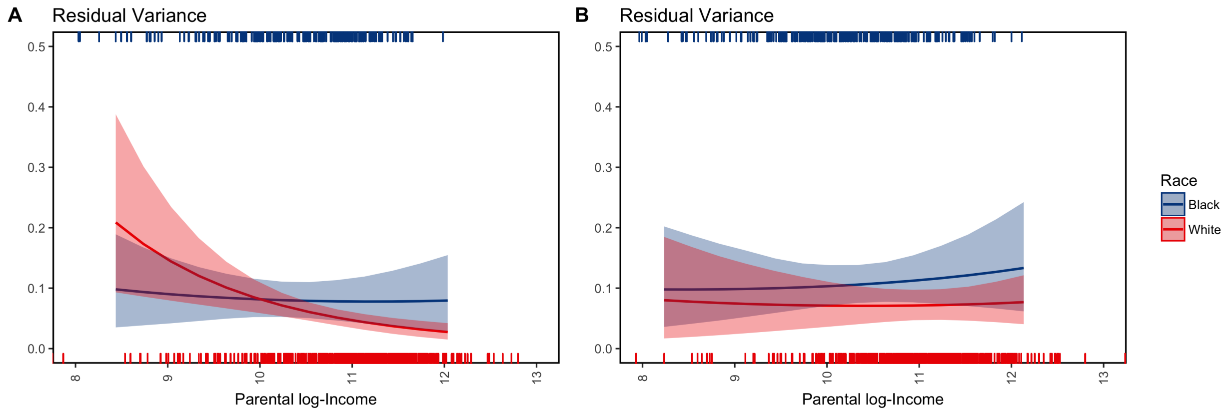


Figure 6: Residual variance as a function of race and parents' income. **A:** Women, **B:** Men.

Taken together, these results shed light on the factors driving the lack of heterogeneity in between-family variance, that is, dispersion in the family component of individual income. I find that all three components contribute to the similarity in between-family variance as they do not differ across races nor across different levels of parental income. Especially relevant is the results for the intergenerational elasticity, which confirm a common linear IGE of about 0.25 for Blacks and Whites and for men and women. This finding is consistent with previous research reporting a similar IGE for the cohorts of Black and White men born in the early 1960s (Bloome and Western, 2011).

# 6 Discussion

This article revisits two long-standing empirical puzzles in the sociological literature on intergenerational mobility: first, whether Blacks experience a "perverse openness" with respect to mobility and, second, whether such openness (or the absence thereof) is rooted in the disparate socioeconomic standings of Blacks and Whites. To answer these questions, I develop an analytic framework – heterogeneous sibling correlations – that allows me to incorporate different facets of the mobility process. In doing so, I propose a comprehensive and direct answer to these long-standing puzzles.

If by immobility, one means the strength of the association between children's and parents' income, then my results suggest no such openness. To the contrary, I find that intergenerational income elasticity does not vary across Blacks and Whites, regardless of gender. Similarly, the association between parents' and children's income is of similar magnitude for families of all economic origins. However, if one understands mobility as uncertainty - that is, the share in one's income is not associated with one's social background - then I find clear evidence of a "perverse openness" among Blacks compared to Whites, driven by the substantial heterogeneity in income among Black brothers and, to a lesser extent, among Black sisters. These findings reconcile the seemingly contradictory results of previous studies that reported both similar IGE for Blacks and Whites (Bloome and Western, 2011) but lower sibling correlation for the former (Conley and Glauber, 2005).

The significance of these results for our understanding of the mobility process critically depends on whether the comparatively large within-family variance found for Blacks arises from idiosyncratic factors or if it, instead, is socioeconomically "structured". In this regard, I find that both race and economic origins drive the perverse openness among Blacks: while Black-White gaps in within-family heterogeneity tend to close as parental income increases, racial gaps exist among children of similar economic origins. Overall, these results suggest no simple answer to the debate about racial and socioeconomic explanations for Black-White differences in mobility, as both race and economic origins interact to create racial differences in social mobility. More research is needed to determine the drivers of these unexplained racial gaps.

