Appendices

A VERY UNEVEN PLAYING FIELD:
ECONOMIC MOBILITY IN THE UNITED STATES

Pablo A. Mitnik
Stanford Center on Poverty and Inequality
pmitnik@stanford.edu

Victoria Bryant
Schar School of Policy and Government, George Mason University
vbryant@masonlive.gmu.edu

David B. Grusky
Stanford Center on Poverty and Inequality
grusky@stanford.edu

June, 2018

The Stanford Center on Poverty and Inequality is a program of the Institute for Research in the Social Sciences at Stanford University.
CONTENTS

B. Survey-Data Evidence on Nonlinearities
C. Construction of Income and Parental-Age Variables
D. Construction of Sampling Weights
E. Comparison between the Base Sample of the SOI-M Panel and the CPS
F. Additional Information on Nonparametric and Spline-Based Estimation and Bootstrap-Based Inference
G. Use of Sampling Weights in Model Estimation
H. Evidence on Attenuation Bias
I. Evidence on Lifecycle Bias
J. Mechanical Relationship between the IGEs, Dispersion and the Rank-rank Slope
K. Additional Robustness Checks
L. Earnings Elasticities and the Missing Poor
M. Supplementary Estimates of Region-Specific IGEs
N. Productivity of Investments in Human Capital in Denmark and the United States

In an influential paper, Corak (2006) carried out both (a) a quantitative meta-analysis of the existing estimates in the literature of the (constant) IGE of men’s earnings, and (b) a qualitative analysis of that literature. We argued in the main text that these analyses were strongly influenced by the results published in Mazumder (2005). This is for two reasons. First, Corak (2006) selected Grawe’s (2004) PSID-based estimate of 0.47 as the preferred estimate for the IGE of men’s earnings, largely because this value was consistent with Mazumder’s estimates with a similar number of years of parental information. Second, when we repeated Corak’s (2006) meta-analysis excluding Mazumder’s (2005) results, we obtained substantially lower predicted IGEs. We provide here additional details on Corak’s meta-analysis and our follow-up analysis.

The right-hand variables in Corak’s meta-analysis were the number of years of parental information employed, the age at which father’s earnings are measured, and a dummy variable indicating whether OLS or IV estimation was conducted. Because Corak included only those studies in which all three of these right-hand variables were available, his meta-analysis is based on 22 of the 41 estimates he listed in his review. We repeated Corak’s (2006) meta-analysis using the same right-hand variables but kept only the survey-based estimates. That is, we excluded the administrative-data results reported in Mazumder (2001), which is a preliminary version of Mazumder (2005). As in Corak’s analysis, we generated predicted IGEs under the assumptions that (a) father’s earnings were measured around age 45, and (b) 5, 10, and 15 years of parental information were used. Table A1 shows that Corak’s (2006) predictions are substantially higher than those we obtain after excluding Mazumder’s estimates from the meta-analysis.

B. Survey-Data Evidence on Nonlinearities

We noted in the main text that, because of sample-size constraints, only a few researchers have used survey data to examine possible nonlinearities in the relationship between parents’ and
children’s income and earnings in the United States (in log-log space). We also noted that the results of these studies were inconclusive. Here we discuss the evidence behind those claims.

In his seminal PSID-based paper, Solon (1992) made a heroic attempt to assess nonlinearities with two very small samples, an attempt that proceeded by adding the square of log parental earnings into the standard specification. The corresponding point estimates suggested that the IGE of men’s earnings increased with parental earnings, but those estimates did not reach statistical significance (Solon 1992, p. 404). In another analysis of PSID data using the same approach, Behrman and Taubman (1990) reported statistically significant results that were qualitatively similar to Solon’s, but they were obtained by pooling men and women. By contrast, Mulligan’s (1997) analysis of PSID data did not find evidence of nonlinearities, which he suggested might be the result of the PSID’s underrepresentation of very rich individuals. Also using the PSID, Hertz (2005) analyzed the Black and White subsamples separately, employing nonparametric methods to uncover the relationship between children’s and parents’ log income. The estimated curves show a convex pattern that is particularly marked for Blacks, but Hertz neither estimated the pooled curve nor provided any inferential information for the curves he did estimate.

The NLS-based findings on nonlinearities are also inconclusive. Although some scholars have reported results indicating that the men’s earnings IGE increases with father’s earnings (Lillard 2001; Bratsberg et al. 2007; Couch and Lillard 2004), those results appear highly sensitive to the details of the specification, indeed several specifications suggest that the IGE decreases in at least some ranges of father’s earnings. There is also very substantial variability across the studies in the implied values of the IGEs under the estimated curves.

It follows that neither the PSID nor the NLS have delivered clear evidence on the question at hand. To be sure, the available studies suggest convex curves on balance, but the evidence is very far from definitive. Moreover, the approaches typically employed to assess nonlinearities relax the
constant-elasticity assumption in a quite limited way, with the implication that the actual pattern of nonlinearities in the data remains ambiguous.

C. Construction of Income and Parental-Age Variables

We provide here additional information on the income and parental-age variables used in our analyses. We discuss (a) how the children’s parents were defined from the information available in tax returns and the rules used to pool income across parents’ returns (when necessary), and (b) the data used to conduct mean imputation for nonfiler children with no W-2 earnings and no UI income.

