I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

The Intergenerational Transmission of Health in the United States: A Latent Variables Analysis

Timothy J. Halliday
Bhashkar Mazumder
Ashley Wong

The Intergenerational Transmission of Health in the United States: A Latent Variables Analysis

Timothy J. Halliday
University of Hawai'i at Mānoa and IZA

Bhashkar Mazumder
Federal Reserve Bank of Chicago

Ashley Wong
Northwestern University

ABSTRACT

The Intergenerational Transmission of Health in the United States: A Latent Variables Analysis

Social scientists have long documented that many components of socioeconomic status such as income and education have strong ties across generations. However, health status, arguably a more critical component of welfare, has largely been ignored. We fill this void by providing the first estimates of the Intergenerational Health Association (IHA) that are explicitly based on a non-linear latent variable model. We develop an estimation procedure for a non-linear model with categorical outcomes in which the latent index is a mixed-linear model and contains covariates that might not vary within cross-sectional units. Adjusting for only age and gender, we estimate an IHA of 0.3 indicating that about one third of a parent's health status gets transmitted to their children. Once we add additional mediators to the model, we show that education, and particularly children's education, is an important transmission channel in that it reduces the IHA by one third. Finally, we show that estimates of the IHA from non-linear models are only moderately higher than those from linear models, while rank-based mobility estimates are identical.

Corresponding author:
Timothy J. Halliday
University of Hawaii at Mānoa
2424 Maile Way
Saunders Hall 533
Honolulu, HI 96822
USA

E-mail: halliday@hawaii.edu

1 Introduction

In recent years, there has been a growing focus on intergenerational mobility in socioeconomic status. By and large, the motivation of this literature has been to elucidate the extent to which there is equality of opportunity. This literature has largely focused on mobility in earnings and education. However, health status has been largely ignored. This is unfortunate since, according to many, health is one of the most important components of human welfare (*e.g.* Jones and Klenow (2016)). Moreover, health status in early childhood has been shown to mediate many important economic outcomes in adulthood (see Case et al. (2005), for example). Consequently, the combination of a high association between health and economic status in conjunction with a high intergenerational transmission of health status has potentially important implications for the transmission of economic status across generations. However, despite the importance of the topic, there is a paucity of work on the intergenerational transmission of health status.

The primary reason for the relative dearth of studies on the intergenerational transmission of health status vis-à-vis economic status is that health is relatively harder to measure. It is certainly true that income, education, and health are all measured with error. However, we would argue that the degree of measurement error is more prominent in health status. The reason for this is that many health measures are blunt proxies for some underlying latent variable. In principle, either income or education could be perfectly measured but they typically are not due to recall or rounding errors. This is particularly the case with income when it is either top coded or measured in brackets. In contrast, because the nature of health status is inherently nebulous, survey respondents cannot provide a numerical value that exactly summarizes their morbidity level. Hence, while in principle, income or education could be measured perfectly (although it rarely is), we would argue that an omnibus measure of morbidity can never be.

A common solution to this problem is to use Self-Reported Health Status (SRHS)

which is a categorical variable that respondents use to summarize their own health state. The different integer values of SRHS correspond to different health categories *e.g.* excellent, very good, *etc.* The approach that we adopt in this paper is that SRHS is a proxy for a more fundamental latent variable that corresponds to the respondent's actual health status. In this sense, our approach is similar to the approach adopted by Halliday and Mazumder (2017) who also employed latent variable models to estimate sibling correlations in health status which can be viewed as another measure of mobility.

In Halliday et al. (2018), we address many of these critiques and provide the first set of estimates of the intergenerational health association (IHA) in the United States. We estimated an IHA on the order of 0.2-0.25. This implies that a child can expect to "inherit" 20-25% of their parents' health capital. Note that this is substantially lower than analogous estimates of the intergenerational correlation in earnings which tend to be 0.5 or higher in the United States (*e.g.* Mazumder (2005), Mazumder (2016)). Based on these estimates, there is substantially more mobility in health status across generations than in economic status.

An important feature of this earlier work was that we took long time averages of SRHS for parents and children with the intent of extracting a more fundamental latent variable to compute the IHA. However, the models in that paper were linear. This paper builds on that work by estimating a non-linear model of intergenerational latent health transmission in which the IHA is an explicit parameter. The fundamental question that we aim to answer in this work is how the estimation of the IHA is affected by employing non-linear, latent variable models.

In this work, we focus exclusively on SRHS. As in our earlier work, the reason is that any study of intergenerational correlations requires a long enough panel so that both the parents and the children can be observed for a sufficiently long time period. We employ the Panel Study of Income Dynamics (PSID) which is the longest running panel in the world. In addition, the PSID has been collecting information on SRHS since 1984 which

provides us with thirty years of data on health status. We know of no other data source in the world in which this is possible.[1]

We estimate an IHA that varies between around 0.2 and 0.3 depending on the specification. Without any controls beyond age and gender, the IHA is 0.3 which is about one-fifth larger than the corresponding estimate from our earlier work of 0.23. We then include a battery of variables that exhibit little variation within families such as race and education to see how these "moderators" impact the estimation of the IHA. We show that race and education can explain roughly one-third of the IHA. In particular, education, and especially children's education, appears to be an important mechanism by which health is transmitted across generations.

On the whole, we show that employing a more rigorous non-linear model does affect estimates of the IHA, albeit modestly. Failure to model the latent variable properly results in an underestimation of the IHA on the order of 10-20% depending on the specification. The bias from estimating linear models is smallest when a large number of moderators is included in the model. Notably, employing non-linear models matters less for estimates of the rank-rank correlation and, in fact, we show that there is no difference in rank-based mobility estimates in linear and non-linear models.

While we are among the first set of authors to estimate intergenerational health mobility within a latent variable framework, there is a (mostly) nascent literature that is starting to investigate health mobility. Probably, the paper that is closest to this and Halliday et al. (2018) is Andersen (2018) who, in work subsequent to ours, estimates similar models using Danish registers and finds that there is even more health mobility in Denmark than in the United States. Earlier work finds that the intergenerational correlation in life spans is on the order of 0.15-0.30 (*e.g.* Ahlburg (1998) and Yashin and Iachine (1997)). Other work by Darden and Gilleskie (2016) finds a significant transmission of

---

[1]The PSID does collect other information on health status but most of this is only available in the 1999 wave and beyond which makes it difficult to use in study of intergenerational correlations.

smoking behaviors across generations in the Framingham Heart Study. Work by Johnston et al. (2013) investigates intergenerational associations in mental health in the British Cohort Study and estimates correlations between 0.13 and 0.15. Currie and Moretti (2007) document a strong intergenerational transmission of low birth weight. Akbulut-Yuksel and Kugler (2017) employ the National Longitudinal Survey of Youth and find strong intergenerational persistence in weight, height, and body mass index. Finally, Björkegren et al. (2019) use Swedish administrative data on adoptees to investigate the relative importance of biological and adoptive parents in determining the health of the child; they show that the biological parents matter considerably more.

As far as we can tell, we make two methodological innovations in this paper. The first is that we show how to estimate a non-linear model with categorical outcomes in which the index is a mixed linear model as is common in the genetics literature and becoming increasingly common in the economics literature on inequality and economic mobility (e.g. Mazumder (2008)). The technical problem with these models is that they are non-linear and also have a parameter dimension that increases with the sample size. This is the incidental parameters problem first discussed in Neyman and Scott (1948). As such, methods that rely on maximization will typically be biased. To address these issues, we develop a Bayesian estimation procedure that fuses the approach to the estimation of selection models in van Hasselt (2011) with another approach to estimating non-linear binary outcome models in which the index is also a mixed linear model developed in Halliday and Mazumder (2017). Our second innovation is that we show how to include covariates that do not vary within cross-sectional units (families in our case). This innovation builds on Hausman and Taylor (1981). These covariates are what we call "mediators" which are variables that mitigate the persistence of health across generations.

The balance of this paper is organized as follows. In Section 2, we describe our extract from the PSID. After that in Section 3, we discuss our model of health transmission across generations and how we estimate it. Next, in Section 4, we provide some descriptive

evidence on the determinants of SRHS in the PSID. In Section 5, we discuss our findings. We then conclude in Section 6.

2   The Data

We employ the PSID spanning the years 1984-2013. The 30 year time span of the data allows us to collect a sufficiently large amount of information on the SRHS of both parents and children. Note that after 1999, the PSID was conducted biannually, so that we have a total of 22 waves in our sample. Because SRHS was only collected for heads of household and their spouses, our sample is restricted to them. Consequently, we only include adult children, *i.e.* those over age 20, in the sample. In total, we have 243,150 individual/time observations. Of these, we have 52,459 observations for fathers or 21.57% of the sample, 86,818 observations for mothers 35.71% of the sample, and 103,873 observations for children or 42.72% of the sample. We report a break-down of samples sizes by year and family relation in Table A1. In total, we have 6,027 households, defined as the unit that once contained the parents and children, and 20,400 individuals. In Table A2, we report the frequencies for the length of time in the panel. The median length in the panel was eight years.

In Table 1, we report descriptive statistics for the variables that we use. Statistics are reported for the whole sample as well as for fathers, mothers, and children. Our main variable of interest is SRHS which is a categorical variable that takes on integer values between one and five (1=excellent, 2 = very good, 3 = good, 4 = fair, 5 = poor). The average for SRHS in the sample is 2.484 which falls somewhere between "very good" and "good." The average age in the sample is 43.133 with parents having an average age of about 49 and children having an average age of about 35. The difference in ages across generations in the sample is about 14 years which may seem low but this most likely reflects a combination of mortality and the fact that children can be sampled well

6

after parents have died. Finally, we also employ data on some individual and household characteristics including years of schooling, family income, and race dummies.

The family income variable is a measure of permanent household income. The variable that we employ is the average of all available time observations of total family income. For each yearly income observation, we divided family income by the square root of the household size. We then deflated yearly income to 2013 dollars using the CPI urban consumers series.

Our primary health measure is SRHS. Within countries, these subjective health measures are highly predictive of health outcomes including mortality (*e.g.* Ware et al. (1978) and Van Kippersluis et al. (2010)). As discussed by Sen (2002) and Banks and Smith (2012), there are issues when comparing subjective health measures such as SRHS across substantially different groups such as Indian states or countries.[2] However, because we are comparing SRHS within families, these issues are less salient in our context. That said, we will conduct some investigations into robustness to reporting heterogeneity by parents and children as in Lindeboom and Van Doorslaer (2004) later on in the paper.

