

# ONE SIZE DOESN'T FIT ALL: A QUANTILE ANALYSIS OF INTERGENERATIONAL INCOME MOBILITY IN THE U.S. (1980-2010)\*

Juan C. Palomino

Universidad Complutense de Madrid (Spain), EQUALITAS and CEDESOG

Tel. +34 916 710 784 E-mail: juancpal@ucm.es

Gustavo A. Marrero

Universidad de la Laguna (Spain), EQUALITAS and CEDESOG

Tel. +34 922 277 798 E-mail: gmarrero@ull.es

Juan G. Rodríguez

Universidad Complutense de Madrid (Spain), EQUALITAS and CEDESOG

Tel. +34 913 942 516 E-mail: juangabr@ucm.es

May 4, 2018

*"Everybody knows... the poor stay poor, the rich get rich"*

Leonard Cohen

## Abstract

Conventional wisdom and previous literature suggest that economic mobility is lower at the tails of the income distribution. However, the few studies that have estimated intergenerational income elasticity (IGE) at different points of the distribution in the U.S. were limited by small samples, arrived at disparate results, and had not estimated the trend of elasticity over time. Using the PSID database, a large sample of income observations in the 1980-2010 period for the U.S. is built, which allows us to obtain robust quantile estimates of the IGE both for the pooled sample and for each wave. For the pooled sample, the IGE shows a U-shaped relation with the income distribution, with higher values at the tails (0.64 at the tenth percentile and 0.48 at the ninety-fifth percentile) and a minimum value –highest mobility– of 0.38 at the seventieth percentile. The trend evolution of the IGE also varies across the income distribution: at the lower and mid quantiles, income mobility increased during the 80s and 90s but declined in the 00s, while for the higher quantiles it remained relatively stable along the whole period. Finally, the impact of education and race on mobility is evaluated. Both factors are found to be important and related to the position at the income distribution.

*JEL classification:* D31, D63 ; *Keywords:* Intergenerational Mobility; Income Elasticity; Quantile Regression

---

\*We are grateful to two anonymous referees for excellent suggestions that contributed to significantly improve the paper. We would also like to thank Nicole Fortin, Roger Koenker and Joao Santos Silva for valuable technical advice and the participants at the 6th ECINEQ Meeting in Luxembourg for helpful comments. The authors acknowledge the financial support of the Ministerio de Economía y Competitividad of Spain, Palomino and Rodriguez through project ECO2016-76506-C4-1-R and Marrero through project ECO2016-76818-C3-2, and from Comunidad de Madrid (Spain) under project S2015/HUM-3416-DEPOPORCM, and Fundación Caja Canarias (Spain) under project CSOCTRA07. The usual disclaimer applies.

# 1 Introduction

The perception that the US is a “land of opportunities” has often served to overlook its levels of income inequality, considering that the economy enjoyed a high level of economic mobility.<sup>1</sup> In the last decades, however, this commonplace perception has been questioned. Studies estimating the connection between parent and child income through the Intergenerational Income Elasticity (IGE) put the level of opportunity in the US into perspective, both comparing it with other nations and, more recently, showing its trend evolution. Thus, the pioneering works of Solon (1992) and Zimmerman (1992) alerted about a much higher value for IGE than what had been obtained in the scarce previous research on this issue.<sup>2</sup> This finding spurred subsequent research analyzing the IGE in the US and around the world, with the US quite consistently ranking higher than other countries with similar degrees of development (Corak (2006); Björklund and Jäntti (2009); Blanden (2013)).

However, partly because of data availability and computational requirements, most IGE studies derive it from a regression-to-the-mean model using ordinary least squares (OLS), and little attention has been paid to the possible differences in the level of elasticity at different points of the income distribution.<sup>3</sup> The few works that have estimated the IGE in the US at different quantiles of the distribution have not found a clear relation between the IGE and the income distribution, and have considered a cross section with relatively small samples (Eide and Showalter (1999); Grawe (2004); Cooper (2011)), which may cast doubts about the accuracy of their estimates. With regard to the trend evolution of IGE, research up to date has focused only on the OLS evolution of IGE and has arrived at disparate results (Hertz (2007); Aaronson and Mazumder (2008); Lee and Solon (2009)).