Parental income and age. The person or persons who claim a child as a dependent in 1987 were defined as the child’s parents. If only one nondependent adult claimed the child (i.e., there was no “secondary filer” in the return), that nondependent adult was typically defined as the (single) parent of the child. However, whenever the adult claiming the child was married and the spouse filed separately, both spouses were defined as parents. It is of course possible that one or both parents of a child will in subsequent years file with someone else (as with most divorces). In that case, annual parental income was computed by pooling resources across the relevant returns, where the pooling is based on the following rules:

(a) If the two parents divorced, and each filed jointly with a new spouse, parental income was defined as the sum of half the income appearing in each parent’s return (as pooling the full income appearing on the returns for each of the remarried parents would have overstated the income of the child’s parents).

(b) If only one of the two parents filed with a new person, that parent’s imputed income (calculated again by dividing by two) was combined with the full income of the other parent.

(c) If a parent was single in 1987 but later married, the pooled income of the parent and his or her spouse was used.
Parental age in each year was computed by averaging the age of the parents listed in any return that year. Average parental income and age when the children were 15-23 years old were computed by averaging the corresponding annual measures.

**Imputations for nonfiler children.** The Annual Social and Economic Supplement of the Current Population Survey (CPS-ASEC) identifies likely nonfilers using a tax simulation model. Although this information is available for the entire period covered by the SOI-Panel, the CPS-ASEC data after 2003 have serious inconsistencies and cannot be used. We therefore used pooled CPS-ASEC data from 1999 to 2003 to compute the mean income of nonfilers (without earnings or UI income) by gender and age group. The resulting values, which we used for mean imputation, are as follows (all in 2010 dollars): 26-30 year-old men: $4,910; 31-35 year-old men: $5,815; 36-40 year-old men: $6,706; 26-30 year-old women: $5,372; 31-35 year-old women: $6,574; 36-40 year-old women: $7,560.

**D. Construction of Sampling Weights**

As indicated in the main text, the SOI Family Panel is based on a stratified random sample of 1987 tax returns, with a sampling probability that increases with income. The sampling weights of the SOI-M Panel take into account (a) the foregoing sampling probability from the SOI Family Panel, (b) the probability that a return will enter the refreshment segment of the OTA Panel, and (c) estimates of the probability that a dependent child will have a valid Social Security number (and hence will be included in the base sample of the SOI-M Panel) given that he or she is included in a “qualifying return” (i.e., a 1987 return from the SOI Family Panel or a return from the refreshment segment of the OTA Panel). The latter estimates were computed separately for each sampling stratum, with the refreshment segment of the OTA Panel treated as a separate sampling stratum.

**E. Comparison between the base sample of the SOI-M Panel and the CPS**

Because the SOI-M Panel is new, it is especially important to compare it against known high-quality samples, such as the CPS. In Table A2 and Figure A1, we evaluate the representativeness of
the base sample of the SOI-M Panel against 1987 data from the CPS-ASEC. The purpose of Table A2 is to show that the weighted age distribution and gender distribution in the SOI-M Panel approximate well the corresponding distributions in the CPS-ASEC. It is equally important to compare the 1987 weighted family income distributions in the base sample and in the CPS-ASEC (for children 12 to 15 years old). Although Figure A1 reveals that the two distributions are very similar, an important difference is that the share of children at the extreme right tail in the SOI-M Panel is almost twice as large as the corresponding CPS-ASEC share. This result is consistent with research documenting an underreporting of top incomes in the CPS (Fixler and Johnson 2014). We also find that the SOI-M, as compared to the CPS-ASEC, has a larger share of children in families with parental income between $10,000 and $30,000 but a smaller share of children in families with parental income between $50,000 and $70,000. These differences are relatively minor and, overall, the results are again satisfactory. It is especially reassuring that there is no evidence of a deficit of poor children in the SOI-M Panel.

F. Additional information on nonparametric and spline-based estimation and bootstrap-based inference

We provide here some additional information regarding nonparametric and spline-based estimation and bootstrap-based inference. For reasons that will become clear, we also explain here why we opted for using mean instead of multiple imputation for nonfilter children with no W-2 earnings and no UI income.

Nonparametric estimation of global and region-specific elasticities. We employ numerical approximations to estimate either a global or region-specific $\text{IGE}_e$ when nonparametric models are used. The estimated nonparametric curve is divided into 196 segments (between the 1st and the 99th percentiles of parental income) such that each segment covers the same share of children in the population (i.e., 0.5 percent). We then use finite differences in logarithms to approximate the average point elasticity in each segment. This allows us to compute the global and region-specific elasticitie
by averaging the estimated elasticities across all relevant segments. Our computation of the global and region-specific IGEs ignores the curve’s final left and right segments (each covering 1 percent of children). Because the curve is estimated less precisely at the boundaries, these “trimmed estimators” are often more efficient. The point estimates from the trimmed and untrimmed estimators are, however, very similar.