3 A Simple Model of Health Transmissions

The Model

We let h_{pft} and h_{cft} denote the observed health of the parent and child for parent $p = 1, ..., P_f$, child $c = 1, ..., C_f$, family $f = 1, ..., F$, and time $t = 1, ..., T_{jf}$ for $j \in \{c, p\}$. Note that T depends on f and either c or p. Health outcomes are discrete and can take on outcomes $j \in \{1, ..., 5\}$. Outcomes are then determined by the following latent variable structure:

$$h_{pft} = j \iff \alpha_{j-1} < X_{pft}\beta^P + \phi_f^P + \nu_{pft} \leq \alpha_j$$
$$h_{cft} = j \iff \alpha_{j-1} < X_{cft}\beta^C + \phi_f^C + \nu_{cft} \leq \alpha_j \tag{1}$$

---

[2]Some studies that investigate international differences in health consider objective outcomes like age at death e.g. Le Grand (1987).

where X_{pft} and X_{cft} are $1 \times K$ vectors. We assume that

$$\left(\begin{array}{c} \phi_f^P \\ \phi_f^C \end{array} \middle| w_f^P, w_f^C \right) \sim N \left(\left(\begin{array}{c} w_f^P \gamma^P \\ w_f^C \gamma^C \end{array} \right), \left(\begin{array}{cc} 1 & \sigma_{pc} \\ \sigma_{pc} & \sigma_{pc}^2 + \kappa^2 \end{array} \right) \right) \tag{2}$$

and

$$\left(\begin{array}{c} \nu_{pft} \\ \nu_{cft} \end{array} \middle| X_{pf}, X_{cf}, \phi_f^P, \phi_f^C \right) \sim N(0, \sigma^2 \mathbf{I_2})$$

where X_{pf} and X_{cf} collect the time-variant regressors for a given individual and w_f^P and w_f^C are $1 \times L$ vectors of family-specific variables. We normalize $\alpha_0 = -\infty$ and $\alpha_5 = \infty$. The covariance matrix for the time-invariant effects is standard and follows Koop and Poirier (1997) and van Hasselt (2011). Note that we are normalizing the residual variance of ϕ_f^P to be unity while we will be estimating the variance of ν_{pft} and ν_{cft}. The IHA is σ_{pc}. Note that the index in this model is a mixed linear model which (as already discussed) is common in the genetics literature and the economics literature on inequality and economic mobility (e.g. Mazumder (2008)).[3]

Using the prior on the fixed effects in equation (2), we can write

$$\phi_f^P = w_f^P \gamma^P + u_f^P$$
$$\phi_f^C = w_f^C \gamma^C + \sigma_{pc} u_f^P + \eta_f^C$$

where $\eta_f^C | w_f^P, w_f^C \sim N(0, \kappa^2)$, $u_f^P | w_f^P, w_f^C \sim N(0,1)$ and $E[u_f^P \eta_f^C | w_f^P, w_f^C] = 0$. This transformation preserves the original covariance structure of ϕ_f^P and ϕ_f^C. Next, if we let

---

[3]Some of the frontier work in Bayesian econometrics does allow for non-Gaussian distributions for latent variables (*e.g.* van Hasselt (2011). However, the normality assumption can be rationalized as follows. With joint normality, mobility is entirely summarized by the covariance structure of the latent variables or their second moments. Hence, any paper that focuses on the IHA (or an integenerational correlation, for that matter) is consistent with an underlying Gaussian distribution. If the distribution were more general, then the focus should be on rank-rank correlations and not the IHA. In future work, we might explore non-Gaussian distributions, but for now, it is beyond the scope of this work.

$\widetilde{h}_{pft}$ and $\widetilde{h}_{cft}$ denote the latent indices in equation (1), we can rewrite this model as

$$
\begin{aligned}
\widetilde{h}_{pft} &= X_{pft}\beta^P + w_f^P\gamma^P + u_f^P + \nu_{pft} \\
\widetilde{h}_{cft} &= X_{cft}\beta^C + w_f^C\gamma^C + \sigma_{pc}u_f^P + \eta_f^C + \nu_{cft}.
\end{aligned}
\tag{3}
$$

For reasons that we will articulate, we will employ Bayesian methods to estimate this model. This formulation of the model facilitates the use of the Gibbs sampler. For example, it has the desirable property that σ_{pc} can be treated as a regression parameter and, so will have a Normal conditional posterior distribution. Similarly, κ^2 will have an Inverse Gamma conditional posterior distribution.

In the model, there are both individual-specific covariates, X_{pft} and X_{cft}, and family-specific (and time-invariant) covariates, w_f^P and w_f^C. The former includes gender and age. For the latter, we include education, race, and permanent income all of which do not vary over time and are (for the most part) family-specific.[4] Inclusion of these covariates allows us to see how the covariance structure in latent health is altered as they are sequentially included in the model. In other words, they shed light on the mechanisms underlying the intergenerational transmission of health status. Note that the family-specific covariates are collinear with the fixed effects in u_f^P and η_f^C. Accordingly, to estimate the γ parameters, we will employ a step in the estimation procedure that is based on Hausman and Taylor (1981) which follows from the prior given in equation (2). In contrast to the estimation of the β's which relies on within family variation for identification, estimation of the γ's relies on between family variation. We discuss this further in the appendix.

We can stack the individual/family/time specific equations into a system of Seemingly Unrelated Regressions (SUR). The notation is somewhat complicated, so we leave the

---

[4]Specifically, we took averages of these covariates over families.

details for the appendix. The resulting SUR model can be written as:

$$\begin{pmatrix} \widetilde{h}^P \\ \widetilde{h}^C \end{pmatrix} = \begin{pmatrix} X^P & \mathbf{0} \\ \mathbf{0} & X^C \end{pmatrix} \begin{pmatrix} \beta^P \\ \beta^C \end{pmatrix} + \begin{pmatrix} \mathbf{P}w^P & \mathbf{0} \\ \mathbf{0} & \mathbf{C}w^C \end{pmatrix} \begin{pmatrix} \gamma^P \\ \gamma^C \end{pmatrix} + \begin{pmatrix} \mathbf{P} \\ \sigma_{pc}\mathbf{C} \end{pmatrix} \mathbf{u_p} + \begin{pmatrix} \mathbf{0} \\ \mathbf{C} \end{pmatrix} \eta_{\mathbf{c}} + \nu$$

or more compactly as

$$\widetilde{\mathbf{h}} = \mathbf{X}\beta + \mathbf{W}\gamma + \mathbf{Z}\theta + \nu$$

where $\widetilde{\mathbf{h}}$ is $T(\equiv T^P + T^C) \times 1$, $\mathbf{X}$ is $T \times 2K$, $\mathbf{W}$ is $T \times 2L$,

$$\mathbf{Z} \equiv \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \sigma_{pc}\mathbf{C} & \mathbf{C} \end{pmatrix},$$

and $\theta \equiv (\mathbf{u_p}', \eta_{\mathbf{c}}')'$. The matrix, $\mathbf{X}$, contains the individual characteristics and the matrix, $\mathbf{W}$, contains the family characteristics for both the parents and the children. Note that the block diagonal structure allows for different effects of observables on parent and child. Importantly, this SUR structure allows us to investigate if parent- or child-specific mediators are more important in explaining the transmission of health status which is something that we will explore later. Finally, the matrix, $\mathbf{Z}$, is a design matrix that allocates the family fixed effects. The formal definitions of all vectors and matrices are in the appendix.

We impose a couple of restrictions on the model to facilitate identification. First, as already discussed, we normalize $V(u_f^P)$ to unity. While it is typically standard to normalize σ^2 to be unity in this class of models, this is still a valid restriction.[5] The second restriction concerns the cut points. In a standard ordered response model, the cut-points and the variance of the idiosyncratic shocks (σ^2 in this model) are not separately identified since we can always scale the cuts and the change in σ^2. Hence, in the standard model, either

---

[5]Indeed, per the discussion in Heckman (1981) on p. 129, σ^2 is identified subject to a normalization on $V(u_f^P)$ since the intra-class correlation is identified in the one factor model with more than one time period.

the cut-points need to be normalized or σ^2 needs to be normalized. We suspect that that is also the case in our more complicated model. Our solution to this is to normalize the cut points to external estimates using the same sample.[6] The procedure that we employ to obtain the cuts externally is to estimate ordered logit models as in Table 2 and to use the cut-point estimates corresponding to the appropriate set of covariates.

We estimate this model using Bayesian simulation methods which offer some benefits over classical Maximum Likelihood Estimation (MLE). First, our estimation procedure only requires simulation from standard distributions whereas MLE requires an onerous maximization procedure which includes the estimation of a complete set of family fixed effects. Second and related, because MLE involves the estimation of family fixed effects and because its asymptotics are with respect to families, MLE has an incidental parameters problem (Neyman and Scott (1948)). Our Bayesian procedure involves simulation from the exact finite sample posterior of the parameters. Because it does not rely on asymptotic theory for the distribution of the parameters, there is no incidental parameters problem with the Bayesian approach.

We make two methodological innovations. First, we show how to estimate non-linear models with categorical outcomes when the index is a mixed linear model while avoiding the incidental parameters problem. Second, because these models contain cross-sectional fixed effects (i.e. coefficients on dummy variables for the cross-sectional units), we show how to include covariates that do not vary in the cross-section a la Hausman and Taylor (1981).

Interpretation of the IHA and Related Mobility Measures

The IHA obtains from a regression of the latent health of the children onto the latent

---

[6]We made this choice for two reasons. First, the algorithm for sampling the cut-points of an ordered choice model from Albert and Chib (1993) is very slow as the latent variables and cuts are highly correlated. Second, normalizing σ^2 to unity would require estimation of $V(u_f^P)$. While this is certainly feasible by sampling from an Inverse Wishart distribution (with a re-parameterization), normalizing $V(u_f^P) = 1$ allows us to interpret the IHA as the percentage of a one standard deviation increase in the parents' health that gets transmitted to their children. We argue that this is a clean interpretation.

health of the parents. Specifically, if we define $u_f^C \equiv \sigma_{pc} u_f^P + \eta_f^C$, then we can write

$$E\left[u_f^C | u_f^P\right] = \sigma_{pc} u_f^P.$$

The normality assumption in the model implies that the expectation is linear. Because we normalized $V(u_f^P)$ to unity, the IHA delivers the effect of a one standard deviation increase in the latent health of the parents on the expected latent health of their children. In other words, σ_{pc} of one standard deviation of the parents' health capital gets transmitted to the next generation.

We can also use this framework to move beyond the IHA to a discussion of rank-rank correlations as a summary of mobility as in Chetty et al. (2014) and Halliday et al. (2018). To fix ideas, we let q_α^P denote the αth percentile of u_f^P. So, if Φ denotes a Standard Normal CDF then we will have that $\Phi\left(q_\alpha^P\right) = \alpha$ and that $E\left[u_f^C | u_f^P = q_\alpha^P\right] = \sigma_{pc} q_\alpha^P$. Now, using the marginal distribution for the children's distribution, we can construct a measure of rank-rank mobility using

$$\tau(\alpha) = \Phi\left(\frac{\sigma_{pc}}{\sqrt{\sigma_{pc}^2 + \kappa^2}} \Phi^{-1}(\alpha)\right). \tag{4}$$

The interpretation of this object is that if the parent's health is at the αth percentile of their marginal distribution then their child's (expected) health will be at the $\tau(\alpha)^{th}$ percentile of the next generation's marginal distribution. For this reason, we call $\tau(\alpha)$ the *unconditional rank-rank regression*. We then obtain that if $\sigma_{pc} = 1$ and $\kappa^2 = 0$, then $\tau(\alpha) = \alpha$ implying perfect rank-persistence across generations. Conversely, if $\sigma_{pc} = 0$ and $\kappa^2 > 0$, then $\tau(\alpha) = 0.5$ implying zero rank persistence across generations. Finally, if $0 < \left|\frac{\sigma_{pc}}{\sqrt{\sigma_{pc}^2 + \kappa^2}}\right| < 1$, then we will obtain positive persistence across generations with some degree of mean reversion.