This paper contributes to the debate about the level and evolution of IGE in the US in three different ways. First, it shows how IGE estimates progressively change across the income distribution following a quite clear U-shaped pattern: parental income influence on children’s income is thus greater at the

---

<sup>1</sup>The “American Dream” refers to opportunity rather than equality. As J. T. Adams said, it is “that dream of a land in which life should be better and richer and fuller for everyone, with opportunity for each according to ability or achievement” (Adams (2012)). In fact, according to the International Social Survey (2012), 94.4% of the Americans think that hard work is essential or very important to get ahead, while this percentage is 75.8% for the average of respondents from all countries. Analogously, 91.4% percent of US respondents think that ambition is essential or very important to get ahead, while this percentage falls to 71% for the world average.

<sup>2</sup>Former studies for the U.S. highlighted IGE values around 0.2 (see Zimmerman (1992) for a review of these studies). Using better databases and correcting for measurement errors, Solon (1992) and Zimmerman (1992) found IGE estimates of about 0.4. Later on, methodological refinements aimed to better correct for transitory shocks and life cycle bias (Mazumder (2005)) estimated values of about 0.5 which are closer to our results.

<sup>3</sup>Previous research using probabilities transition matrices already pointed at a significant inertia for individuals at the tails of the income distribution. Jantti et al. (2006) show that the chances of remaining in the same quintile for individuals with parents from the bottom of the income distribution are significantly higher in the US than in the UK or the nordic countries.

tails of the income distribution. Using family income data from the Panel Study of Income Dynamics (PSID), we apply conditional Quantile Regression (QR) to estimate the IGE in the US in the whole 1980-2010 period.<sup>4</sup> In particular, we combine QR computation with the model proposed in Lee and Solon (2009), to enlarge the available data and, in this manner, obtain accurate estimations at the tails of the distribution while controlling for measurement error and life cycle bias.<sup>5</sup>

In order to study whether the observed high levels of IGE in the US are a recent or a structural phenomenon, and to check whether the trend evolution of the IGE is homogenous across the income distribution of adult sons, our second contribution is a time series analysis of IGE along the 1980-2010 period at different income percentiles. To the best of our knowledge, this is the first time that the trend of IGE is estimated at different points of the income distribution, and we find that the top quantiles and the mid-bottom quantiles of the distribution seem to have followed different trajectories in the three decades considered. However, as explained in the results section, the null hypothesis that IGE values are equal in all years considered for each given quantile is not rejected in most of the quantiles, which calls for caution when interpreting the trend results.

Finally, we explore the role of sons' education and race as intergenerational transmission channels of parental income, both across the income distribution and along the time trend. In this respect, we find that the impact of both factors, education and race, depends on the point of the distribution under consideration.

The rest of the paper is structured as follows. In Section 2 we present our methodology to estimate IGE across the income distribution for the entire pool and year by year. Section 3 details our choices and treatment of the PSID database, while section 4 presents our main IGE results for the pooled sample, and for its trend from 1980 to 2010, and Section 5 concludes. A sensitivity analysis including different data choices and the unconditional quantile regression estimation is available in an online appendix.

---

<sup>4</sup>We also apply the new unconditional quantile regression method proposed by Firpo et al. (2009) in our sensitivity analysis (see Appendix).

<sup>5</sup>See Table 4 for a review of elasticity estimates for the US using OLS in the literature. In Table 5 we review the existing IGE literature using QR in the US. Previous works had not found the clear U-shape pattern connecting IGE with the son's position at the income distribution. As we will discuss in the results section, this could be due to their small samples and to the use of earnings instead of income as the elasticity variable.

## 2 Methodology

The intergenerational income elasticity refers to the influence of parental income in children’s adult income. In the canonical Galton (1886) regression of a child’s income  $y_{s_{it}}$  on the parent’s income  $y_{p_{it}}$ ,

$$\ln y_{s_i} = \alpha + \beta \ln y_{p_i} + \varepsilon_{it} \quad (1)$$

the constant term  $\alpha$  captures the trend in average incomes across generations due for example to changes in labor market institutions, international trade or technology, while the  $\beta$  coefficient, called intergenerational elasticity, measures the degree of persistence in family’s income across generations. The higher the value of  $\beta$ , the larger the capacity of parental income to predict son’s economic achievement. Accordingly,  $1 - \beta$  is a measure of intergenerational income mobility. Finally, the error term  $\varepsilon_{it}$  represents all other influences on the child’s adult income not correlated with parental income.<sup>6</sup>

The use of this basic model presents some important limitations. First, trying to avoid the life cycle bias, scholars have traditionally restricted the sample to observations at a precise children’s age, thus overlooking a lot of information from income at other ages. As a result, the number of observations to estimate intergenerational mobility has typically been small. Second, the intergenerational income elasticity has been usually estimated by ordinary least squares (OLS), which yields an estimate at the mean of the distribution, but ignores the possible variation of intergenerational mobility across income quantiles, as pointed out by Bishop et al. (2014). Finally, when only parental income is included as an explanatory variable, the model in (1) is incapable of analyzing channels of income transmission between parents and children. Next, we explain the main strategies we have adopted to overcome these limitations.