# 7 Appendix

## 7.1 Income measures

### 7.1.1 Individual income

My income measure is the residuals from an OLS regression that purges the effect of age and period from log-income. Formally:

$$y_{ijt} = y^*_{ijt} - X'_{ijt}\beta \tag{6}$$

where $y^*_{ijt}$ is the natural logarithm of labor income of the individual $i$ of family $j$ in year $t$, transformed to 2015 dollars. $X$ is a vector of covariates (age, age-squared and year) and $\beta$ is a vector of regression coefficients.

### 7.1.2 Parental income

I use the algorithm proposed by Mazumder (2014), which starts with an individual's income at the age of 42. If income is recorded in that year, such data is added to a running sum or ignored otherwise. The algorithm continues searching for income data around the age of 42, favoring younger ages. Therefore, it searches income data at 41, then 43, 40 and so on. The algorithm searches for income data on the age range 30 to 55 and stops adding data to the running sum once a total of 10 data points is reached. Individuals for which fewer than 10 data points are recorded between the ages of 30 to 55 are excluded from the analysis. I apply this procedure to both mothers and fathers separately (or the one for which information exists). Next, I create a measure of permanent parental combining the permanent income of mother and father ($y^m$ and $y^f$, respectively). Formally:

$$y^p_j = ln(y^m_j 1\{y^m_j > 0\} + y^f_j 1\{y^f_j > 0\}) \tag{7}$$

In the case of single parent households this measurement is equivalent to the 'close-to-42" income of the parent present in the household.

## 7.2 Descriptive statistics

Table 6: Descriptive statistics

| | Brothers | | | | Sisters | | | |
|---|---|---|---|---|---|---|---|---|
| | Black | | White | | Black | | White | |
| Children's log-income | 10.19 | (0.91) | 10.74 | (0.81) | 9.83 | (0.96) | 10.10 | (1.02) |
| Parent's log-income near 42 | 10.17 | (0.82) | 10.90 | (0.77) | 10.11 | (0.77) | 10.99 | (0.74) |
| Sibling age spread | 8.76 | (4.94) | 7.27 | (4.33) | 8.64 | (4.70) | 7.42 | (4.20) |
| Sibship size | 4.33 | (1.99) | 3.51 | (1.40) | 4.10 | (1.69) | 3.57 | (1.39) |
| Mother's age at first birth | 21.45 | (4.34) | 22.79 | (3.73) | 20.58 | (3.93) | 22.79 | (3.98) |
| Mother's age at birth | 24.86 | (4.84) | 25.45 | (3.73) | 23.97 | (4.36) | 25.63 | (4.14) |
| Number of person-years | 3754 | | 11051 | | 5420 | | 9180 | |
| Number of persons | 451 | | 1080 | | 614 | | 1017 | |
| Number of mothers | 189 | | 462 | | 249 | | 437 | |

## 7.3 Parameter estimation

### 7.3.1 Sibling correlations

Consider a person's income as a linear combination of both family and individual factors. In particular, let $y_{ijt}$ be the natural logarithm of income for the $ith$ sibling in family $j$ at time $t$, such that:

$$\begin{aligned} y_{ijt} &= a_j + \mu_{ij} + \nu_{ijt} \text{ where} \\ a_j &\sim N(\alpha, \sigma_a^2) \\ \mu_{ij} &\sim N(a_j, \sigma_\mu^2) \\ \nu_{ijk} &\sim N(\mu_{ij}, \sigma_\nu^2) \end{aligned} \tag{8}$$

Here $a_j$ is the family component to individual income, $\mu_{ij}$ is an individual component to income, and $\nu_{ijt}$ represent yearly in individual income. $\sigma_\nu^2$, $\sigma_\mu^2$ and $\sigma_a^2$ are variances corresponding to each component, which I estimate via REML in a multilevel model. Under the assumption of orthogonality between components, the sibling correlation in permanent log-income can be estimated as follows:

$$\hat{\rho} = \frac{\hat{\sigma}_a^2}{\hat{\sigma}_a^2 + \hat{\sigma}_\mu^2} \tag{9}$$