Bayesian Estimation

We now discuss Bayesian estimation of the model's parameters. We define $\lambda \equiv (\beta', \gamma', \sigma_{pc}, \kappa^2, \sigma^2)'$ and $\mathbf{h}$ as the collection of the blunt outcome variables (h_{pft} and h_{cft}) into a $T \times 1$ vector. Note that T is defined precisely in the appendix and corresponds to the total number of individual/time observations across both generations. The model's posterior is then given by

$$p(\widetilde{\mathbf{h}}, \lambda, \theta | \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z}).$$

To estimate λ and θ, we will employ a Gibbs sampler algorithm in which we first sample $\widetilde{h}$ using the data augmentation procedure discussed by Albert and Chib (1993), followed by σ_{pc}, β, θ, γ, and finally by κ^2 and σ^2. Conditional on $\widetilde{\mathbf{h}}$ and employing the structure that we have laid out, the model is a relatively standard linear hierarchical model:

$$
\begin{aligned}
p(\lambda, \theta | \widetilde{\mathbf{h}}, \mathbf{X}, \mathbf{W}, \mathbf{Z}) \quad &\propto \quad p(\widetilde{\mathbf{h}} | \lambda, \theta, \mathbf{X}, \mathbf{W}, \mathbf{Z}) p(\theta | \lambda) p(\lambda) \qquad\qquad (5) \\
&\propto \quad p(\widetilde{\mathbf{h}} | \lambda, \theta, \mathbf{X}, \mathbf{W}, \mathbf{Z}) \prod_{f=1}^{F} p(\eta_{cf} | \kappa) p(\kappa^2) p(\sigma^2) p(\beta) p(\gamma) p(\sigma_{pc})
\end{aligned}
$$

We employ conjugate prior distributions which are defined in the appendix. The Gibbs sampler algorithm that we employ is then defined below.

1. Initialize $\left(\sigma_{pc}^0, \beta^0, \theta^0, \gamma^0, \kappa^{2,0}, \sigma^{2,0} \right)$.

2. Sample from $p(\widetilde{\mathbf{h}} | \sigma_{pc}^{n-1}, \beta^{n-1}, \theta^{n-1}, \gamma^{n-1}, \kappa^{2,n-1}, \sigma^{2,n-1}, \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z})$. Specifically, we draw a total of T values of $\widetilde{h}_{ift}$ for $i \in \{c, p\}$ from

$$N \left(\mathbf{X}\beta^{n-1} + \mathbf{W}\gamma^{n-1} + \mathbf{Z}\theta^{n-1}, \sigma^{2,n-1} \right)$$

truncated on the interval (α_{j-1}, α_j) whenever $h_{ift} = j$ for $i \in \{c, p\}$.

3. Sample from

$$p(\sigma_{pc}|\beta^{n-1}, \theta^{n-1}, \gamma^{n-1}, \kappa^{2,n-1}, \sigma^{2,n-1}, \widetilde{\mathbf{h}}^n, \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z}),$$

$$p(\beta|\theta^{n-1}, \gamma^{n-1}, \kappa^{2,n-1}, \sigma^{2,n-1}, \widetilde{\mathbf{h}}^n, \sigma_{pc}^n, \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z}),$$

$$p(\theta|\gamma^{n-1}, \kappa^{2,n-1}, \sigma^{2,n-1}, \widetilde{\mathbf{h}}^n, \sigma_{pc}^n, \beta^n, \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z}),$$

$$p(\gamma|\kappa^{2,n-1}, \sigma^{2,n-1}, \widetilde{\mathbf{h}}^n, \sigma_{pc}^n, \beta^n, \theta^n, \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z}),$$

$$p(\kappa^2|\sigma^{2,n-1}, \widetilde{\mathbf{h}}^n, \sigma_{pc}^n, \beta^n, \theta^n, \gamma^n, \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z}),$$

$$p(\sigma^2|\widetilde{\mathbf{h}}^n, \sigma_{pc}^n, \beta^n, \theta^n, \gamma^n, \kappa^{2,n}, \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z})$$

where the conditional distributions are defined in the Appendix in equations (A.3) - (A.9).

4. Return to Step 2.

For the second step of the algorithm, we employ an efficient routine developed by Botev (2017) to sample from a truncated normal distribution.

At each iteration of the Gibbs sampler, we can compute a pseudo-R^2 as follows. We can simulate the latent variable by drawing a T vector of Normal random variables with mean zero and variance σ^2. These are then added to $\mathbf{X}\beta + \mathbf{W}\gamma + \mathbf{Z}\theta$ from that iteration. We can then predict a vector of the the categorical outcomes, $\widehat{\mathbf{h}}$. A pseudo-R^2 can then be computed as $\rho(\widehat{\mathbf{h}}, \mathbf{h})^2$ where ρ denotes Spearman's rank correlation to account for the ordinal nature of the data. Note that because of the inclusion of the family fixed effects in the design matrix, inclusion of many time-invariant variables will essentially transfer the explanatory power from $\mathbf{Z}$ to $\mathbf{W}$ and, thus, inclusion of additional time-invariant regressors will do little to affect the pseudo-R^2.

4 Descriptive Evidence

Before we discuss our estimates of the structural model of health transmission, we provide some descriptive evidence on SRHS in the PSID. First, we estimate a series of ordered logit models in which we regress SRHS onto our main exogenous covariates: age, sex, race, education, and income. Second, we provide some evidence on the intergenerational transmission of health status that relies on linear models that is similar to the estimates in Halliday et al. (2018). This set of estimates will provide us with an indication of how much of a difference employing the non-linear model makes.

In Table 2, we report the ordered logit estimates which display some common relationships between SRHS and typical exogenous covariates. First, SRHS declines with age. This is important because it suggests that people do not merely report their health status relative to others of the same age which has also been pointed out by Deaton and Paxson (1998). Second, we see the common result that women have worse SRHS than men. Third, we see common patterns across races with whites having better SRHS than blacks in the second column. Inclusion of race increases the R^2 by 47% relative to the previous column. Next, we add education in the third column. Again, we see the common pattern of increasing SRHS in years of schooling. Inclusion of the education variable further increases the R^2 by 48%. Finally, we add our measure of permanent family income in the fourth column. We see the common result that richer households have better SRHS. Note that this does attenuate the education variable somewhat. However, both variables remain highly significant. Inclusion of family income further increases the R^2 by 6%. Hence, conditional on race and education, income does little to improve the predictive power of the model.

Next, we provide a set of "reduced form" estimates of the IHA employing linear models. These estimates are based on a linear analogue to the model from the previous section. We will adopt the notation that $\overline{x_f}^j$ denotes the average over time periods and individuals

(within a given generation) of x_{jft} for either the parent ($j = P$) or child ($j = C$). We
then consider the following linear system:

$$
\begin{aligned}
\overline{h}_f^P &= \overline{X_f}^P \widetilde{\beta^P} + \overline{W_f}^P \widetilde{\gamma^P} + \widetilde{u_f^P} \\
\overline{h}_f^C &= \overline{X_f}^C \widetilde{\beta^C} + \overline{W_f}^C \widetilde{\gamma^C} + \widetilde{u_f^C} \\
\widetilde{u_f^C} &= \widetilde{\sigma_{pc}} \widetilde{u_f^P} + \widetilde{\eta}_f
\end{aligned}
$$

To economize on notation, we are suppressing the individual/time averages of the ν_{cft}
and ν_{pft} which will be small in a long panel. The dependent variables in the first two
equations of this system are averages of the cardinal five-point SRHS variables (denoted
h_{jft} earlier). This system is what essentially would obtain if the econometrician could
observe the system of latent variables in equation (3) directly. A measure of the bias that
is created by ignoring the latent variable structure from the previous section is then given
by:

$$
bias = \frac{\sigma_{pc} - \widetilde{\sigma_{pc}}}{\sigma_{pc}}. \tag{6}
$$

We will report this estimate of the bias in the coming section. Note that this system
delivers the same IHA estimates as in Halliday et al. (2018).

We report estimates from five specifications in Table 3. In the first column, we only
adjust for age and gender and we obtain a reduced form IHA of 0.247. Further adjusting
for race in the next column, we obtain a slightly smaller estimate of 0.228. Next, we
include education in the model in the third column and we obtain an IHA of 0.160.
So, adjusting for education results in an estimate that is about 30% lower than the
specification in the second column. In the fourth column, we add income to the model
and the estimate is about the same as it was in the third column when we only adjusted for
education. Accordingly, conditional on education, income does not appear to explain the
intergenerational transmission of health. This is consistent with Table 2 in the sense that

the inclusion of income in the ordered logit models had little impact on the R^2. Finally, in the fifth and final specification, we include family income but we exclude education. We obtain an IHA of 0.197. This is substantially higher than the IHA in the third column of 0.163 where we controlled for education but not earnings. What this suggests is that education is relatively more important in explaining the intergenerational transmission of health than income is.

5   Structural IHA Estimates

Core Estimates

We now report the core estimates of the IHA from the latent variable model. For each estimation, we drew 1500 simulations from the posterior distribution of the model's parameters. We discarded the first 500 of these due to the "burning-in" of the sampler. In the appendix, we discuss the convergence of the sampler.

We estimate five specifications. All specifications include age and gender as family-variant covariates. In the first specification, we include no additional controls. In the next three specifications, we sequentially include race, years of schooling, and permanent income as family-specific covariates. Note that each of these variables are averages that are specific to the family in a given generation. For example, education is the average years of schooling for all children or all parents in a given family. We think of these variables as indices specific to the family for education, race, or income.

The means of the simulated parameters and their corresponding confidence intervals are reported in Table 4. The first panel of the table reports the estimates of the parent-specific covariates. The second panel reports estimates of the children-specific covariates. The bottom panel reports the estimates of covariance parameters.

The estimates of the covariate coefficients reveal standard patterns and are broadly in-line with those in Table 2. The coefficient estimates on age and gender are very similar to

17

those in Table 2 and are very stable across generations and specifications. The coefficient estimates on education and income are also fairly stable and are numerically close to those from Table 2. For all estimations, the pseudo-R^2 is about 25%.

The coefficient estimates on the white and black dummies in columns two through four warrant some discussion. First, in both generations, we see that education and income attenuate, but do not eliminate, the black/white gap in health status. Second, unlike education and income, we do see generational differences in the impact of race on health status with there being less of a race gap in the younger generation.