**Estimation of global arc elasticities.** We employ a numerical approximation to estimate the global AIGE. This summary persistence measure can be written as:

$$\text{Global AIGE}_e = \int_0^{100} \int_0^{100} AIGE_e(Y|x, \min(u, v), \max(u, v)) \, du \, dv,$$

where to simplify notation we have ignored that the space of integration has to exclude those cases in which $u = v$. We approach the value of this integral by computing:

$$\text{Global AIGE}_e \approx \frac{1}{40,000} \sum_{i=1}^{40,000} AIGE_e(Y|x, \min(u, v)_i, \max(u, v)_i),$$

where the $(u, v)_i$ are obtained by randomly drawing one value from each of 40,000 cells spanning the space of integration. These cells are defined so that each contains the same share of “pairs of children.” The procedure is as follows. First, the estimated curve is divided into 200 segments, such that each segment covers the same share of children in the population (i.e., 0.5 percent). A matrix with 40,000 cells is then generated by cross-tabulating a variable indexing the segments of the curve with itself. Next, for each cell, a value of parental income is randomly drawn from each of the two segments of the curve defining the cell. With these 40,000 pairs of parental-income values in hand, we compute for each one the pair of corresponding conditional children’s expectations. We do so in each case by linearly interpolating the needed expectations from the estimated expected values associated to the values of parental income defining the corresponding segments of the curve. Finally, we compute the arc elasticity for each sampled pair of families, and the average arc elasticity across all pairs. The latter is our estimate of the global AIGE.
Bootstrap-based inference. In analyses based on the spline and nonparametric models, all inference (save the test of the constant-elasticity assumption, which is described in the note to Table 3) relies on the nonparametric bootstrap, using 2,000 bootstrap samples. We generate bootstrap samples via simple random sampling of primary sampling units, with replacement, within strata. When estimating any nonparametric model with bootstrap samples, we keep the smoothing parameter fixed at the value selected for the original sample. This is equivalent to keeping the functional form fixed when carrying out bootstrap-based inference with parametric models (Racine and Parmeter 2014, p. 313 and note 12).

To assess the uncertainty of our spline-model estimates, we use the percentile method to compute confidence intervals (Efron and Tibshirani 1986). However, we can only compute “variability bounds” (Racine 2008) or “confidence bands” (Wasserman 2006) for our nonparametric estimates (for which we also use the percentile method). Indeed, with a nonparametric regression the bootstrap does not deliver true confidence intervals, as bias does not disappear asymptotically. It follows that the bootstrap-generated intervals are not centered around the true “parameters.” Under the reasonable assumption that the bias is small, the variability bounds nonetheless provide approximations to true confidence intervals (e.g., Wasserman 2006, p. 89).

We test one-sided null hypotheses by computing “type-2 p-values,” which were developed specifically for bootstrap-based tests (Singh and Berk 1994; see also Liu and Singh 1997; Efron and Tibshirani 1998). These type-2 p-values, which we call “p-values” for simplicity, are computed as the proportion of bootstrap samples in which the null hypothesis is true. They can be interpreted as standard p-values.

Mean imputation instead of multiple imputation. Nonparametric estimation, numerical approximation, and bootstrap-based inference are all computer intensive. Using them together, as we do here, is therefore very demanding, as it requires conducting nonparametric estimation and numerical approximation with each bootstrap sample. In the context of bootstrap-based inference,
multiple-imputation needs to be nested within the bootstrap resampling process, thus multiple imputation greatly exacerbates “computing costs.” For instance, conducting bootstrap inference with 2,000 samples and five imputed datasets would have required us to repeat 10,000 times the nonparametric estimation of each model we estimate, plus the associated numerical approximations. This may be compared to repeating it a mere 2,000 times with mean imputation. For this reason, we opted for using mean imputation instead of multiple-imputation when imputing the income of nonfiler children with no W-2 earnings and no UI income. The use of mean imputation should lead to a small underestimation of the uncertainty of our estimates (Little and Rubin 2002).

G. Use of Sampling Weights in Model Estimation

Although there are circumstances in which the use of weights in model estimation is not the best approach (e.g., Winship and Radbill 1994), we apply sampling weights in all of our analyses. This decision to use weights is partly an empirical one. As Pfeffermann and Sverchkov (2009, p. 455) clearly put it, if the sampling weights are “related to the values of the model outcome variable even after conditioning on the model covariates,” ignoring the information contained in sampling weights “may yield large biases and erroneous inference.” When we tested for bias in constant-elasticity models without weights (see Du Mouchel and Duncan 1983; Nordberg 1989), the null hypothesis of no bias was systematically rejected, a result that informed our decision to apply weights in our analyses.

As is well known, using weights in model estimation entails a loss of precision (e.g. Winship and Radbill 1994), a loss that increases when the weights cover a wide range (e.g., Skinner and Mason 2012). This is unfortunately the case in the SOI-M Panel (i.e., final weights go from 1 to 5,400). In our analyses, we carry out standard Horvitz-Thompson (e.g., Fuller 2009) estimation with sampling weights, which is the usual approach employed in the social sciences. Alternative approaches proposed by Pfeffermann and Sverchkov (1999) and Skinner and Mason (2012), which
were developed to improve precision when weights cover a wide range, did not lead to consistent improvements in precision with our data.

**H. Evidence on Attenuation Bias**

Since the early 1990s, mobility researchers sought to reduce attenuation bias by using averages of 3-5 years of parental income or earnings, with the assumption that such averages proxied reasonably well for lifetime average parental income or earnings. In his now-classic article, Mazumder (2005) argued that averaging over 3-5 years was not nearly enough, indeed he suggested that about 16 years of parental information, and perhaps more, are needed to eliminate or nearly eliminate attenuation bias. When Mazumder used progressively more years of parental information (from SSA records), the estimated earnings elasticities increased substantially. This research led to a growing consensus that a long time series of parental data was needed to secure good estimates. Against this consensus, Chetty et al. (2014) reported that (constant-elasticity) estimates from tax data nearly stabilize once 5 years of information are employed: The income IGE_{g} increased only 6.4 percent, from 0.344 to 0.366, when Chetty et al. used 15 years of parental information instead of 5 (2014, Table 1 and Online Appendix E).