#### 7.3.2 Bayesian hierarchical linear model with "organized dispersion"

Let $y_{ijt}$ be the log-income for the $ith$ sibling in family $j$ at time $t$, which can be described as a linear combination of both family and individual factors. Unlike in the standard setting described above, now all variance components are allowed to be heteroskedastic and depend on (potentially different) variance covariates. In addition, the family component of individual income ($a$) is decomposed into the part explained by parental income and the part independent of it. Formally,

$$
\begin{aligned}
y_{ijt} &= a_j + \mu_{ij} + \nu_{ijt} \\
y_{ijt} &= \alpha + \beta_j y_j^p + \varepsilon_j + \mu_{ij} + \nu_{ijt} \text{ where} \\
\beta_j &= \beta_0 + \beta_1 r_j \text{ and} \\
a_j &\sim N(\alpha + \beta_j y_j^p, \beta_j^2 \sigma_{y^p j}^2 + \sigma_{\varepsilon j}^2) \\
\mu_{ij} &\sim N(a_j, \sigma_{\mu ij}^2) \\
\nu_{ijk} &\sim N(\mu_{ij}, \sigma_{\nu ijk}^2)
\end{aligned}
\tag{10}
$$

In the current study, all dispersion models use the same set of variance covariates: race, parental income and the interaction between the two, all of which are measured at the family level. Here $X$ is a vector of covariates in the variance models and $\delta$, $\gamma$, $\lambda$ and $\zeta$ are the corresponding regression coefficients. Formally:

$$
\begin{aligned}
\hat{\sigma}_{y^p j} &= exp(X_j^{'}\delta) \\
\hat{\sigma}_{\varepsilon j} &= exp(X_j^{'}\gamma) \\
\hat{\sigma}_{\mu ij} &= exp(X_j^{'}\lambda) \\
\hat{\sigma}_{\nu ijk} &= exp(X_j^{'}\zeta)
\end{aligned}
\tag{11}
$$

$\sigma$ parameters are modeled in a log scale because they are strictly positive. Consequently, standard deviations are assumed to follow a log-normal distribution with mean $\hat{\sigma}$ and dispersion $\tau$.

$$
\begin{aligned}
\sigma_{y^p j}^2 &\sim LN(\hat{\sigma}_{y^p j}, \tau_{y^p}) \\
\sigma_{\varepsilon j}^2 &\sim LN(\hat{\sigma}_{\varepsilon j}, \tau_{\varepsilon}) \\
\sigma_{\mu ij}^2 &\sim LN(\hat{\sigma}_{\mu ij}, \tau_{\mu}) \\
\sigma_{\nu ijk}^2 &\sim LN(\hat{\sigma}_{\nu ijk}, \tau_{\nu})
\end{aligned}
\tag{12}
$$

Equation 13 describes the prior distributions for the parameters in the model. All parameters for the mean-part of variance models (vectors $\delta$, $\gamma$, $\lambda$ and $\zeta$) were given a prior distributed $Cauchy(0, 2.5)$. These priors are chosen following Gelman et al. (2008) and constitute so-called weakly informative priors, as they assign similar probabilities to all plausible values of the parameter but very low probabilities to extreme values. Substantively such prior distribution represents conservative prior beliefs, in the sense that they "let the data speak" but rule out unreasonable parameter values. The dispersion parameters in the variance models – $\tau_{y^p}$,$\tau_{\varepsilon}$, $\tau_{\mu}$ and $\tau_{\nu}$– are assumed to distribute $Gamma(1, 1)$. These parameters represent uncertainty around variance estimates.

For each model, I run four parallel chains for 7000 HMC iterations. I assess Markov chains' convergence by inspecting trace plots and Gelman-Rubin potential scale reduction factor (Gelman and Rubin, 1992). I evaluate auto-correlation in the Markov chains and the consequent loss in effective sample size via visual inspection of auto-correlation plots and by computing the ratio of effective to total sample size.

$$
\begin{aligned}
\delta &\sim Cauchy(0, 2.5I) \\
\gamma &\sim Cauchy(0, 2.5I) \\
\lambda &\sim Cauchy(0, 2.5I) \\
\zeta &\sim Cauchy(0, 2.5I) \\
\tau_{y^p} &\sim Gamma(1, 1) \\
\tau_{\varepsilon} &\sim Gamma(1, 1) \\
\tau_{\mu} &\sim Gamma(1, 1) \\
\tau_{\nu} &\sim Gamma(1, 1)
\end{aligned}
\tag{13}
$$