We now move on to the covariance parameters and, specifically, the IHA. We visually depict the IHA estimates using box plots in Panel A of Figure 1 which has five box plots corresponding to the five IHA estimates from Table 4. Without any family-specific controls the IHA is 0.291. In the second column when race is added as a control, the IHA decreases to 0.257. What this says is that factors associated with race reduce the IHA by about 12%. Further adding education in column three reduces the IHA substantially; the IHA becomes 0.170 once race and education are included. This is 34% lower than the IHA from the second column and 42% lower than the baseline IHA estimate from first column. Adding permanent income to the other covariates in the fourth columns does not change the IHA from the previous column. The IHA is 0.184 and the confidence interval overlaps substantially with the confidence interval in the previous column.[7]

Based on the estimates of the IHA from the first four columns of Table 4, it is not clear whether education or earnings are more important in explaining the intergenerational transmission of health capital. To shed light on this, in the fifth column, we include

---

[7]It is important to note that inclusion of additional family-specific covariates increases the variance of components of the residual that are not correlated across generations. The variances, σ^2 and κ^2, measure the within and between components of the residual variation in latent health, respectively. Inclusion of family-specific variables increases the relative contribution of the within component of the variance. Because of the normalization on the variance of u_f^P, the parameter, σ^2, must increase. On a similar note, the variance, κ^2, is the variance component of the family-specific fixed effect in the children's generation that is *not* transmitted across generations. Inclusion of observable covariates that are transmitted across generations, therefore, increases the relative contribution of the the components that are not transmitted across generations.

earnings (and race) but exclude education from the model. We obtain an IHA of 0.223 which is 13% lower than the IHA in the second column that only includes race. By way of contrast, the IHA from the third column that includes education and race was 33% lower. A similar finding is obtained in Table 3.

There are three main take-aways from these estimates of the IHA. The first is that, without conditioning on any family-specific covariates, a child can expect to receive about 30% of their parent's health status. Of this, education and, to a lesser extent race explain over one-third of the transmission. Third, ignoring the latent variable matters somewhat but not tremendously. We report the bias created by ignoring the latent variable in the third row from the bottom of Table 4 given in equation 6. Estimates of the IHA that are obtained when the latent variable is not properly extracted are between 7 and 15 percent smaller than those obtained by estimating our latent variable model.

Cut Point Heterogeneity

The results from the core model, while allowing the β's and γ's to be different for parents and children, assume that the cut points are the same. This issue is discussed by Lindeboom and Van Doorslaer (2004) who propose using a simple likelihood ratio statistic to test for heterogeneity in the cut points across groups. In this spirit, we calculate test statistics for heterogeneity in (a) the regression coefficients and the cut points and (b) the cut points based on the models estimated in Table 2. For all four specifications, the likelihood ratio tests indicate that there is heterogeneity in both the regression coefficients and the cut points; all eight p-values are essentially zero.

Allowing for heterogeneity in the cut points is useful for two reasons. First, this allows us to see if the results in Table 4 are robust to this generalization. Second, allowing for heterogeneity in the cuts across generations enables us to include different sets of covariates for parents and children which we will do later on in the paper.[8]

---

[8]This point is not obvious. The reason has to do with the cuts needing to change with the selection of covariates. Every time the set of covariates changes, the estimation of the cut points in a standard ordered logit model changes. Because we needed to utilize cuts from an external estimation due to the

In Table 5, we report estimation results from the generalized model with the different cuts by generation. We consider the same five specifications as in Table 4. Also, as with the estimates in Table 4, we estimated the cut-points externally except with these estimates we did so separately by parent and child. Finally, we also report box plots of the IHA from this generalized model in Panel B of Figure 1.

The main take-away from these estimates is that leaving the cut-points unrestricted by generation results in slightly larger estimates of the IHA for all five specifications. In the most parsimonious specification in the first column, we obtain an estimate of the IHA of 0.308, whereas the corresponding estimate in Table 4 was 0.291. However, note that the 95% confidence intervals from the first column of each table do overlap. This is the case for the other four specifications as well. The point estimates of the IHA are systematically higher in Table 5, but the confidence intervals of the estimates always overlap. This is clearly shown in both panels of Figure 1.

Unconditional Rank-Rank Regressions

We now compute the unconditional rank-rank regressions as defined by $\tau(\alpha)$ in equation (4). For both the models with homogeneous cut points in Table 4 and heterogeneous cut points in Table 5, we compute $\tau(\alpha)$ for the specification in the first column with no controls (aside from age and gender) and in the fourth column that controls for race, education, and income. We report our estimates of $\tau(\alpha)$ from these two specifications in Figure 2. Each panel of the figure reports lines for $\tau(\alpha) = 0.5$ which corresponds to perfect health mobility and $\tau(\alpha) = \alpha$ which corresponds to perfect health persistence. Finally, in the bottom of each panel in the figure, we report the rank-rank or Spearman correlation which can be obtained from an OLS regression of $\tau(\alpha)$ onto α.[9]

In all four panels of Figure 2, the plots of $\tau(\alpha)$ are closer to $\tau(\alpha) = 0.5$ than to $\tau(\alpha) = \alpha$ indicating a large degree of health mobility. This is especially true for the

---

identification issues articulated earlier, using different covariates across generations requires different cuts across generations as well.

[9]We restricted the regressions to values of α between 0.1 and 0.9.

panels in the second column from the models that include controls for race, education, and income. In addition, looking across the rows, we see that accounting for heterogeneity in the cut points makes little difference in mobility as estimated by $\tau(\alpha)$ whereas it did have some impact on estimates of the IHA.

The implied rank-rank correlations reported at the bottoms of each of the four panels are about 0.28 in the panels in the first column and 0.17 in the panels in the second column of Figure 2. Note that in each of the two columns, these correlations are very similar indicating that accounting for heterogeneity in the cuts makes little difference in these mobility measures. Finally, in Table A3 in the appendix, we report estimates of "reduced form" rank-rank correlations analogous to what we did in Table 3 for the IHA. The corresponding correlations are 0.281 and 0.163 and can be found in the first and fourth columns of the table. This is almost identical to the rank-rank correlations from the structural model. In Figure A3 of the appendix, we report estimates of $\tau(\alpha)$ for the remaining three specifications in Tables 4 and Table 5 and we still see that the implied rank-rank correlations from the latent variable model are remarkably similar to those from Table A3. This suggests that computing the rank-rank correlations from the linear model as was done in Halliday et al. (2018) does not result in any meaningful bias.

Why are the rank-based mobility estimates from the non-linear and "reduced form" models approximately equal? The reason is that for any two random variables, X and Y, the rank correlation is the same for any monotonic transformations of the two variables. Because the variables $\overline{h}_f^P$ and $\overline{h}_f^C$ are close to monotonic transformation of u_f^P and u_f^C, we should expect to find similar rank correlations in both the non-linear and linear models.

Generation-Specific Mediators

Previous results depicted in Figure 1 suggested that education was the primary mediator in the link between parent and child health. However, in those results, we included the education of both the parent and the child as mediators. As such, it is not clear from previous findings whether it is the education of the child or the parent that is the most

important. To investigate this, we employ heterogeneous cut points as we did above while adjusting only for child or parent education but not both.

We present these estimations in Figure 3. The figure displays box plots of the IHA from four estimations. The first includes only race as a control and is the same as the estimate from the second column of Table 4. The second adds parent education but not child education. The third includes child education but not parent education. The fourth includes both and is therefore identical to the estimate in the third column of Table 4.

The results indicate that, while both parent and child education matter, the latter matters more. Adjusting for only race, the IHA is 0.257. Further inclusion of parent education drops the mean to 0.228. However, adding child education but excluding parent education as in the third box plot drops the mean further to 0.190. Finally, including both parent and child education results in an IHA of 0.184.

6 Conclusion

In this paper, we provided the first structural estimates of the Intergenerational Health Association based on a non-linear, latent variable model of health determination. Using the PSID, we compute an IHA of 0.291 when only adjusting for age and gender. This is roughly 20% larger than the only other estimate of the IHA in the literature from Halliday et al. (2018). Our ability to extract a more accurate measurement of health status is what underlies the larger estimates. In addition, the non-linear model delivers an unadjusted rank-rank correlation of 0.28 and a correlation of 0.16 when adjusting for race, education, and income. Both the IHA and rank-rank correlation indicate a large degree of health mobility across generations. Notably, rank-based estimates from our non-linear model are identical to those from analogous linear models.

Based on this, estimation of linear models as was done in our previous work should be sufficient for some aspects of the study of intergenerational health mobility. However,

researchers should bear two issues in mind. First, as shown in Halliday et al. (2018), taking long time averages of self-rated health is critical. Second, we show in this paper that estimations of the IHA with linear models have a small bias, but rank-based mobility estimations have no bias.

We also include a battery of family-specific covariates such as proxies for race, education, and permanent household income. We show that the combination of race and education can explain about 40% of the IHA. Of these, education matters the most as its inclusion reduces the IHA by roughly one third relative to a specification that only includes race. In addition, we find that children's education matters more than parent's education. This suggests that parents regardless of their educational attainment (or policy makers) may be able to weaken intergenerational links in health capital by investing in children's education.

One interesting outcome of this work is that there is substantially less persistence in health status across generations than in economic status which typically exhibits an intergenerational correlation of 0.5 or higher (see Mazumder (2005), for example). This is the case even after we account for proper measurement of health status using latent variable models. Future work should investigate why this is the case and what it implies for the association between health and economic status in later parts of the life-cycle.

Finally, we have developed a framework that can be employed when looking at intergenerational transmissions in other contexts where latent variables are important. Our methodology allows researchers to estimate non-linear models with categorical outcomes in which (a) the index is a mixed linear model and (b) some covariates do not exhibit variation within cross-sectional units. The methods that we have developed in this paper could be applied to study intergenerational persistence in non-cognitive skills (or other personality traits) where we may have multiple measurements that are error-ridden proxies of an underlying latent variable. We hope that other researchers adopt the approach that we have developed in this work to other contexts so that they can arrive at a better

understanding of persistence in other non-economic outcomes across generations.

References

Ahlburg, Dennis, "Intergenerational transmission of health," *The American Economic Review*, 1998, *88* (2), 265–270.

Akbulut-Yuksel, Mevlude and Adriana Kugler, "Heterogeneity in the long term health effects of warfare," *Economics & Human Biology*, 2017, *27*, 126–136.

Albert, James H and Siddhartha Chib, "Bayesian analysis of binary and polychotomous response data," *Journal of the American statistical Association*, 1993, *88* (422), 669–679.

Andersen, Carten, "Intergenerational Health Mobility: Evidence from Danish Registers," 2018. Working paper.

Banks, James and James P Smith, "International comparisons in health economics: evidence from aging studies," *Annual Review of Economics*, 2012, *4*, 57–81.

Björkegren, Evelina, Mikael Lindahl, Marten Palme, and Emilia Simeonova, "Pre-and Post-Birth Components of Intergenerational Persistence in Health and Longevity: Lessons from a Large Sample of Adoptees," 2019.

Botev, ZI, "The normal law under linear restrictions: simulation and estimation via minimax tilting," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2017, *79* (1), 125–148.

Case, Anne, Angela Fertig, and Christina Paxson, "The lasting impact of childhood health and circumstance," *Journal of health economics*, 2005, *24* (2), 365–389.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *," *The Quarterly Journal of Economics*, 2014, *129* (4), 1553–1623.

Currie, Janet and Enrico Moretti, "Biology as destiny? Short-and long-run determinants of intergenerational transmission of birth weight," *Journal of Labor economics*, 2007, *25* (2), 231–264.

Darden, Michael and Donna Gilleskie, "The Effects of Parental Health Shocks on Adult Offspring Smoking Behavior and Self-Assessed Health," *Health economics*, 2016, *25* (8), 939–954.

Deaton, Angus S and Christina Paxson, "Health, income, and inequality over the life cycle," in "Frontiers in the Economics of Aging," University of Chicago Press, 1998, pp. 431–462.

Grand, Julian Le, "Inequalities in health: some international comparisons," *European economic review*, 1987, *31* (1-2), 182–191.