### 2.1 The model

To use all the available information, and still tackle the life cycle bias, we follow the approach in Lee and Solon (2009). This methodology permits the exploitation of the entire pool of data, estimating the IGE with all available pairwise observations of adult sons and parents’ income, while controlling for the influence of the life cycle on income of both parents and children. The equation to be estimated

---

<sup>6</sup>Although the relation between parental income and son’s income cannot be affected by reverse causality, there could be omitted variables that prevent us from establishing a causal relationship. Also, the value of IGE can be influenced by other variables involved (education quantity, education quality, race, location, social connections, etc.) In fact, our study of education and race discloses part of the role of these variables on IGE: controlling for race and for the amount of education reduces our measure of IGE in more than 30% at most quantiles and even more than 60% at the 5th percentile (see Table 2).

is the following:

$$\ln y_{sit} = \alpha + \beta \ln y_{pi} + \sum_{n=1}^4 \gamma_n A_i^n + \sum_{n=1}^4 \delta_n C_{it}^n + \sum_{n=1}^4 \theta_n [\ln y_{pi}] C_{it}^n + \varepsilon_{it} \quad (2)$$

where  $y_{sit}$  is the real household income (in logs) of adult sons from family  $i$  at year  $t = 1980, 1981, \dots, 2010$ ;  $y_{pi}$  is the averaged parental household income (in logs) of family  $i$  when the son was a child between 13 and 19 years old; the rest of terms control for the influence of the life cycle on parental and son's income. Variable  $A_i^n$ , parameters  $\gamma_1$  to  $\gamma_4$ , represents the age of the parent in family  $i$  when the children was 16 years old. Variable  $C_{it}^n$ , parameters  $\delta_1$  to  $\delta_4$ , controls for the son's age when his income is measured. It is expressed as the difference between the age of the son and the age of 40 years old at each year  $t$  in which income is computed, thus centering our estimates at the age of 40. If  $c$  is the birth cohort of the individual,  $t - c$  is the age at which income is reported, and thus  $C = t - c - 40$ . The third variable  $[\ln y_{pi}] C_{it}^n$ , parameters  $\theta_1$  to  $\theta_4$ , represents the interaction between parental income and the age of the son, and it tries to account for the possible divergences in life-income patterns depending on parental income. Age related variables (A and C) are quartic in order to control for different possible functional shapes when time interacts with income.

We first estimate (2) for the entire pool of data, thus obtaining IGE in the US for the entire sample. Later, we estimate the time trend of  $\beta$  between 1980 and 2010 using all available information. For this purpose, we need to modify (2) as follows:

$$\ln y_{sit} = \alpha_t D'_t + \beta_t [\ln y_{pit} D'_t] + \sum_{n=1}^4 \gamma_n A_i^n + \sum_{n=1}^4 \delta_n C_{it}^n + \sum_{n=1}^4 \theta_n [\ln y_{pit}] C_{it}^n + \varepsilon_{it} \quad (3)$$

where  $D_t$  is a vector of yearly dummy variables whose first element takes the value of 1 for 1980 and 0 otherwise, the second element takes the value of 1 for 1981 and 0 for all the rest, and so on. Thus, estimating (3) gives us a different intercept  $\alpha$  and slope  $\beta$  for each PSID wave at  $t = 1980, 1981, \dots, 2010$ . The age-controlling variables are estimated for the pooled data and the model assumes they are time invariant.

## 2.2 Quantile regression

We use Quantile Regression to contrast if intergenerational mobility varies across the income distribution. This method offers the possibility of obtaining point estimates at any selected quantile of the son's conditional income distribution. Using the entire pool of data, we run QR for equation (2) and

estimate IGE at every ventile, i.e. quantiles 0.05, 0.10... 0.95. Initially, the QR estimates are obtained for the pooled 1980-2010 sample. The large size of this sample allows us to obtain highly accurate QR estimates at the tails. Later, we estimate the QR version of (3) and characterize the time trend evolution of IGE at different percentiles for the 1980 – 2010 period. Despite that these estimations are less accurate because the sample must be split, they permit to analyze the particular trend of IGE at different quantiles all along the 1980 – 2010 period.