#### 7.3.3 Weighting

I implement inverse probability weighting by multiplying the log-posterior of each observation with its respective weights. The construction of these weights is described in equation 14:

$$
w_i = \frac{N}{\sum_i \frac{1}{\hat{p}_i}} \frac{1}{\hat{p}_i} \tag{14}
$$

where $\hat{p}$ is the predicted probability of belonging to the analytic sample, estimated with a logistic regression which predictors are race, parental income and the interaction of these two variables. The model is estimated separately to predict belonging to the brothers' sample and the sisters' sample.

## 7.4 Additional findings

Table 7: Sibling correlation in permanent log-income, HGLM estimates

| | Men | | Women | |
|---|---|---|---|---|
| | Black | White | Black | White |
| $\sigma_a^2$ | 0.18 (0.14-0.22) | 0.14 (0.12-0.16) | 0.23 (0.19-0.28) | 0.20 (0.17-0.24) |
| $\sigma_\mu^2$ | 0.49 (0.44-0.56) | 0.19 (0.17-0.21) | 0.44 (0.40-0.49) | 0.46 (0.42-0.51) |
| $\sigma_\nu^2$ | 0.48 (0.46-0.49) | 0.26 (0.26-0.27) | 0.53 (0.51-0.54) | 0.52 (0.51-0.54) |
| $\rho$ | 0.26 (0.21-0.32) | 0.42 (0.38-0.47) | 0.34 (0.30-0.39) | 0.30 (0.26-0.35) |

Table 8: Variance model for permanent log-income (Men)

| | HGLM Estimates | | | Bayesian Estimates | | |
|---|---|---|---|---|---|---|
| | log-$\sigma_a^2$ | log-$\sigma_\mu^2$ | log-$\sigma_\nu^2$ | log-$\sigma_a^2$ | log-$\sigma_\mu^2$ | log-$\sigma_\nu^2$ |
| Intercept | -1.85 ( 0.12) | -0.85 ( 0.06) | -0.87 ( 0.02) | -1.93 ( 0.21) | -0.83 ( 0.08) | -0.87 ( 0.02) |
| White | -0.10 ( 0.15) | -0.74 ( 0.08) | -0.37 ( 0.02) | -0.03 ( 0.24) | -0.75 ( 0.10) | -0.37 ( 0.02) |

Table 9: Variance model for permanent log-income (Women)

| | HGLM Estimates | | | Bayesian Estimates | | |
|---|---|---|---|---|---|---|
| | log-$\sigma_a^2$ | log-$\sigma_\mu^2$ | log-$\sigma_\nu^2$ | log-$\sigma_a^2$ | log-$\sigma_\mu^2$ | log-$\sigma_\nu^2$ |
| Intercept | -1.50 ( 0.10) | -0.81 ( 0.05) | -0.64 ( 0.02) | -1.50 ( 0.10) | -0.81 ( 0.05) | -0.64 ( 0.02) |
| White | -0.15 ( 0.14) | 0.06 ( 0.07) | -0.01 ( 0.02) | -0.15 ( 0.14) | 0.06 ( 0.07) | -0.01 ( 0.02) |

Table 10: Variance model for permanent log-income (Men)

| | Blacks | | Whites | |
|---|---|---|---|---|
| | HGLM Estimate | Bayesian Estimate | HGLM Estimate | Bayesian Estimate |
| $\sigma_a^2$ | 0.16 (0.12-0.19) | 0.15 (0.09-0.21) | 0.15 (0.08-0.25) | 0.14 (0.11-0.18) |
| $\sigma_\mu^2$ | 0.43 (0.38-0.48) | 0.44 (0.37-0.51) | 0.20 (0.16-0.27) | 0.21 (0.18-0.23) |
| $\sigma_\nu^2$ | 0.42 (0.41-0.44) | 0.42 (0.40-0.44) | 0.29 (0.27-0.32) | 0.29 (0.28-0.30) |
| $\rho$ | 0.27 (0.22-0.32) | 0.25 (0.16-0.35) | 0.42 (0.28-0.56) | 0.41 (0.33-0.48) |