Halliday, Timothy J and Bhashkar Mazumder, "An Analysis of Sibling Correlations in Health Using Latent Variable Models," *Health Economics*, 2017.

_ , _ , **and Ashley Wong**, "Intergenerational Health Mobility in the US," 2018. IZA Working Paper.

Hausman, Jerry A and William E Taylor, "Panel data and unobservable individual effects," *Econometrica: Journal of the Econometric Society*, 1981, pp. 1377–1398.

Heckman, James J, "Statistical Models for Discrete Panel Data," in C. Manksi and D. McFadden, eds., *Structural Analysis of Discrete Data and Econometric Applications*, Cambridge: MIT Press, 1981, chapter 3, pp. 114–178.

Johnston, David W, Stefanie Schurer, and Michael A Shields, "Exploring the intergenerational persistence of mental health: Evidence from three generations," *Journal of health economics*, 2013, *32* (6), 1077–1089.

Jones, Charles I and Peter J Klenow, "Beyond GDP? Welfare across countries and time," *The American Economic Review*, 2016, *106* (9), 2426–2457.

Kippersluis, Hans Van, Owen O'Donnell, Eddy Van Doorslaer, and Tom Van Ourti, "Socioeconomic differences in health over the life cycle in an Egalitarian country," *Social science & medicine*, 2010, *70* (3), 428–438.

Koop, Gary and Dale J Poirier, "Learning about the across-regime correlation in switching regression models," *Journal of Econometrics*, 1997, *78* (2), 217–227.

Lindeboom, Maarten and Eddy Van Doorslaer, "Cut-point shift and index shift in self-reported health," *Journal of health economics*, 2004, *23* (6), 1083–1099.

Mazumder, Bhashkar, "Fortunate sons: New estimates of intergenerational mobility in the United States using social security earnings data," *The Review of Economics and Statistics*, 2005, *87* (2), 235–255.

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

_ , "Estimating the Intergenerational Elasticity and Rank Association in the United States: Overcoming the Current Limitations of Tax Data," in "Inequality: Causes and Consequences," Emerald Group Publishing Limited, 2016, pp. 83–129.

Neyman, Jerzy and Elizabeth L Scott, "Consistent estimates based on partially consistent observations," *Econometrica: Journal of the Econometric Society*, 1948, pp. 1–32.

Sen, Amartya, "Health: perception versus observation: self reported morbidity has severe limitations and can be extremely misleading," 2002.

van Hasselt, Martijn, "Bayesian inference in a sample selection model," *Journal of Econometrics*, 2011, *165* (2), 221–232.

Ware, JE, Allyson Davies-Avery, and Cathy A Donald, "Conceptualization and measurement of health for adults in the health insurance study: Vol. V, general health perceptions," *Santa Monica, CA: Rand Corporation*, 1978, *479.*

Yashin, Anatoli I and Ivan A Iachine, "How frailty models can be used for evaluating longevity limits: Taking advantage of an interdisciplinary approach," *Demography*, 1997, *34* (1), 31–48.

Yu, Bin and Per Mykland, "Looking at Markov samplers through cusum path plots: a simple diagnostic idea," *Statistics and Computing*, 1998, *8* (3), 275–286.

Table 1: Descriptive Statistics

	Everybody	Fathers	Mothers	Children
SRHS	2.484	2.489	2.726	2.279
	(1.089)	(1.129)	(1.120)	(0.998)
Age	43.133	49.325	48.627	35.415
	(14.439)	(14.224)	(14.750)	(9.994)
Sex (1=female)	0.591	0	1	0.546
	(0.492)	-	-	(0.498)
Years of Schooling	12.899	12.706	12.351	13.454
	(2.629)	(3.053)	(2.693)	(2.199)
White	0.609	0.711	0.560	0.600
	(0.488)	(0.453)	(0.492)	(0.490)
Black	0.363	0.256	0.412	0.376
	(0.481)	(0.436)	(0.492)	(0.484)
Family Income (2013 dollars)	42,996.14	50,798.24	38,506.15	42,808.62
	(52,653.94)	(57,307.60)	(47,438.79)	(53,876.90)

Notes: This table report reports the means and standard deviation (in parentheses) of each variable. There were 6027 households in the sample.

Table 2: Ordered Logit Estimates

	(1)	(2)	(3)	(4)
Age	0.039***	0.042***	0.036***	0.038***
	(0.001)	(0.001)	(0.001)	(0.001)
Sex (1=female)	0.414***	0.314***	0.310***	0.277***
	(0.020)	(0.020)	(0.020)	(0.020)
White	-	-0.536***	-0.298***	-0.281***
		(0.075)	(0.068)	(0.066)
Black	-	0.316***	0.349***	0.256***
		(0.076)	(0.070)	(0.067)
Years of Schooling	-	-	-0.203***	-0.176***
			(0.005)	(0.005)
Family Income (2013 dollars)	-	-	-	-0.193
				(0.009)
R^2	0.0350	0.0516	0.0763	0.0807
F	6027	6027	6027	6027

* significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level
Notes: This table reports the coefficient estimates from a series of ordered logit models in which SRHS is regressed on exogenous covariates. Standard errors are reported in parentheses and are clustered by households.

Table 3: Reduced form Estimates of the IHA

	(1)	(2)	(3)	(4)	(5)
IHA ($\widetilde{\sigma_{pc}}$)	0.247***	0.228***	0.160***	0.166***	0.197***
	(0.011)	(0.012)	(0.013)	(0.013)	(0.012)
Age	X	X	X	X	X
Sex	X	X	X	X	X
White	-	X	X	X	X
Black	-	X	X	X	X
Years of Schooling	-	-	X	X	-
Family Income	-	-	-	X	X
F	6027	6027	6027	6027	6027

* significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level
Notes: We report estimates of the IHA from the "reduced form" model in Section 4. Robust standard errors are reported in parentheses.

Table 4: Parameter Estimates: Core Results

	(1)	(2)	(3)	(4)	(5)
Parents					
Age	0.0468	0.0444	0.0437	0.0450	0.0476
	(0.0461,0.0472)	(0.0410,0.0472)	(0.0401,0.0477)	(0.0403,0.0500)	(0.0440,0.0511)
Sex (1=female)	0.234	0.181	0.201	0.173	0.165
	(0.219,0.248)	(0.131,0.222)	(0.149,0.262)	(0.112,0.241)	(0.114,0.216)
White	-	-0.273	-0.405	-0.0711	-0.137
		(-0.485,-0.014)	(-0.660,-0.185)	(-0.216,0.072)	(-0.294,0.011)
Black	-	0.627	0.327	0.530	0.539
		(0.415,0.880)	(0.079,0.551)	(0.391,0.678)	(0.393,0.682)
Years of Schooling	-	-	-0.209	-0.159	-
			(-0.219,-0.199)	(-0.174,-0.144)	
Family Income (2013 dollars)	-	-	-	-0.248	-0.335
				(-0.293,-0.208)	(-0.360,-0.311)
Children					
Age	0.0424	0.0400	0.0427	0.0425	0.047
	(0.0416,0.0434)	(0.0377,0.0427)	(0.0397,0.0453)	(0.0397,0.0454)	(0.0438,0.0510)
Sex (1=female)	0.289	0.219	0.279	0.225	0.212
	(0.269,0.310)	(0.174,0.266)	(0.226,0.337)	(0.168,0.280)	(0.154,0.274)
White	-	-0.081	-0.432	-0.277	-0.246
		(-0.238,0.070)	(-0.605,-0.260)	(-0.474,-0.0767)	(-0.432,-0.051)
Black	-	0.341	-0.151	-0.122	-0.056
		(0.181,0.490)	(-0.323,0.018)	(-0.310,0.060)	(-0.242,0.126)
Years of Schooling	-	-	-0.189	-0.135	-
			(-0.204,-0.176)	(-0.153,-0.116)	
Family Income (2013 dollars)	-	-	-	-0.234	-0.318
				(-0.266,-0.202)	(-0.343,-0.295)
σ^2	1.566	1.667	1.849	1.866	1.824
	(1.555,1.578)	(1.654,1.681)	(1.833,1.866)	(1.850,1.881)	(1.805,1.845)
κ^2	1.325	1.419	1.553	1.535	1.500
	(1.260,1.394)	(1.352,1.489)	(1.473,1.635)	(1.455,1.618)	(1.421,1.587)
IHA (σ_{pc})	0.291	0.257	0.170	0.184	0.223
	(0.262,0.313)	(0.217,0.288)	(0.122,0.205)	(0.147,0.221)	(0.189,0.263)
bias	15 %	11 %	7 %	10 %	12 %
pseudo-R^2	0.251	0.250	0.252	0.251	0.248
	(0.248,0.253)	(0.247,0.254)	(0.249,0.256)	(0.248,0.255)	(0.244,0.251)
F	6027	6027	6027	6027	6027

Notes: Means of parameter simulations are reported based off of 1500 simulations from the posterior distribution (with the first 500 draws being discarded). 95 % confidence intervals are reported below. *bias* is calculated per the formula in equation 6 and corresponds to the bias in the IHA estimates in Table 3.
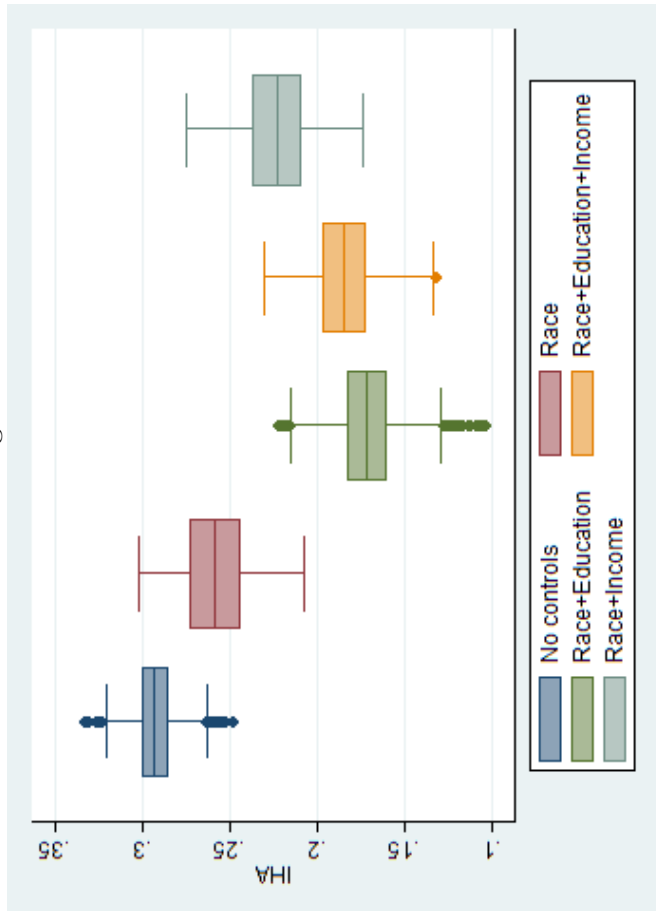
Table 5: Parameter Estimates: Results - Different Cuts by Generation