In contrast with OLS, which minimizes squared errors and yields the estimates at the mean of the distribution, QR minimizes absolute errors at any particular quantile of the conditional  $Y|X$  distribution (Koenker and Bassett (1978), Koenker (2005)).<sup>7</sup> Suppose that we want to calculate the QR estimate of the quantile  $\tau$ . Then, those absolute errors corresponding to observations below the quantile  $\tau$  are weighted with the weight  $1 - \tau$ , while the absolute errors for those observations above the quantile  $\tau$  are weighted (asymmetrically) with  $\tau$ . This asymmetrical weighting can make the QR estimates less robust at the tails of the distribution. This is not a problem for samples that are sufficiently large, but with small samples, a change in only some of the data might alter the coefficient quite significantly. For this reason, we apply the proposal in Lee and Solon (2009), which allows us to use the entire 1980-2010 pool of data to estimate IGE at all ventiles of the distribution. In our yearly IGE trend estimates, when the estimation is 'split' by years, we have excluded the most extreme quantiles ( $\tau = 0.05$  and  $\tau = 0.95$ ) from the graphical representation of the results due to the high standard errors of the estimation at those quantiles.

### 2.2.1 Conditional and unconditional quantile regression

Unlike the OLS estimator, which is valid for both the conditional and unconditional distribution, conditional QR estimates cannot be used to represent the estimates at the unconditional quantiles of the distribution. Trying to address this problem, Firpo et al. (2009) have proposed a new 'unconditional quantile regression' method to estimate the impact of the different X covariates at the unconditional quantiles. In order to check the robustness of our results, in the online appendix we have compared conditional and unconditional quantile estimations of IGE for an age restricted-subsample. We find the results in the conditional and unconditional regressions to be overall quite similar and that our main qualitative findings (i.e. the U-shape pattern of the IGE and the different role of education and

---

<sup>7</sup>The use of absolute errors instead of squared errors makes QR less sensitive to outliers than OLS. Also, as pointed out by Mitnik et al. (2015) OLS estimates of elasticity using log transformed income are in fact centered at the geometric mean instead of the arithmetic mean; however, in contrast to the mean, the median and the quantiles estimated by QR are unaffected by a log transformation of the variables.

race across the distribution) still hold.

## 2.3 Factors of intergenerational income transmission

It is a challenging issue to understand the main channels and factors that condition the transmission of income from parents to children. In principle, education, connections, race and other genetic traits are potential candidates. Unfortunately, the availability of data to test some of these factors is limited.<sup>8</sup>

We focus on two possibly explaining variables that are time-consistent along the PSID panel: son's 'years of education' and 'race'. We attempt to measure the importance of these factors in the transmission of income across the children's income distribution, first for the entire pool of data, and then at each PSID wave along the last three decades.

To estimate the impact of education for the entire pool, we first add in equation 2 the 'years of education' variable  $e_{s_i}$ .

$$\ln y_{sit} = \alpha + \beta \ln y_{pi} + \lambda e_{s_i} + \sum_{n=1}^4 \gamma_n A_i^n + \sum_{n=1}^4 \delta_n C_{it}^n + \sum_{n=1}^4 \theta_n [\ln y_{pi}] C_{it}^n + \varepsilon_{it} \quad (4)$$

where  $\lambda_{s_i}$  is the partial direct impact of the variable  $e_{s_i}$  on son's income, given parental income and all other controls in 4. How can we interpret a possible change in the  $\beta$  coefficient after the inclusion of the variable  $e$ ? Let us consider an extreme situation in which the education variable  $e$  is uncorrelated with parental income. In this case, even when the variable  $e$  is significant to explain children income, including this variable in the regression does not modify the influence that parental income has on son's income, thus the primitive  $\beta$  (as estimated in (2)) will remain unchanged. On the opposite case, if the variable  $e$  is strongly correlated with parental income, the new  $\beta$  will significantly drop when the variable  $e$  is included in the regression. Hence, we can interpret that the smaller the change in  $\beta$  when we include variable  $e$  in the regression, the weaker the role of this variable as a transmission channel. Analogously, comparing the elasticity ( $\beta$ ) from equation (4) with the one obtained in equation (2) can measure the share of elasticity 'mediated' by education:  $(\beta_{baseline} - \beta_{edu})/\beta_{baseline}$