Table 11: Variance model for permanent log-income (Women)

| | Blacks | | Whites | |
|---|---|---|---|---|
| | HGLM Estimate | Bayesian Estimate | HGLM Estimate | Bayesian Estimate |
| $\sigma_a^2$ | 0.22 (0.18-0.27) | 0.22 (0.16-0.29) | 0.20 (0.12-0.31) | 0.17 (0.12-0.24) |
| $\sigma_\mu^2$ | 0.45 (0.40-0.49) | 0.46 (0.40-0.51) | 0.48 (0.37-0.60) | 0.48 (0.43-0.55) |
| $\sigma_\nu^2$ | 0.53 (0.51-0.55) | 0.53 (0.51-0.54) | 0.53 (0.49-0.56) | 0.52 (0.51-0.54) |
| $\rho$ | 0.33 (0.29-0.39) | 0.33 (0.25-0.40) | 0.29 (0.19-0.41) | 0.26 (0.19-0.35) |

Table 12: Bayesian Variance model for permanent log-income (Men)

| | log-$\sigma_a^2$ | log-$\sigma_\mu^2$ | log-$\sigma_\nu^2$ |
|---|---|---|---|
| Intercept | -0.24 ( 2.27) | 0.47 ( 1.25) | 1.32 ( 0.32) |
| White | 3.44 ( 2.61) | -4.11 ( 1.63) | -1.77 ( 0.38) |
| Parents's logY | -0.14 ( 0.22) | -0.15 ( 0.12) | -0.22 ( 0.03) |
| White * Parents's logY | -0.34 ( 0.25) | 0.34 ( 0.15) | 0.15 ( 0.04) |

Table 13: HGLM Variance model for permanent log-income (Men)

| | log-$\sigma_a^2$ | log-$\sigma_\mu^2$ | log-$\sigma_\nu^2$ |
|---|---|---|---|
| Intercept | -1.49 ( 1.95) | 2.01 ( 1.06) | -1.79 ( 0.30) |
| White | 4.60 ( 2.31) | -5.14 ( 1.33) | -0.27 ( 0.36) |
| Parents's logY | -0.02 ( 0.19) | -0.31 ( 0.10) | 0.11 ( 0.03) |
| White * Parents's logY | -0.45 ( 0.22) | 0.45 ( 0.13) | -0.04 ( 0.04) |

Table 14: Bayesian Variance model for permanent log-income (Women)

| | log-$\sigma_a^2$ | log-$\sigma_\mu^2$ | log-$\sigma_\nu^2$ |
|---|---|---|---|
| Intercept | -4.72 ( 2.53) | 2.68 ( 1.22) | -0.37 ( 0.29) |
| White | -0.74 ( 3.39) | -5.44 ( 1.74) | 0.08 ( 0.37) |
| Parents's logY | 0.31 ( 0.24) | -0.36 ( 0.12) | -0.03 ( 0.03) |
| White * Parents's logY | 0.02 ( 0.31) | 0.55 ( 0.16) | 0.00 ( 0.03) |

Table 15: HGLM Variance model for permanent log-income (Women)

| | log-$\sigma_a^2$ | log-$\sigma_\mu^2$ | log-$\sigma_\nu^2$ |
|---|---|---|---|
| Intercept | -5.23 ( 1.78) | 4.52 ( 0.93) | -3.74 ( 0.26) |
| White | 1.00 ( 2.34) | -7.64 ( 1.24) | 2.87 ( 0.35) |
| Parents's logY | 0.36 ( 0.17) | -0.54 ( 0.09) | 0.30 ( 0.03) |
| White * Parents's logY | -0.13 ( 0.22) | 0.76 ( 0.12) | -0.28 ( 0.03) |

# References


Behrman, J. R., Pollak, R. A., and Taubman, P. (1982). Parental Preferences and Provision for Progeny. *Journal of Political Economy*, 90(1):52–73.

Bhattacharya, D. and Mazumder, B. (2011). A nonparametric analysis of black-white differences in intergenerational income mobility in the United States. *Quantitative Economics*, 2(3):335–379.