	(1)	(2)	(3)	(4)	(5)
			Parents		
Age	0.0460 (0.0454,0.0467)	0.0436 (0.0411,0.0461)	0.0431 (0.0402,0.0458)	0.0462 (0.0425,0.0502)	0.0470 (0.0436,0.0508)
Sex (1=female)	0.238 (0.223,0.253)	0.180 (0.140,0.217)	0.201 (0.157,0.245)	0.182 (0.136,0.233)	0.159 (0.114,0.212)
White	-	-0.190 (-0.381,0.004)	-0.386 (-0.576,-0.187)	-0.118 (-0.279, 0.032)	-0.198 (-0.353,-0.052)
Black	-	0.694 (0.507,0.888)	0.334 (0.148,0.524)	0.452 (0.300,0.605)	0.430 (0.284,0.573)
Years of Schooling	-	-	-0.204 (-0.213,-0.195)	-0.153 (-0.166,-0.139)	-
Family Income (2013 dollars)	-	-	-	-0.284 (-0.319,-0.247)	-0.374 (-0.400,-0.350)
			Children		
Age	0.0418 (0.0407,0.0428)	0.0417 (0.0392,0.0447)	0.044 (0.0410,0.0470)	0.0487 (0.0454,0.0524)	0.0483 (0.0450,0.0517)
Sex (1=female)	0.239 (0.220,0.258)	0.218 (0.172,0.270)	0.283 (0.225,0.338)	0.261 (0.207,0.320)	0.214 (0.159,0.271)
White	-	-0.407 (-0.584,-0.253)	-0.425 (-0.623,-0.243)	-0.273 (-0.469,-0.074)	-0.184 (-0.368,-0.013)
Black	-	0.036 (-0.143,0.190)	-0.132 (-0.312,0.053)	-0.102 (-0.283,0.095)	0.047 (-0.136,0.243)
Years of Schooling	-	-	-0.194 (-0.209,-0.179)	-0.139 (-0.159,-0.117)	-
Family Income (2013 dollars)	-	-	-	-0.227 (-0.262,-0.189)	-0.290 (-0.313,-0.264)
σ^2	1.591 (1.579,1.603)	1.698 (1.684,1.712)	1.867 (1.852,1.884)	1.930 (1.911,1.951)	1.857 (1.836,1.880)
κ^2	1.463 (1.394,1.539)	1.569 (1.493,1.649)	1668 (1.587,1.755)	1.700 (1.615,1.796)	1.652 (1.575,1.739)
IHA (σ_{pc})	0.308 (0.272,0.343)	0.275 (0.244,0.308)	0.187 (0.152,0.223)	0.195 (0.162,0.242)	0.241 (0.213,0.277)
bias	20 %	17 %	17 %	15 %	18 %
pseudo-R^2	0.249 (0.246,0.252)	0.249 (0.246,0.252)	0.252 (0.248,0.255)	0.251 (0.248,0.255)	0.246 (0.243,0.249)
F	6027	6027	6027	6027	6027

Notes: Per Table 4

Figure 1: IHA Box Plots
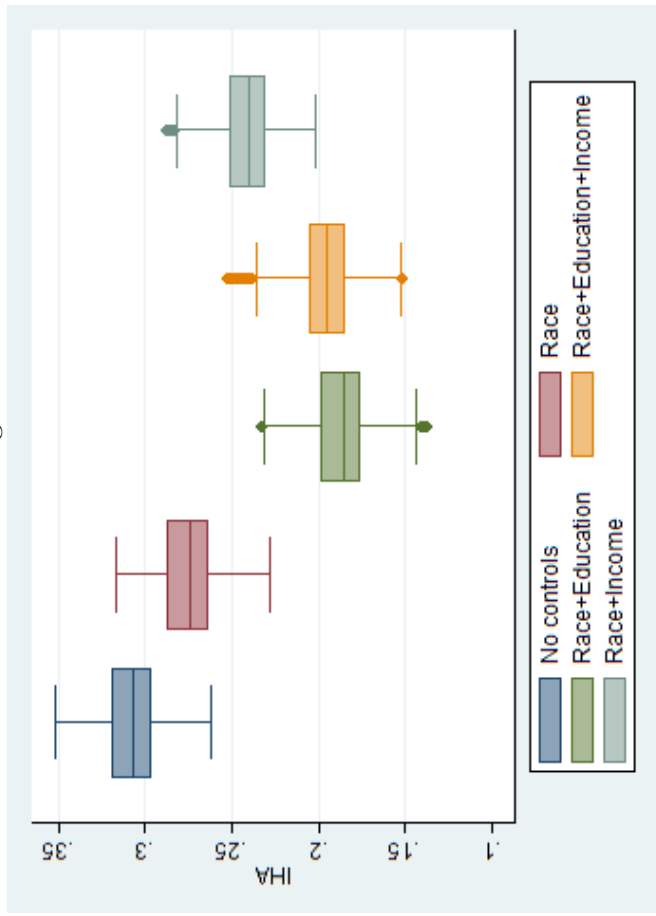
Figure 2: Unconditional Rank-Rank Regressions
Includes Race, Education, and Income
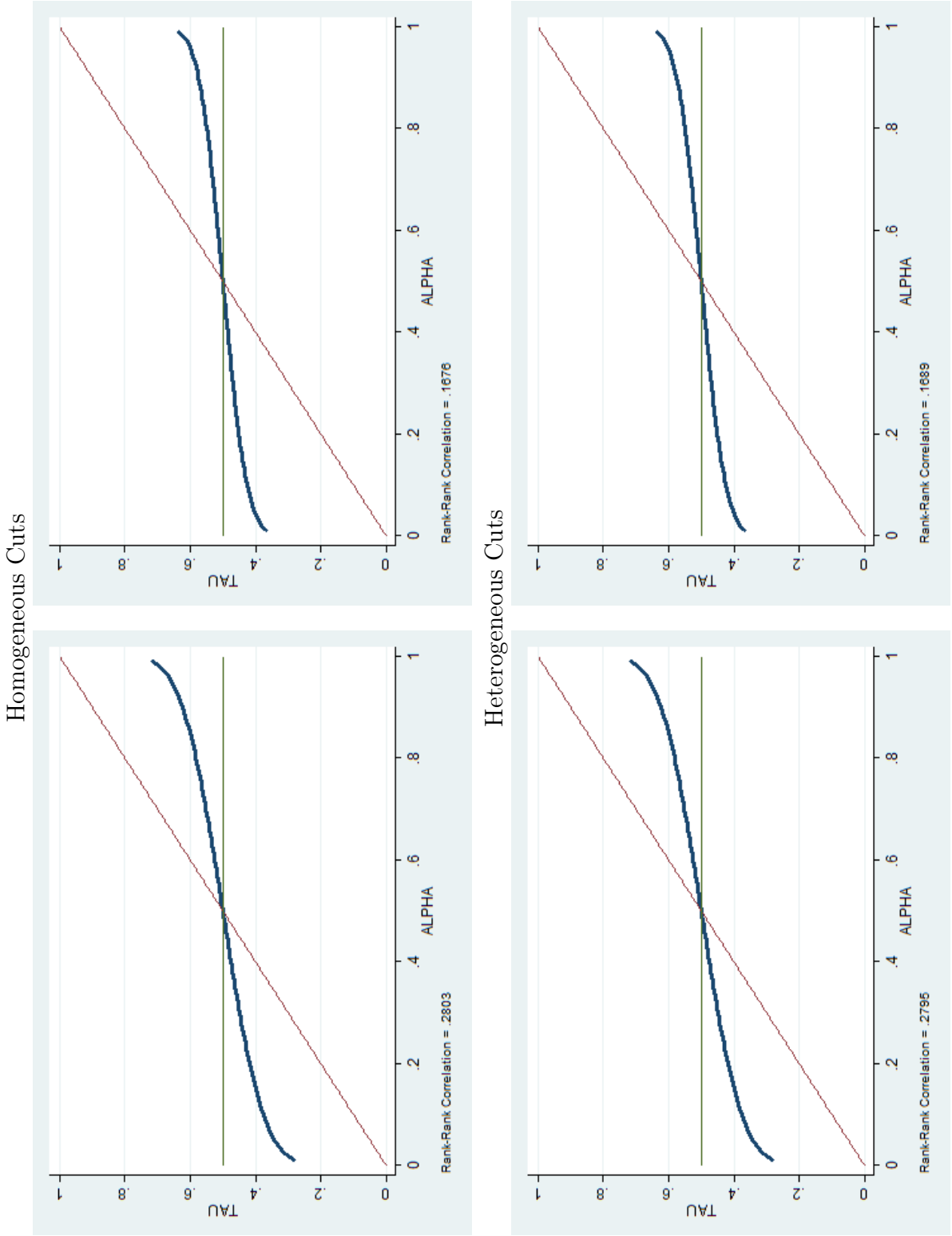
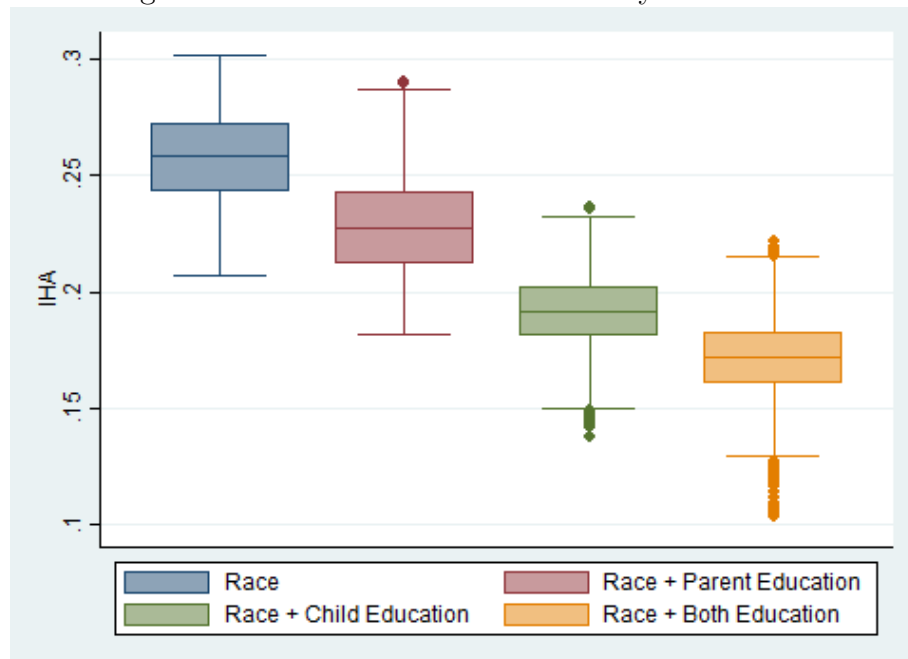Homogeneous Cuts