To control for the additional effect of race, we have added the race variable in equation 5. Variable  $r_{s_i}$  is a dummy variable that takes the value 1 for white individuals and is 0 otherwise. Then, the impact

---

<sup>8</sup>Anger and Heineck (2010) find a positive relation between parental and children cognitive abilities, even controlling for education and economic background. It is hard, however, to connect this transmission of abilities with the transmission of income, and studies about this transmission channel are rare. Data availability has made scholars focus mainly on variables like education and race (Hertz (2006); Torche (2013)). Bowles and Gintis (2002) is a prominent exception, finding the impact of intelligence on income transmission to be relatively small, accounting for a 12.5% share of the intergenerational correlation.

of race can be calculated comparing the elasticity  $\beta$  from equation (5) and the  $\beta$  from (4), and relating it with the original baseline beta:  $(\beta_{edu} - \beta_{race+edu})/\beta_{baseline}$

$$\ln y_{sit} = \alpha + \beta \ln y_{pi} + \lambda e_{si} + \varphi r_{si} + \sum_{n=1}^4 \gamma_n A_i^n + \sum_{n=1}^4 \delta_n C_{it}^n + \sum_{n=1}^4 \theta_n [\ln y_{pi}] C_{it}^n + \varepsilon_{it} \quad (5)$$

Finally, we have analogously included the variables  $e_{si}$  and  $r_{si}$  in our trend estimation (3), in order to analyze the influence of son's education and race in the time evolution of IGE.

### 3 Database

To measure intergenerational income mobility, we use the 'core' sample of the PSID database. The PSID is a household panel maintained by the University of Michigan that began in 1968 and is still running. The survey was conducted annually from 1968 to 1997, and then every other year.

The income variable used is total family income, which aggregates the total income of the household, including taxable incomes and transfers received by the head, the head's spouse and other family members, and which is consistently included in the PSID since its creation (note that income reported refers to the year prior to the interview). All values are transformed to 2010 US dollars using the average Consumer Price Index (CPI) from the Bureau of Labor Statistics and outlier observations are removed. We follow Lee and Solon (2009) and exclude observations for which income is less than \$100 or more than \$150,000 in 1967 dollars. In total, around 200 observations (less than 1% of the sample) were dropped.<sup>9</sup>

We match sons and parents using the individual and family codes provided by the PSID, creating an unbalanced panel. Parental observations include family incomes of households with both male and female heads, and the sample of children is restricted to those sons that later become household heads.<sup>10</sup>

We have averaged yearly parental family income when the child was between 13 and 19 years old (seven years), provided there were at least three observations over this period (6.24 observations on average), in order to reduce transitory shocks (see Solon (1992), Zimmerman (1992)), Mazumder (2005)

---

<sup>9</sup>Note that we also exclude outlier observations in which reported parental age -when the child was 16- is smaller than 30 or greater than 70.

<sup>10</sup>Our preliminary results showed that adult daughters' IGE depended strongly on their marital status. A rigorous analysis for women should consider assortative mating (Chadwick and Solon (2002), Black and Devereux (2011)) and the structural change in women's access to the labor market occurred in the decades analyzed, which is beyond the scope of this paper. In this respect, note also that race data for wives is only available from 1984.



or Mitnik et al. (2015)).

In the online appendix we discuss the sensitivity of these data choices to different specifications. Our choice of the PSID 'core sample' seems robust to the use of only the Survey Research Component sample of the PSID. Also, we test our results to the use of different outliers thresholds and also include an illustrative example of the effect of different numbers of parental years averaged, which supports our choice of up to seven years averaged as a good measure of permanent income.

When observing the income of the adult children, a 'life cycle' bias can arise in the estimation depending on the age at which this income is observed. Previous works on intergenerational elasticity have concluded that observing income at the middle of the life cycle is the best proxy of permanent income (Black and Devereux (2011)).<sup>11</sup> However, restricting the sample to observations at a precise children's age, implies ignoring a lot of information from income at other ages that might be available and could be exploited. To use this information, but still tackle the life cycle bias, we follow the approach in Lee and Solon (2009). As mentioned in Section 2, instead of shortening the age range of children, we use all available observations of income from the whole working life of individuals, but include age-dependent covariates in the regression to control for the different age at which family income is observed, centering our estimates at the age of 40. For consistency, we control in the regressions also for parental age in order to tackle the potential parental life cycle bias. Since the age controls are estimated for the period 1980-2010 as a whole, it is important that there is no significant change in the age-income pattern across cohorts. We have checked that in our sensitivity analysis (see online appendix) and found no significant change.