We revisit here the issue of attenuation bias with the SOI-M Panel and the IGE_{e}. We proceed, as is conventional, by examining how constant-elasticity estimates change as additional years of parental information are incorporated. We complete our analyses with two different approaches, one based on a common sample, and another based on common sample-inclusion rules (shortened to “common rules” from hereon in).\textsuperscript{5}

The goal of the first approach is to use the same sample as we increment the number of years of parental information (for any income measure and gender). We would of course ideally use precisely the samples employed in our main research to reestimate the IGE_{e} as we successively increment the number of years of parental income. However, because a 9-year average of parental income may be positive while an average based on fewer years may not, some observations have to
be dropped when using parental variables computed with fewer than 9 years. As a result, the goal of this approach cannot be fully attained, but it can be approximated to a very large degree.

Under the common rules approach, we apply our sample-selection rules with each of the parental measures (based on one through nine years of parental information), with the implication that the samples will not likely be the same across these measures. The observations that are excluded, for example, because parental income is above $7,000,000 for a 3-year average may stay in the sample when the average is computed with a different number of years. A further complication arises when applying our sample-inclusion rule that children should have at least 6 years of parental information available (when they are 15-23 years old). This rule may be interpreted to require an absolute number of 6 years or to require two-thirds of the maximum number of years. We implemented both interpretations, but here we only report the results generated under the second (as the results are much the same for either approach).\textsuperscript{6}

We start by examining the results obtained with the common sample approach. In Figures A2 and A3, we present the estimates of the total-income IGE\textsubscript{e} and earnings IGE\textsubscript{e} for men, and of the total-income IGE\textsubscript{e} for women, respectively. We also present corresponding total-income estimates after pooling the samples for men and women (see Figure A4).\textsuperscript{7}

For both genders, the income and earnings estimates increase over the full range of years of parental information, but they increase more rapidly between one and 4-5 years of parental information than for further increases. The estimates for women increase only marginally after 7 years of information, whereas the estimates for men increase at a slower rate after four years but then jump up noticeably between the 7-year and 8- and 9-year parental measures. For men and women pooled, the estimates increase smoothly at a markedly decreasing rate over the full range of years of parental information.

We next consider the common rules approach (see Figures A5-A7). With this approach, the total bias that results from using a one-year instead of 9-year variable is much smaller than with the
common sample approach, regardless of sample (men, women, all) and income measure. The relationship between the estimated elasticities and the number of years of information is also closer to linear. The estimates, however, still increase substantially in all cases when comparing 4-5 years of parental information with 8-9.

The evidence in Figures A2-A7 suggests a plateau by year 9. Although we might conclude that the 9-year measure eliminates the bulk of attenuation bias, we cannot rule out that the curves continue growing very slowly but without reaching any plateau (or reaching it much later). There is also an alternative explanation for the tapering-off that emphasizes the deteriorating quality of the additional parental years that are incorporated. After 6 years of parental information are used to compute parental income in the SOI-M data, measures that include additional years go “in the wrong direction,” as they pertain to parents who are increasingly advanced in their earnings lifecycle. As Mazumder (2005, pp. 247-248) noted, attenuation bias is best reduced by adding parental information from parents’ prime-age period, not by adding information when they are in their fifties. It is possible that the results that Chetty et al. (2014) report are likewise affected by the noisiness of the additional years they are incorporating.

In summary, our evidence indicates that (a) attenuation bias is greatly reduced by using 9 years of parental information, (b) it is possible that some bias still remains, and (c) a decision to use 5 years instead of 9 years would result in a non-negligible increase in bias.

I. Evidence on Lifecycle Bias

We consider here “left-side” lifecycle bias. This refers to the lifecycle bias that arises because children from different parental-income backgrounds have different age-income or age-earnings profiles. In particular, measuring children’s income or earnings too early in their lifecycle leads to underestimating the corresponding elasticity. As we indicated in the main text, there is evidence suggesting that estimating IGEs with parents’ and children’s information taken around age 40 come closest to representing lifetime IGEs (see Mitnik 2017 for the IGE of the expectation; and Haider and
Solon 2006; Böhlmark and Lindquist 2006; Nyborn and Stuhler 2016 for the conventional IGE).

Chetty et al. (2014) claim, however, that income IGEs stabilize around age 30, both in the case of the IGEg (p. 1580 and Online Appendix Figure IIa) and the IGEe (Online Appendix C and Online Appendix Figure Ib). We revisit this issue here with our SOI-M data.

Unlike in the attenuation-bias analysis, here we have included confidence intervals in all figures, as this helps interpret some of the results. Also to facilitate interpretation, we have added a second horizontal axis at the top of the figures, showing the mean age across our four cohorts in each year. We begin by considering lifecycle bias for earnings elasticities. In Figure A8, we present our estimates of the (constant-elasticity) IGEe of men’s earnings from 2001 to 2010. This figure, which reveals that the men’s earnings IGEe rises swiftly and nearly continuously, is in close agreement with the findings in the previous literature.