Björklund, A. and Jantti, M. (2007). Intergenerational income mobility and the role of family background. *Handbook of Economic Inequality*, pages 1–66.

Björklund, A. and Kjellström, C. (2002). Estimating the return to investments in education: How useful is the standard Mincer equation? *Economics of Education Review*, 21:195–210.

Black, S. E. and Devereux, P. J. (2011). Chapter 16 - Recent Developments in Intergenerational Mobility. volume 4, Part B of *Handbook of Labor Economics*, pages 1487–1541. Elsevier.

Blau, P. M. and Duncan, O. D. (1967). American Occupational Structure.

Bloome, D. and Western, B. (2011). Cohort Change and Racial Difference in Educationl and Income Mobility. *Social Forces*, 90(2):375–395.

Conley, D. and Glauber, R. (2005). Sibling Similarity and Difference in Socioeconomic Status: Life Course and Family Resource Effects. *NBER Working Paper*.

Conley, D. and Glauber, R. (2008). All in the family?. Family composition, resources, and sibling similarity in socioeconomic status. *Research in Social Stratification and Mobility*, 26:297–306.

Conley, D., Pfeiffer, K. M., and Velez, M. (2007). Explaining sibling differences in achievement and behavioral outcomes: The importance of within- and between-family factors. *Social Science Research*, 36(3):1087–1104.

Erola, J. and Jalovaara, M. (2016). The Replaceable: The Inheritance of Paternal and Maternal Socioeconomic Statuses in Non-Standard Families. *Social Forces*, 95(3):971–995.

Featherman, D. L. and Hauser, R. M. (1976). Changes in the Socioeconomic Stratification of the Races, 1962-73. *American Journal of Sociology*, 82(3):621–651.

Gelman, A., Jakulin, A., Pittau, M. G., and Su, Y. S. (2008). A weakly informative default prior distribution for logistic and other regression models. *Annals of Applied Statistics*, 2(4):1360–1383.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4):457–472.

Hout, M. (1984). Occupational Mobility of Black Men : 1962 to 1973 OCCUPATIONAL MOBILITY OF BLACK MEN : 1962 to 1973 *. *American Sociological Review*, 49(3):308–322.

Manduca, R. (2018). Income Inequality and the Persistence of Racial Economic Disparities. *Sociological Science*, 5:182–205.

Mazumder, B. (2008). Sibling similarities and economic inequality in the US. *Journal of Population Economics*, 21(3):685–701.

Mazumder, B. (2011). Family and Community Influences on Health and Socioeconomic Status: Sibling Correlations Over the Life Course. *The B.E. Journal of Economic Analysis & Policy*, 11(3).

Mazumder, B. (2014). Black-White Differences in Intergenerational Economic Mobility in the US. *Economic Perspectives*, XXXVIII(1).

Meng, X.-L. (2009). Decoding the H-likelihood. *Statistical Science*, 24(3):280–293.

Rönnegård, L., Shen, X., and Alam, M. (2010). Hglm: A package for fitting hierarchical generalized linear models. *The R Journal*, 2(2):20–28.

Solon, G. (1992). Intergenerational Income Mobility in the United States. *The American Economic Review*, 82(3):393–408.

Solon, G., Corcoran, M., Gordon, R., and Laren, D. (1991). A Longitudinal Analysis of Sibling Correlations in Economic Status. *The Journal of Human Resources*, 26(3):509–534.

Solon, G., Page, M. E., and Duncan, G. J. (2000). Correlations between Neighboring Children in Their Subsequent Educational Attainment. *Review of Economics and Statistics*, 82(August):383–392.

Sorensen, T. and Vasishth, S. (2015). Bayesian linear mixed models using Stan: A tutorial for psychologists, linguists, and cognitive scientists.

Stan Development Team (2016). Rstan: the R interface to Stan.

Warren, J. R., Sheridan, J. T., and Hauser, R. M. (2002). Occupational Stratification across the Life Course: Evidence from the Wisconsin Longitudinal Study. *American Sociological Review*, 67(3):432.

Western, B. (2009). Variance Function Regressions for Studying Inequality.

Wilson, W. J. (1978). The Declining Significance of Race ch7.pdf.