In sum, at each year from 1980 to 2010, we keep the observations of sons who are between 25 and 65 years old, provided that they are the head of the household and live in the family home. By the year 1980 we already have sufficient individuals who were between 13 and 19 years old in 1968 (when the PSID began) and have already established their own household. In table 1 we show the number of observations that abide all these criteria for all years in the period 1980-2010, and include the mean and standard deviation of age and real family income in logs for parents and sons. Our sample consists of a total of 25,084 observations from 3088 different individuals. On average, each individual appears in 8.12 waves of the survey, with a standard deviation of 6.39 times. As discussed in section 4.1, we have taken this into account in the computation of standard errors (see Footnote 12).

In addition to the main total family income variable, we also consider from the PSID the individual variables 'years of education' and 'race', aiming to study their importance in the transmission of

---

<sup>11</sup>Using Finnish data, Lucas and Kerr (2013) find that IGE estimates increase with the son's age considered until approximately the age of 40.

parental income. The education variable represents the actual grade of school completed, ranging 1-17 where a code value of 17 indicates that the individual completed at least some postgraduate work. In the case of race, we transform the categorical variable 'race of head' into a dummy variable that takes the value 1 when the race of the son is white and zero otherwise. Using a dummy for white race, we implicitly assimilate black to the other non-white races. Please see the online appendix for a discussion on the use of educational categories instead of years and of a black race dummy instead of a white race dummy. Overall, results do not change significantly.

## 4 Intergenerational Income Elasticity results

In the first part of this section we present the results of our pooled data regression. In particular, using the entire 1980 - 2010 sample, we show the value of IGE at each quantile. We also measure the importance of education and race as channels of intergenerational income transmission. In the second part, we study the evolution of IGE between 1980 and 2010 at different points of the distribution of income and the role of education and race along that period and across the distribution.

### 4.1 IGE by quantiles: a pooled regression analysis for the 1980-2010 period

The  $\beta$  intergenerational income elasticity estimates obtained from the pooled (1980-2010) sample at the mean and at all conditional ventiles are displayed in Table 2 and Figure 1.<sup>12</sup> The OLS estimation yields a value of 0.47, which is in line with the literature (see Table 4). More importantly, if we enrich the picture with the conditional QR estimations, we observe a U-shaped relationship. The intergenerational income elasticity is highest at the lower percentiles of the distribution –reaching a value of around 0.6 at the 5<sup>th</sup> – 20<sup>th</sup> percentiles. Then, it declines steadily, reaching a minimum around 0.38 at the 70<sup>th</sup> percentile. At the top part of the distribution, the IGE increases again, reaching a value of almost 0.5 at the 90<sup>th</sup> – 95<sup>th</sup> percentiles.<sup>13</sup>

---

<sup>12</sup> Given that individuals appear in our sample during several survey waves, our observations can be considered to be 'clustered' in individuals, and standard errors must take this into account. For that purpose, we have applied the clustered version of the bootstrap method in the 'quantreg' R package, which is based on the proposal of Hagemann (2016). When possible all figures plot a standard error bar centered at the point estimate. We thank an anonymous referee for pointing this out to us.

<sup>13</sup> Our target variable is total family income, which is computed after transfers but before taxes, and is not directly affected by differential taxation overtime (we ignore here possible behavioral effects). Although (Mitnik et al., 2015, p. 71) do not find a significant difference between using pre-tax and post-tax income in the measurement of the IGE, our total family income might be affected by different transfer policies overtime. Transfers could be specially relevant for the lower part of the distribution and could downward bias the IGE estimates at the lowest quantiles.

These results indicate that the 'inheritance' of family income in the US varies when we move along the conditional income distribution of adult sons. For example, a hypothetical shift in one dollar of parental income would shift average son's income in 0.47 dollars (our OLS estimate), while the 10<sup>th</sup> quantile of the conditional income distribution would shift by 0.64 dollars and the 70<sup>th</sup> quantile by just 0.38. Children at the upper middle part of the conditional distribution show the smallest degree of intergenerational persistence, while top incomes and, specially, low incomes are very much conditioned by their childhood economic circumstances, represented here by parental income. Previous studies estimating the IGE at different quantiles have relied on much smaller samples and have found disparate results. For example, Grawe (2004), using a sample of only 354 observations, found that intergenerational elasticity is higher at the median than at the tails, i.e., an inverse U-shaped. Eide and Showalter (1999) using a sample of 612 observations, and Cooper (2011) with a sample of 1,424 observations found a continuous –almost linear– decrease in the IGE as we go up the income distribution. According to these authors there is not a significant increase in the IGE at the upper part of the distribution.<sup>14</sup>