Is there evidence of an earlier stabilization when we turn to total-income elasticities? In Figures A9 and A10, we present our estimates of the total-income IGEe for men and women respectively, again from 2001 to 2010. The elasticities for women grow at first but indeed seem to stabilize when women are in their early 30s (although there is a very small uptick at the end of the series). Although the elasticities for men exhibit a pattern that may seem more difficult to interpret, the key consideration to keep in mind is that they are likely affected by period as well as lifecycle effects. We advance the twofold hypothesis that (a) the dip in the income IGEe in 2008 and 2009 is the result of the short-term income compression produced by the Great Recession, and (b) the subsequent uptick of the income IGEe in 2010 reflects the growth of inequality and the restoration of more nearly normal age-income profiles. Based on a larger sample that is less affected by sampling variability, the results for all children, presented in Figure A11, exhibit a clearer pattern that is consistent with this hypothesis. The figure shows that the IGEe increases quite smoothly between 2001 and 2007, falls in 2008 and 2009, and then returns to its pre-recession level in 2010.
Given the evidence in the previous literature and our results, it seems clear that measuring the elasticity of men’s earnings when they are relatively young should generate substantial lifecycle bias. In addition, taking into account that the cohorts represented in the SOI-M Panel were 29-32 years old in 2004, our results for the pooled income IGE_e suggest that estimates of that elasticity around age 30 involves a substantial amount of lifecycle bias.

J. Mechanical Relationship between the IGE_e, Dispersion and the Rank-rank Slope

In the main text, we stated that we anticipated a low earnings elasticity for women (compared to men’s), as women from relatively affluent backgrounds tend to have higher-income partners and to work fewer hours (or not at all) when they have young children (Raaum et al. 2007). We now provide a second justification for this expectation that is based on the mechanical relationship between the IGE_e, dispersion, and the rank-rank slope.

It is well known that the (constant) IGE_g is equal to the product of (a) the linear correlation between children’s and parents’ log incomes, and (b) the ratio between the standard deviations (SD) of the children’s and parents’ log incomes. Similarly, we will show here that the (constant) IGE_e (a) increases with the SD of children’s income and with the correlation between children’s income and log parental income, and (b) decreases with the SD of log parental income.

Mitnik (2017) has shown that the probability limit of the PPML estimator of the constant IGE_e, \( \alpha_1 \), is approximately equal to:

\[
\alpha_1 \approx \frac{1}{\text{Cov}(Y, \ln X)} - \left[ \frac{1}{\text{Cov}(Y, \ln X)} \right]^2 - \frac{2}{\text{Var}(\ln X)} \right]^{1/2},
\]

where \( \left[ \frac{1}{\text{Cov}(Y, \ln X)} \right]^2 > \frac{2}{\text{Var}(\ln X)} > 0 \). We may then write:

\[
\alpha_1 \approx \frac{1}{SD(Y)SD(\ln X)\text{Corr}(Y, \ln X)} - \left[ \frac{1}{SD(Y)SD(\ln X)\text{Corr}(Y, \ln X)} \right]^2 - \frac{2}{[SD(\ln X)]^2} \right]^{1/2}
\]
It is apparent that $\alpha_1$ falls when $SD(ln X)$ increases. The partial derivative of $\alpha_1$ with respect to $A$ is:

$$\frac{\partial \alpha_1}{\partial A} = 1 - \frac{A}{(A^2 - 2)^{\frac{1}{2}}} < 0.$$ 

As $A$ is inversely related to both $SD(Y)$ and $Corr(Y, ln X)$, it follows that $\alpha_1$ increases with both $SD(Y)$ and $Corr(Y, ln X)$.

The rank-rank slope between two variables in levels is closely related to the linear correlation between the logarithms of the variables (Chetty et al. 2014, p. 1561). It’s reasonable to assume that the correlations between children’s earnings and log parental income are ordered across genders in the same way in which the rank-rank slopes and the correlations between the logged income variables are ordered. From the fact that the rank-rank slope relating parental income to children’s individual earnings is flatter for women than for men (Chetty et al. 2014, Table 1), it follows that $Corr(Y, ln X)$ should be smaller for women than for men. In addition, the distribution of women’s earnings is substantially less dispersed than that for men (e.g., Weinberg 2007, Table 5), so $SD(Y)$ can be expected to be smaller for women than for men. Therefore, as $SD(ln X)$ should be approximately the same for men and women, we expect the women’s IGE to be much smaller than the men’s IGE.

K. Additional Robustness Checks

The purpose of this appendix is to describe some of our supplementary analyses of robustness. We first consider whether controls for parental age should be included. In our baseline
analyses, these controls were always omitted, as the age at which parents have their children is not likely to be exogenous to their income. Because high-income parents are more likely to delay childrearing, and because children with older parents have better life chances and higher lifetime income (e.g., Liu et al. 2011, Myrskylä and Fenelon 2012, Myrskylä et al. 2013, Powell et al. 2006), some of the association between parental income and children’s income is a consequence of those delay-of-childrearing decisions. When other scholars control for parental age, they are eliminating that indirectly-generated portion of the total association, and their estimates may therefore be downwardly biased. For this reason, our baseline approach omits controls for parental age, even though other scholars often include them to account for age-related differentials in the measurement error for lifetime income.