No Controls

Heterogeneous Cuts

Figure 3: IHA Box Plots: Education by Generation

A.1 Appendix: Deriving the SUR Model

We now derive the SUR model. If we stack the model in equation (3) over t and all parents or children in family f, we obtain

$$\widetilde{h}_f^P = X_f^P \beta^P + \mathbf{1_{T_f^P}} * w_f^P \gamma^P + \mathbf{1_{T_f^P}} * u_f^P + \nu_{pf} \qquad (A.1a)$$

$$\widetilde{h}_f^C = X_f^C \beta^C + \mathbf{1_{T_f^C}} * w_f^C \gamma^C + \mathbf{1_{T_f^C}} * \sigma_{pc} u_f^P + \mathbf{1_{T_f^C}} * \eta_f^C + \nu_{cf} \qquad (A.1b)$$

where equation (A.1a) is a $T_f^P (\equiv \sum_{p=1}^{P_f} T_{pf})$ - and equation (A.1b) is a $T_f^C (\equiv \sum_{c=1}^{C_f} T_{cf})$ - system of equations and $\mathbf{1_X}$ denotes an X - vector of ones.[10] Next, we stack one more time over families to obtain:

$$\widetilde{h}^P = X^P \beta^P + \mathbf{P} w^P \gamma^P + \mathbf{P} \mathbf{u_p} + \nu_\mathbf{p} \qquad (A.2a)$$

$$\widetilde{h}^C = X^C \beta^C + \mathbf{C} w^C \gamma^C + \sigma_{pc} \mathbf{C} \mathbf{u_p} + \mathbf{C} \eta_\mathbf{c} + \nu_\mathbf{c}. \qquad (A.2b)$$

Equation (A.2a) is a $T^P (\equiv \sum_{f=1}^{F} T_f^P)$ - system in which X^P is $T^P \times K$, $w^P \equiv (w_1^{P'}, \dots w_F^{P'})'$ is $F \times L$, $\mathbf{P}$ is $T^P \times F$ and is defined as

$$\mathbf{P} \equiv \begin{pmatrix} \mathbf{1_{T_1^P}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{1_{T_F^P}} \end{pmatrix},$$

$\mathbf{u_p} \equiv (u_1^P, \dots, u_F^P)'$ which is $F \times 1$, and $\nu_\mathbf{p}$ collects the idiosyncratic shocks and is $T^P \times 1$. Similarly, equation (A.2b) is a $T^C (\equiv \sum_{f=1}^{F} T_f^C)$ - system in which X^C is $T^C \times K$, w^C is $F \times L$, $\mathbf{C}$ is $T^C \times F$ and is defined analogously to $\mathbf{P}$, $\eta_\mathbf{c}$ is defined analogously to $\mathbf{u_p}$ and is $F \times 1$, and $\nu_\mathbf{c}$, which also collects the idiosyncratic shocks, is $T^C \times 1$. Finally, we can stack the parent and child systems into one grand system. Doing this, we obtain

$$\begin{pmatrix} \widetilde{h}^P \\ \widetilde{h}^C \end{pmatrix} = \begin{pmatrix} X^P & \mathbf{0} \\ \mathbf{0} & X^C \end{pmatrix} \begin{pmatrix} \beta^P \\ \beta^C \end{pmatrix} + \begin{pmatrix} \mathbf{P} w^P & \mathbf{0} \\ \mathbf{0} & \mathbf{C} w^C \end{pmatrix} \begin{pmatrix} \gamma^P \\ \gamma^C \end{pmatrix} + \begin{pmatrix} \mathbf{P} \\ \sigma_{pc} \mathbf{C} \end{pmatrix} \mathbf{u_p} + \begin{pmatrix} \mathbf{0} \\ \mathbf{C} \end{pmatrix} \eta_\mathbf{c} + \nu$$

or more compactly as

$$\widetilde{\mathbf{h}} = \mathbf{X} \beta + \mathbf{W} \gamma + \mathbf{Z} \theta + \nu$$

where $\widetilde{\mathbf{h}}$ is $T (\equiv T^P + T^C) \times 1$, $\mathbf{X}$ is $T \times 2K$, $\mathbf{W}$ is $T \times 2L$,

$$\mathbf{Z} \equiv \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \sigma_{pc} \mathbf{C} & \mathbf{C} \end{pmatrix},$$

and $\theta \equiv (\mathbf{u_p}', \eta_\mathbf{c}')'$.

---

[10]Note that T_f^P and T_f^C correspond to the total number of time periods and either parents or children in a given family.

A.2 Appendix: Prior Distributions

We employ the following conjugate prior distributions:

$$\sigma_{pc} \sim N\left(0, \underline{s}_{pc}^{-1}\right)$$
$$\beta \sim N\left(0, \underline{H}_{\beta}^{-1}\right)$$
$$\gamma^{P} \sim N\left(0, \underline{H}_{\gamma^{P}}^{-1}\right)$$
$$\gamma^{C} \sim N\left(0, \underline{H}_{\gamma^{C}}^{-1}\right)$$
$$\kappa^{2} \sim IG_{2}\left(s_{\kappa}, \nu_{\kappa}\right)$$
$$\sigma^{2} \sim IG_{2}\left(s_{\sigma}, \nu_{\sigma}\right).$$

Note that, as already discussed, $\eta_{cf} \sim N(0, \kappa^{2})$.

A.3 Appendix: Conditional Posterior Distributions

We now discuss the conditional distributions that we use in the Gibbs sampler. We begin by discussing the sampling of all of the parameters except for γ which warrants a separate discussion. These conditional posteriors are based off of the factorization of the posterior given in equation (5). This factorization of the posterior is written in terms of the fixed effects given in θ which are orthogonal to the family-specific covariates in w_{f}^{P} and w_{f}^{C}. To understand the conditional posterior for γ, we will write the posterior in terms of ϕ^{P} and ϕ^{C} which are inclusive of w_{f}^{P} and w_{f}^{C}.

A.3.1 Conditional Posteriors for $\sigma_{pc}, \beta, \theta, \kappa^{2}$ and σ^{2}

The conditional distribution of σ_{pc} is given by

$$\sigma_{pc}|\beta, \theta, \gamma, \kappa^{2}, \sigma^{2} \sim N\left(\sigma_{pc}^{*}, \sigma^{2} * H_{\sigma_{pc}}^{-1}\right) \tag{A.3}$$

where

$$
\begin{aligned}
\mathbf{D} &\equiv \mathbf{Cu_{p}} \\
H_{\sigma_{pc}} &\equiv \mathbf{D'D} + \underline{s}_{pc} \\
\mathbf{W}^{C} &\equiv \mathbf{C}w^{C} \\
\widehat{\sigma}_{pc} &\equiv (\mathbf{D'D})^{-1}\mathbf{D'}\left[\widetilde{\mathbf{h}}^{C} - \mathbf{X}^{C}\beta^{C} - \mathbf{W}^{C}\gamma^{C} - \mathbf{C}\eta_{\mathbf{c}}\right] \\
\sigma_{pc}^{*} &\equiv H_{\sigma_{pc}}^{-1}(\mathbf{D'D})\widehat{\sigma}_{pc}.
\end{aligned}
$$

Second, the conditional distribution of β is given by

$$\beta|\theta, \gamma, \kappa^{2}, \sigma^{2}, \sigma_{pc} \sim N\left(\beta_{*}, \sigma^{2} * H_{\beta}^{-1}\right) \tag{A.4}$$

where

$$
\begin{aligned}
H_\beta &\equiv \mathbf{X}'\mathbf{X} + \underline{H}_\beta \\
\widehat{\beta} &\equiv (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}' \left[\widetilde{\mathbf{h}} - \mathbf{W}\gamma - \mathbf{Z}\theta \right] \\
\beta_* &\equiv H_\beta^{-1} (\mathbf{X}'\mathbf{X}) \widehat{\beta}.
\end{aligned}
$$

Note that $\mathbf{Z}$ contains σ_{pc} which was sampled in the previous step. Next, the conditional distribution of θ is given by

$$
\theta | \gamma, \kappa^2, \sigma^2, \sigma_{pc}, \beta \sim N \left(\theta_*, \sigma^2 * H_\theta^{-1} \right) \tag{A.5}
$$

where

$$
\begin{aligned}
H_\theta &\equiv \mathbf{Z}'\mathbf{Z} + \underline{H}_\theta \\
\underline{H}_\theta &\equiv \begin{pmatrix} \mathbf{I}_F & \mathbf{0} \\ \mathbf{0} & \kappa^{-2}\mathbf{I}_F \end{pmatrix} \\
\widehat{\theta} &\equiv (\mathbf{Z}'\mathbf{Z})^{-1} \mathbf{Z}' \left[\widetilde{\mathbf{h}} - \mathbf{X}\beta - \mathbf{W}\gamma \right] \\
\theta_* &\equiv H_\theta^{-1} (\mathbf{Z}'\mathbf{Z}) \widehat{\theta}.
\end{aligned}
$$

Finally, the conditional distributions of κ^2 and σ^2 are given by

$$
\kappa^2 | \sigma^2, \sigma_{pc}, \beta, \theta, \gamma \sim IG_2 \left(\kappa_*^2, F - 2 + \nu_\kappa \right) \tag{A.6}
$$

where

$$
\kappa_*^2 \equiv \eta_\mathbf{c}' \eta_\mathbf{c} + s_\kappa.
$$

and

$$
\sigma^2 | \sigma_{pc}, \beta, \theta, \gamma, \kappa^2 \sim IG_2 \left(\sigma_*^2, T - 2 + \nu_\sigma \right) \tag{A.7}
$$

where

$$
\begin{aligned}
\sigma_*^2 &\equiv s_\sigma + s \\
s &\equiv \left[\widetilde{\mathbf{h}} - \mathbf{X}\beta - \mathbf{W}\gamma - \mathbf{Z}\theta \right]' \left[\widetilde{\mathbf{h}} - \mathbf{X}\beta - \mathbf{W}\gamma - \mathbf{Z}\theta \right].
\end{aligned}
$$

A.3.2   Conditional Posterior for γ

To derive the conditional posteriors for γ, we write the posterior in terms of ϕ instead of θ:

$$
p(\widetilde{\mathbf{h}}, \lambda, \phi | \mathbf{h}, \mathbf{X}, \mathbf{W}, \mathbf{Z}).
$$

Given this formulation of the posterior, the information on γ comes through the prior in equation (2). Conditional on w_f^P, w_f^C, the posterior on $(\phi_f^P, \phi_f^C, \lambda)$ can be written as:

$$
\begin{aligned}
p(\phi_f^P, \phi_f^C, \lambda | w_f^P, w_f^C) &= p(\phi_f^P, \phi_f^C | \lambda, w_f^P, w_f^C) p(\lambda) \\
&= p(\phi_f^C | \phi_f^P, \lambda, w_f^P, w_f^C) p(\phi_f^P | \lambda, w_f^P, w_f^C) p(\lambda) \\
&= p(\phi_f^C | \phi_f^P, \gamma^C, \sigma_{pc}, \kappa^2, w_f^P, w_f^C) p(\phi_f^P | \gamma^P, w_f^P, w_f^C) p(\lambda) \\
&\propto \kappa^{-1} \exp\left(-\kappa^{-2} \left(\phi_f^C - w_f^C \gamma^C - \sigma_{pc} u_f^P \right)^2 \right) \times \\
&\qquad \exp\left(- \left(\phi_f^P - w_f^P \gamma^P \right)^2 \right) p(\lambda)
\end{aligned}
$$

Based on this, we can employ standard calculations from the Bayesian literature and derive the conditional distribution of γ:

$$
\gamma^P | \phi^P, \phi^C, \beta, \kappa^2, \sigma^2, \sigma_{pc} \sim N\left(\gamma_*^P, H_{\gamma^P}^{-1} \right) \tag{A.8}
$$

where

$$
\begin{aligned}
H_{\gamma^P} &\equiv w^{P\prime} w^P + \underline{H}_{\gamma^P} \\
\widehat{\gamma}^P &\equiv \left(w^{P\prime} w^P \right)^{-1} w^{P\prime} \phi^P \\
\gamma_*^P &\equiv H_{\gamma^P}^{-1} \left(w^{P\prime} w^P \right) \widehat{\gamma}^P
\end{aligned}
$$

and

$$
\gamma^C | \phi^P, \phi^C, \beta, \kappa^2, \sigma^2, \sigma_{pc} \sim N\left(\gamma_*^C, \kappa^2 * H_{\gamma^C}^{-1} \right) \tag{A.9}
$$

where

$$
\begin{aligned}
H_{\gamma^C} &\equiv w^{C\prime} w^C + \underline{H}_{\gamma^C} \\
\widehat{\gamma}^C &\equiv \left(w^{C\prime} w^C \right)^{-1} w^{C\prime} \left[\phi^C - \sigma_{pc} \mathbf{u}^P \right] \\
\gamma_*^C &\equiv H_{\gamma^C}^{-1} \left(w^{C\prime} w^C \right) \widehat{\gamma}^C
\end{aligned}
$$

and the vectors ϕ^P and ϕ^C are $F \times 1$ vectors that collect the ϕ_f parameters.