Besides the much bigger sample used in our research, there exists another reason that could explain why these previous studies do not find an increase of the IGE in the US from the 70<sup>th</sup> percentile onwards. While we use parents and sons' total household income, Eide and Showalter (1999) regress son's *earnings* on parental earnings/income, and Cooper (2011) measures intergenerational elasticity for sons' and parents labor earnings. A great deal of the correlation between parental and children incomes at the upper part of the distribution could occur through capital income, which is included in the total household income variable. If so, values of intergenerational elasticity of sons' earnings would underestimate actual intergenerational elasticity of income at the top quantiles. In this sense, Jantti et al. (2006), using transition matrices to measure intergenerational mobility also of earnings, find higher inertia at both ends of the distribution in the U.S., but with more intensity at the bottom than at the top.<sup>15</sup>

Studies measuring intergenerational elasticity applying QR in other countries are scarce, but seem to

---

<sup>14</sup>Recall that our sample consists of 25,084 observations from 3,088 individuals. See Table 5 for a summary review of the results of the literature using QR for IGE estimation in the US.

<sup>15</sup>Bowles and Gintis (2002) find that wealth explains 0.12 out of a 0.32 correlation between parental and children income, more than a third of the value. Wealth –and therefore the capital income derived from it– is concentrated at the top percentiles of the distribution. Levine (2012) reports that in 2010 the top 1% of the households ordered by wealth had a share of 34.5% of the net worth in the U.S. while the bottom 50% possessed only 1%. Fräðdorf et al. (2011) show that the share of inequality in household income explained by capital income is increasing in the U.S. Outside the US, Lucas and Kerr (2013) have also found –using a nested model– that intergenerational transmission of income is significantly greater than intergenerational transmission of earnings in Finland.

concide in finding less mobility at the bottom of the income distribution. In line with our results, Tejada et al. (2015), in their estimation of intergenerational elasticity of income for the 1982 born cohort in the city of Pelotas (Brasil), find higher values of the IGE at both ends of the income distribution. On the other hand, Bratberg et al. (2007) apply QR for earnings data from Norway cohorts born in 1950-1960, and find the relation between the IGE and the position at the earnings distribution to be decreasing, with higher IGE at the bottom tail, but more mobility (lower IGE) at the top of the earnings distribution. Again, the distinction between income and earnings discussed above could explain that at the top of the *earnings* distribution mobility is higher than at the middle, while the opposite happens when we consider *income*.

#### 4.1.1 Education and race impact on elasticity in the pool

Next, we focus on the role of education and race as possible channels of income transmission across generations. Our results –see Table 2– show that when education is included in the regression (equation (3)), the estimated the IGE decreases a share of 0.274 (27.4%) at the mean (OLS estimation). This OLS result is similar to Eide and Showalter (1999) or Cooper (2011) who find approximately a 30% mediating role of education in the persistence of income across generations at the mean in the distribution; other works (Torche (2013)); Blanden et al. (2014)) find an even higher explaining role of education.<sup>16</sup> Our QR results find a share of the IGE mediated by education between 20% and 48% depending on the quantile. This share is lower in the range of the 20th-70th percentiles –representing around 20% of the IGE- and increases significantly when approaching the extremes of the distribution (see Table 2 and Figure 2a). Thus, even though we cannot control for factors like the quality of the schools or the peer-effects, between one fifth and half of intergenerational income transmission is explained just by the different amount of education –measured in years– that parents can provide to their children.<sup>17</sup>

With respect to race, the OLS regression yields an additional decrease in the IGE of 10% when we include the dummy variable ‘race’ as an additional control in equation (5). Thus, at the mean, one tenth of the ‘inheritance’ of parental income can be attributed to the race of the individual (Table 2), controlling by the years of education. Looking at the impact of race on the IGE across the income distribution -which, to the best of our knowledge, has never before been studied in the literature- we find it to be of around 10% at the bottom half percentiles, the influence being much smaller (about

---

<sup>16</sup>See table 4 for a review of the most relevant previous literature on this issue.

<sup>17</sup>Needless to say, the years of education mediating role could englobe other factors cross-correlated with the number of years of schooling, parental income and son’s income (e.g. parental motivation). Also note that, for robustness, we have also run the analysis using educational levels instead of years of education, finding a similar impact on the IGE (see Figures 2a and A5a).