These concerns with endogeneity are likely, we think, to trump any benefits that accrue to accounting for differential error in proxying lifetime income. It is nonetheless reassuring that the position one takes on this issue does not much matter. In Table A3, we present estimates from the constant elasticity and spline models (see Equations [3] and [4] respectively), where those models now include a third-degree polynomial in parental age. This table also provides p-values from the corresponding constant-elasticity tests.\textsuperscript{10} We find that the results with the parental age controls are much the same as our baseline results and, moreover, consistent with other key conclusions reached on the basis of the estimates in Table 3 (and Figure 2). Most importantly, the constant-elasticity assumption is still rejected, and the estimates for women are still lower than the corresponding ones for men.

In a second set of supplementary analyses, we reconsider the standard approach of calculating total-income IGES, an approach that may mislead because total income is an imperfect measure of the capacity to consume and invest. Because this capacity is more directly measured by disposable income than total income, we have used the SOI-M Panel to provide an approximate measure of disposable income (by subtracting net federal taxes from total income).\textsuperscript{11} This measure
can then be used to answer the following simple question: Does the conventional total-income IGE misrepresent the extent to which economic advantage is transmitted across generations?

In Figure A12, we compare our estimates of total-income and disposable-income global elasticities, again applying our baseline spline and nonparametric specifications (for men and women). The estimates are very similar across these two income measures: When averaged across models, the effect of using disposable-income measures is to reduce the IGE by about 5 percent for men and by less than 3 percent for women. The total-income IGE, while slightly higher than its disposable-income counterpart, does not misrepresent in any fundamental way the persistence of economic status.

L. Earnings Elasticities and the Missing Poor

In this appendix, we examine the effects of the “missing poor” on the estimation of the IGE of men’s individual earnings, thus supplementing our analysis of selection bias for total-income elasticities (as presented in the main text). The IGE estimates presented in the main text were based on the assumption that those without reported W-2 earnings have zero earnings. We were able to retain offspring with zero earnings in our analyses because the IGE, unlike the conventional IGE, is defined for variables including zero in their support. This assumption is almost surely false because some people without W-2 earnings receive earnings from work in the informal economy. Even so, our hypothesis is that it provides a good approximation, and that IGE estimates are robust to the exact assumption that is made as long as this assumption is consistent with the notion that the typical earnings of those with zero W-2 earnings are low and do not vary much by parental income. We test this hypothesis here. In doing so, we also assess the effects of following the practice, conventional in the literature, of dropping not only nonearners but also low earners from the analysis.

While conventional IGE-based analyses yield wildly different estimates under different assumptions about the earnings of nonearners (Mitnik et al. 2015, Tables 12 and 13), the nonparametric estimates of the global IGE, presented in Table A4, are very robust. In fact, when
low-earners are retained, the estimates hardly change under the various assumptions about the mean earnings of those without reported earnings. The estimates do fall substantially when low-earners are excluded (something that also happens when the estimand is the IGEg), attesting to the importance of avoiding this conventional practice.

The foregoing results thus indicate that excising nonearners and low-earners produces downward biases, but that, as long as nonearners are retained, estimates of the global IGEc are very robust to the values imputed to those with zero W-2 earnings. This implies that those estimates are likely to be close to the mark even though W-2 reported earnings exclude earnings from work in the informal economy. We can conclude that using administrative information does solve the problem of the missing poor for the estimation of the IGE of men’s earnings, but only as long as that elasticity is the IGEc.

M. Supplementary Estimates of Region-Specific IGEs

We present here supplementary estimates of all region-specific IGEs after dropping six observations (four for men and two for women). These observations had a very disproportionate influence on the spline-based estimates for the below-P10 region (as indicated in the main text). The spline-model estimates are presented in Table A5, while the nonparametric estimates are presented in Table A6.

N. Productivity of Investments in Human Capital in Denmark and the United States

We pointed out in the main text that, although it is difficult to compare the efficiency of the U.S. and Denmark in transforming money into human capital, the available literature suggests that Denmark is not likely to be, overall, any less efficient. We briefly discuss the evidence here.

The key results are in Bogetofl et al. (2014, Table 2), where it is reported that (a) the cost per primary student was similar in Denmark and the United States, (b) the cost per secondary student (lower and upper) was somewhat larger in the United States than in Denmark, (c) the cost per tertiary student was substantially larger in the United States than in Denmark, and (d) graduation rates were
very similar or somewhat higher in Denmark. In a related article, Hanushek and Woessmann (2011, Figure 2.3) show that the cumulative educational expenditure per student ages 6-15 is higher in the United States than in Denmark, but that the math performance of students, as measured by their PISA scores, is higher in the latter. Sutherland et al. (2009, Figure 1b) report a similar result for student performance in the combined reading, scientific, mathematical, and problem solving PISA scales.

Some studies have employed the data-envelopment-analysis (DEA) technique (e.g., Thanassoulis et al. 2008) to compare the relative efficiency of different countries. Bogetofl et al. (2014, Table 5) report that Denmark and the United States are very similar in terms of the efficiency of their overall educational systems when the output in consideration is the number of students enrolled, but that Denmark is more efficient than the United States at the secondary-education level. They also report that Denmark is in general more efficient than the United States when the output in consideration is the graduation rate in a cohort, although they are similarly efficient at the upper secondary education level when differences in the PISA test scores of the entering students are taken into account (Bogetofl et al. 2014, Table 7). Lastly, they report that the two countries are similarly efficient or that the U.S. is more efficient (depending on education level), when the output in consideration is the completion rate in relation to the number of students who begin an educational program (Bogetofl et al. 2014, Tables 9 and 10).