A.4   Appendix: Convergence of the Sampler

To assess the convergence of the Gibbs' sampler, we compute the CUMSUM statistic from Yu and Mykland (1998). For a sequence of simulations $\{\theta^n\}_{n=1}^T$, the statistic is:

$$
CS_t = \left(\frac{1}{t} \sum_{n=1}^t \theta^n - \mu_\theta \right) / \sigma_\theta
$$

where μ_θ and σ_θ are the mean and standard deviations taken over all T draws. If the sampler converges, then the statistic will converge smoothly towards zero. We report the

CUMSUM statistics for the estimation in the fourth column of Table 4 in Figure A1. As can be seen, the CUMSUM converges smoothly towards zero for the beta's, gamma's, and the covariance parameters.

A final issue is how long it takes the sampler to "burn-in." To shed light on this, we present Figure A2 which displays the time series of the 1500 draws of the three covariance parameters, σ^2, κ^2, and σ_{pc} from the same specification of Table 4. As can be seen, the sampler reaches the stationary distribution in well under 500 draws.

Table A1: Sample Sizes by Year

	Everybody	Fathers	Mothers	Children
1984	10,231	2,669	4,266	3,296
1985	10,490	2,683	4,295	3,512
1986	10,614	2,688	4,285	3,641
1987	10,779	2,702	4,286	3,791
1988	10,841	2,660	4,261	3,920
1989	10,856	2,631	4,218	4,007
1990	11,289	2,718	4,400	4,171
1991	11,343	2,666	4,346	4,331
1992	11,647	2,668	4,386	4,593
1993	11,899	2,642	4,453	4,804
1994	12,270	2,719	4,543	5,008
1995	12,199	2,659	4,473	5,067
1996	12,161	2,596	4,411	5,154
1997	9,949	2,321	3,742	3,886
1999	10,076	2,259	3,672	4,145
2001	10,360	2,198	3,597	4,565
2003	10,592	2,126	3,534	4,932
2005	10,813	1,964	3,418	5,431
2007	10,969	1,876	3,257	5,836
2009	11,264	1,798	3,171	6,295
2011	11,248	1,658	2,991	6,599
2013	11,260	1,558	2,813	6,889
Total	243,150	52,459	86,818	103,873

Table A2: Frequencies for Time Length in Panel

Years in Panel	Frequency	Percent	Cumulative
1	20,400	8.39	8.39
2	19,240	7.91	16.30
3	18,125	7.45	23.76
4	16,943	6.97	30.73
5	15,711	6.46	37.19
6	14,815	6.09	43.28
7	13,984	5.75	49.03
8	13,197	5.43	54.46
9	12,496	5.14	59.60
10	11,709	4.82	64.41
11	11,124	4.57	68.99
12	10,591	4.36	73.34
13	9,867	4.06	77.40
14	8,005	3.29	80.69
15	7,526	3.10	83.79
16	7,074	2.91	86.70
17	6,630	2.73	89.43
18	6,171	2.54	91.96
19	5,693	2.34	94.30
20	5,203	2.14	96.44
21	4,694	1.93	98.37
22	3,952	1.63	100.00

Table A3: Reduced form Estimates of Rank-Rank Correlations

	(1)	(2)	(3)	(4)	(5)
Rank-Rank Correlation	0.281***	0.242***	0.172***	0.163***	0.204***
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Age	X	X	X	X	X
Sex	X	X	X	X	X
White	-	X	X	X	X
Black	-	X	X	X	X
Years of Schooling	-	-	X	X	-
Family Income	-	-	-	X	X
F	4882	4882	4882	4882	4882

* significant at the 10% level; ** significant at the 5% level; *** significant at the 1% level
Notes: We report estimates of the rank-rank correlation based off of the "reduced form" model in Section 4. We restricted the samples to values for which the parent's rank was between the 10th and 90th percentile. Robust standard errors are reported in parentheses.
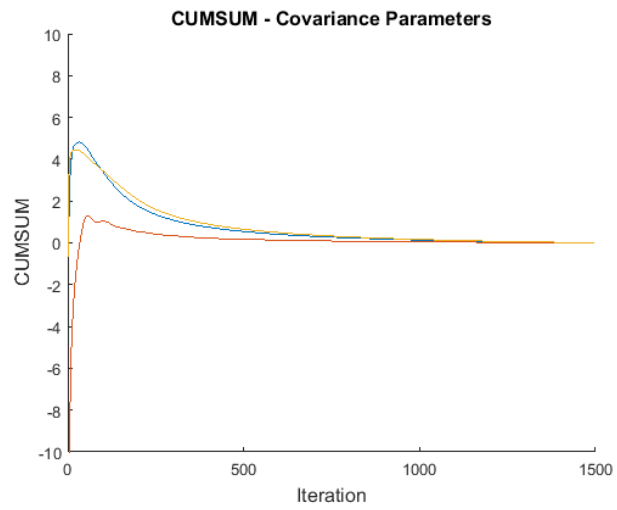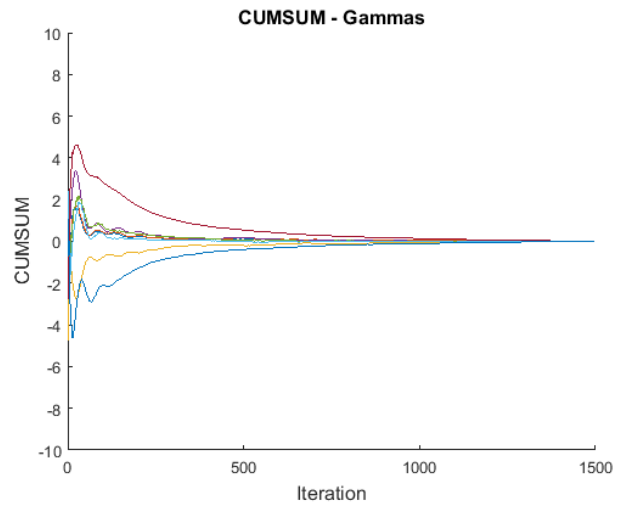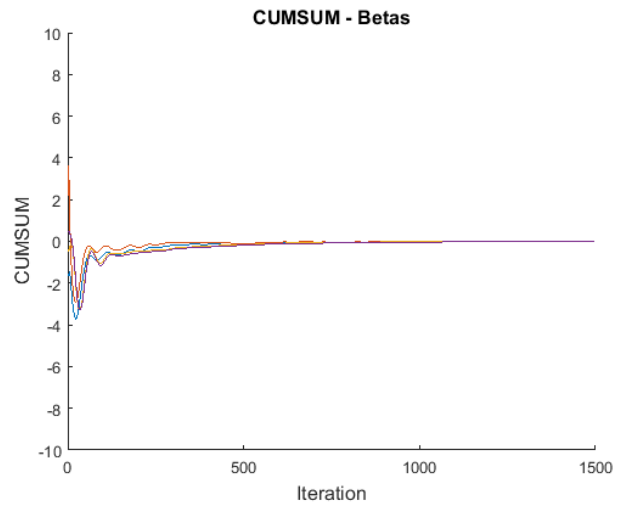
Figure A1: CUMSUM Plots

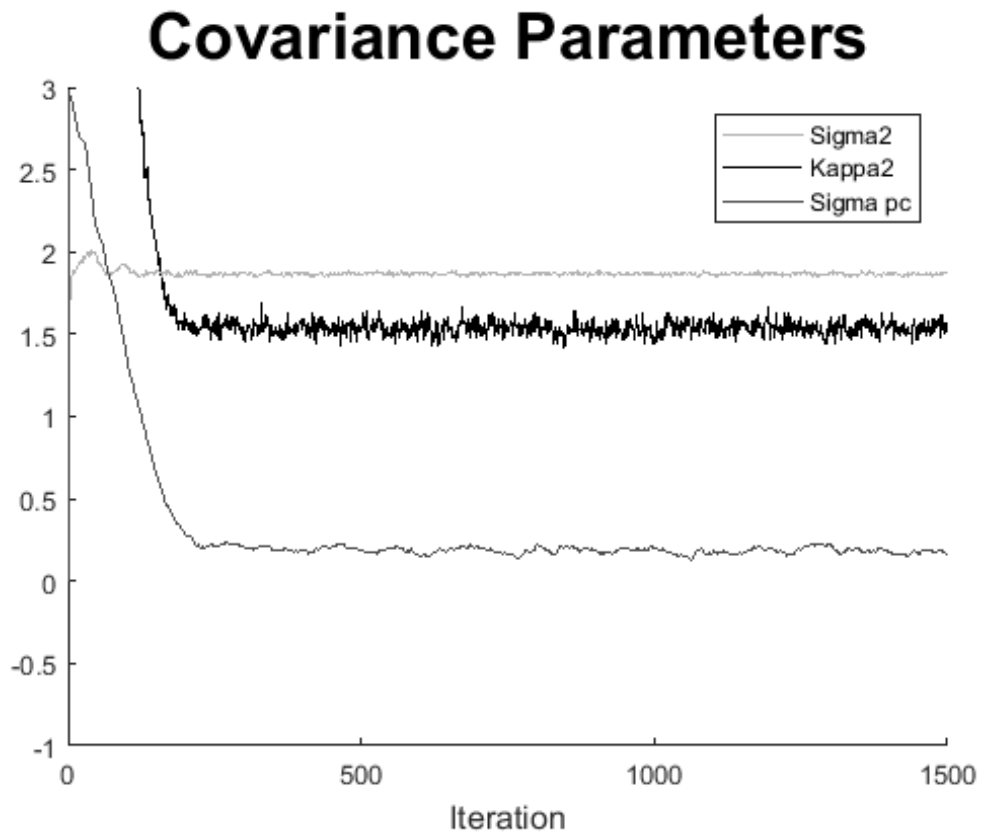Figure A2: Time Series Plot of Covariance Parameters

Figure A3: Additional Unconditional Rank-Rank Regressions

Race   Includes Race and Education   Includes Race and Income

Homogeneous Cuts

Heterogeneous Cuts