Using the same technique, Dolton and Marcenaro-Gutierrez (2016, Table 2) report that both countries are similarly efficient when the output in consideration is PISA scores. Sutherland et al. (2009, Table 2) report estimates of efficiency at the school level across countries, using a synthetic indicator of the average school-level PISA scores of 15-year-old students. Depending on the exact specification, the median school in Denmark is slightly or somewhat more efficient than the median school in the United States. Several other results they report are also consistent with the notion that the United States is not more efficient than Denmark.
Notes

1 For example, some of Couch and Lillard’s specifications (2004, Tables 8.4 and 8.5) imply a decreasing IGE across quintiles of father’s earnings, whereas others imply an IGE that first increases and then decreases.

2 When Couch and Lillard (2004) modify the usual specification by adding a quadratic term in the log of father’s earnings, the average elasticity within each of the quintiles of father’s earnings are 0.12, 0.21, 0.23, 0.25, and 0.29 (from the lowest to the highest). When Bratsberg et al. (2007) estimate a similar model, they report IGES that are about three times larger: 0.49 (10th percentile of parental earnings), 0.58 (50th percentile of parental earnings), and 0.65 (90th percentile of parental earnings).

3 For this comparison, we use the 1988 CPS-ASEC, as it provides annual family-income information for 1987.

4 The true “parameters” of interest include, for example, the expectation of children’s income at some value of parental income or the expected IGES in some region of parental income.

5 Mazumder (2005) reports results using both approaches (see his Tables 4 and 5).

6 We implemented the second interpretation by allowing, for each parental income variable, the following maximum number of years of missing information: 3 missing years for the 9-year variable; 2 missing years for the 6-, 7-, and 8-year variables; 1 missing year for the 3-, 4-, and 5-year variables; and no missing year for the 1- and 2-year variables.

7 The main reason for producing estimates with pooled data is that the larger sample size makes the trend clearer.

8 It should be noted that Chetty et al. (2014) consider this possibility but reject it on the argument that their Figure IIb (in their Online Appendix) shows that “estimates of mobility are not sensitive to varying the age in which parent income is measured over the range observed in our dataset” (2014a, Online Appendix E). The evidence in the figure in question, however, pertains to the rank-rank slope. The rank of parents may remain the same as they get older even as the differences between their incomes do not.

9 The expression assumes, without any loss of generality, that $E(Y) = E(\ln X) = 1$. There is no loss of generality because this can always be achieved by simply changing the monetary units used to measure income (i.e., by dividing the children’s income variable by its mean, and the parental income variable by the exponential of the mean of its logarithmic values minus one).
We also estimated constant-elasticity and spline models including (a) cohort dummies (thus simultaneously controlling for cohort and children’s age), and (b) cohort dummies and a third degree polynomial in parental age. We have not reported results from these specifications because they are very similar to those that are reported.

It should be recalled that, because we do not subtract state taxes and do not include some non-taxable transfers (e.g., TANF), our measure only provides an approximation to true disposable income. This measure does, however, incorporate the Earned Income Tax Credit (EITC) and other refundable credits.

When estimating nonparametric models with disposable income, we do not select the smoothing parameter by finding the global minimizer of $AIC_c$ (Hurvich, Simonoff, and Tsai 1998) within the range $.08, 1$, as we do with all other nonparametric models. Instead, we reuse the same smoothing parameters selected for the corresponding total income models. Proceeding this way ensures that any difference is due to the change in the income measure (rather than the change in the smoothing parameter).

We tested the null hypothesis that the global $IGE_e$ for disposable income is not smaller than the global $IGE_e$ for total-income. For the nonparametric model, this hypothesis is rejected for men ($p = 0.000$), but it cannot be rejected for women ($p = 0.259$). For the spline model, it is rejected for both men ($p = 0.000$) and women ($p = 0.005$).
References


**Figure A1:** Parental Income in SOI-M Base Sample and CPS-ASEC

![Income distribution chart]

- **Y-axis:** Percent of children
- **X-axis:** Parental income in 1987

- **Legend:**
  - CPS-ASEC
  - SOI-M

The chart shows the distribution of parental income levels across different income brackets for the SOI-M Base Sample and CPS-ASEC in 1987.
Figure A2: $IGE_e$ as a function of years of parental information, men (common-sample approach)
Figure A3: Total-income IGE\textsubscript{a} as a function of years of parental information, women (common-sample approach)
Figure A4: Total-income IGE as a function of years of parental information, all children (common-sample approach)
Figure A5: $\text{IGE}_a$ as a function of years of parental information, men (common-rules approach)
Figure A6: Total-income IGE as a function of years of parental information, women (common-rules approach)
Figure A7: Total-income $\text{IGE}_n$ as a function of years of parental information, all children (common-rules approach)
Figure A8: $\text{IGE}_a$ of men's earnings, 2001-2010
Figure A10: IGEs of women's total income, 2001-2010
Figure A11: \( \text{IGE}_s \) of children's total income, 2001-2010
Figure A12: Global Total- and Disposable-Income IGE

- □ Total Income
- ▲ Disposable Income

<table>
<thead>
<tr>
<th></th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spline Models</td>
<td>0.51</td>
<td>0.46</td>
</tr>
<tr>
<td>Nonpar. Models</td>
<td>0.48</td>
<td>0.52</td>
</tr>
</tbody>
</table>

**Spline Models** | **Nonpar. Models** | **Spline Models** | **Nonpar. Models**
<table>
<thead>
<tr>
<th>Years of parental information</th>
<th>Included IGE Estimates</th>
<th>All</th>
<th>Survey-based</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0.40</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0.46</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>0.52</td>
<td>0.39</td>
<td></td>
</tr>
</tbody>
</table>
Table A2: SOI-M Base Sample and 1987 CPS-ASEC, Weighted Frequencies

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Frequency</td>
<td>Weighted Frequency</td>
<td>Weighted Frequency</td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1,768</td>
<td>1,806,303</td>
<td>1,662,106</td>
</tr>
<tr>
<td>13</td>
<td>1,784</td>
<td>1,745,662</td>
<td>1,656,149</td>
</tr>
<tr>
<td>14</td>
<td>1,875</td>
<td>1,861,325</td>
<td>1,721,148</td>
</tr>
<tr>
<td>15</td>
<td>1,926</td>
<td>1,703,241</td>
<td>1,844,131</td>
</tr>
<tr>
<td>Total</td>
<td>7,353</td>
<td>7,116,532</td>
<td>6,883,534</td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>1,629</td>
<td>1,656,620</td>
<td>1,582,240</td>
</tr>
<tr>
<td>13</td>
<td>1,692</td>
<td>1,586,241</td>
<td>1,581,609</td>
</tr>
<tr>
<td>14</td>
<td>1,729</td>
<td>1,671,537</td>
<td>1,643,910</td>
</tr>
<tr>
<td>15</td>
<td>1,949</td>
<td>1,811,879</td>
<td>1,751,337</td>
</tr>
<tr>
<td>Total</td>
<td>6,999</td>
<td>6,726,277</td>
<td>6,559,096</td>
</tr>
<tr>
<td></td>
<td>Constant-Elasticity Models</td>
<td>Spline Models</td>
<td>P-value from CE test</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-----------------------------</td>
<td>---------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Men's earnings</td>
<td>0.46</td>
<td>0.51</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.40-0.52)</td>
<td>(0.45-0.58)</td>
<td></td>
</tr>
<tr>
<td>Men's total income</td>
<td>0.45</td>
<td>0.48</td>
<td>0.013</td>
</tr>
<tr>
<td></td>
<td>(0.40-0.49)</td>
<td>(0.42-0.55)</td>
<td></td>
</tr>
<tr>
<td>Women's total income</td>
<td>0.41</td>
<td>0.42</td>
<td>0.041</td>
</tr>
<tr>
<td></td>
<td>(0.37-0.46)</td>
<td>(0.37-0.48)</td>
<td></td>
</tr>
</tbody>
</table>

*Note:* The constant elasticity (CE) test is an F-test of the null hypothesis that all coefficients of the spline model of Equation [4], save \( a_0, a_1 \) and the age coefficients, are zero.
<table>
<thead>
<tr>
<th>Included Men</th>
<th>Imputation for Nonearners</th>
<th>IGE_e</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings above $ 3,000</td>
<td></td>
<td>0.45</td>
</tr>
<tr>
<td>Earnings above $ 1,500</td>
<td></td>
<td>0.47</td>
</tr>
<tr>
<td>Earnings above $ 600</td>
<td></td>
<td>0.47</td>
</tr>
<tr>
<td>Earnings above $ 100</td>
<td></td>
<td>0.47</td>
</tr>
<tr>
<td>Earnings above $ 0</td>
<td></td>
<td>0.47</td>
</tr>
<tr>
<td>All</td>
<td>$0</td>
<td>0.56</td>
</tr>
<tr>
<td>All</td>
<td>$1</td>
<td>0.56</td>
</tr>
<tr>
<td>All</td>
<td>$100</td>
<td>0.56</td>
</tr>
<tr>
<td>All</td>
<td>$1,000</td>
<td>0.56</td>
</tr>
<tr>
<td>All</td>
<td>$3,000</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>Up to P10</td>
<td>P10-P50</td>
</tr>
<tr>
<td>----------------------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td><strong>Men's earnings</strong></td>
<td>0.19</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>(-0.10-0.64)</td>
<td>(0.34-0.63)</td>
</tr>
<tr>
<td><strong>Men's total income</strong></td>
<td>0.34</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>(0.07-0.68)</td>
<td>(0.29-0.55)</td>
</tr>
<tr>
<td><strong>Women's total income</strong></td>
<td>0.52</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>(0.27-0.80)</td>
<td>(0.18-0.44)</td>
</tr>
<tr>
<td></td>
<td>Up to P10</td>
<td>P10-P50</td>
</tr>
<tr>
<td>------------------------</td>
<td>-----------</td>
<td>---------</td>
</tr>
<tr>
<td>Men's earnings</td>
<td>0.31</td>
<td>0.56</td>
</tr>
<tr>
<td></td>
<td>(0.03-0.63)</td>
<td>(0.39-0.73)</td>
</tr>
<tr>
<td>Men's total income</td>
<td>0.38</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>(0.23-0.55)</td>
<td>(0.33-0.55)</td>
</tr>
<tr>
<td>Women's total income</td>
<td>0.53</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>(0.31-0.75)</td>
<td>(0.20-0.52)</td>
</tr>
</tbody>
</table>