5%) from the 60<sup>th</sup> percentile upwards (Table 2 and Figure 2b).

The role of education shows no strong trend pattern across the income distribution in the period analyzed -see online appendix- although it seems to increase slightly in the 00s for all quantiles. The importance of race in the IGE shows a similar pattern at all quantiles analyzed: decreasing in the 80s and -especially- in the 90s, but regaining importance from the mid 2000s.

## 4.2 Evolution of the IGE in the US between 1980 and 2010

As seen above, for the period 1980-2010 as a whole, high-income quantiles and, above all, low-income quantiles show greater elasticity than middle-income quantiles. But, how was the evolution of the IGE for the entire distribution and by quantiles during this period? For illustrative purposes, we present the results graphically by groups of quantiles and excluding the two extreme ventiles: the low-income group (10<sup>th</sup> to 30<sup>th</sup> percentiles); the mid-low income group (percentiles 35<sup>th</sup> to 50<sup>th</sup>); the mid-high income group (percentiles 55<sup>th</sup> to 70<sup>th</sup>); and the high-income group (percentiles 75th to 90th). The estimation at the mean (OLS) is plotted with the mid-low income group that includes the median (Figure 3).<sup>18</sup>

We must caution the reader about the statistical significance of the trends of IGE at the different quantiles. Using the Wald test, we have tested the hypothesis that all IGE estimates from different years at a given quantile are equal (see Table 3). The hypothesis can be rejected at the 5% significance level only for the 20th, 35th and 55th percentiles and for the linear projection estimated with OLS. Also note that the test only statistically rejects that the coefficients are equal, but does not evaluate the slope or the direction of the possible trend.

Our OLS estimation at the mean (see Figure 3, top-right), the intergenerational elasticity shows a decreasing trend in the first two decades analyzed, followed by an increase in the 2000s. This result contrasts with Aaronson and Mazumder (2008), who found an increase in the IGE over the 1980-2000 period, and with Hertz (2007) and Lee and Solon (2009) who found no trend for that same period. Mayer and Lopoo (2005), on the other hand, found a decreasing trend of the IGE for the period 1984-94.<sup>19</sup> With a bigger sample using tax records and the Statistics of Income annual cross sections, Chetty et al. (2014b) estimate rank-rank relative intergenerational mobility for cohorts born after 1971,

---

<sup>18</sup>For space reasons, the tables with the estimations of the IGE at each ventile for each PSID wave have not been included. Neither have the tables with the trend estimates for the IGE controlling for education and race. They are available upon request.

<sup>19</sup>Mayer and Lopoo analyze trends by cohorts. The period 1984-94 corresponds to the years in which the cohorts are 30 years old, the age at which they estimate the IGE in their rolling groups regression (Mayer and Lopoo, 2005, p. 176)

measuring son's family income when the son is 29-30 years old. For cohorts born in the 70s –which would correspond to our estimates in the 2000s decade- they find a stable trend in relative mobility, which would be consistent with our OLS estimation of an increasing trend in IGE for that period, given that inequality in the United States increased during that decade.<sup>20</sup>

Concerning the trend at different points of the income distribution, our estimates show that, for the linear projection and for all the quantiles below the median, the IGE decreased in the 80s and 90s and increased in the 2000s. This pattern is more pronounced at the lowest quantiles. At the upper part of the distribution, however, both the mid-high and the high-income quantiles maintained a steady value of the IGE along the three decades analyzed, and show only a very mild decreasing pattern in the 90s that turns increasing in the 2000s. It is worth noting that the IGE at the low-income quantiles has always been the highest, this group consistently suffering from lower mobility than the rest of income groups.

With more intensity for the lower part of the distribution the change of century seems to be a turning point in the trend of the IGE for all groups. Elasticity raised in all income groups since 2002, above all with the Great Recession (2007-2009). After the Great Recession intergenerational elasticity generally decreased in 2010, although more observations will be required to confirm this new trend in the IGE series.

We have also run an analysis of the share of the IGE related to education and race along the 1980-2010 period. The results, which again must be interpreted with caution due to the data limitations when we estimate IGE for each wave separately, are displayed in the online appendix (Figures A8 and A9). They show no clear trend for education at the different quantiles and, for race, they point at an increase in importance in the 2000s after having decreased in the previous two decades.

## 5 Conclusion

Despite the extensive literature on the subject of measuring the magnitude of the IGE in the US, most of the works estimate it at the mean of the income distribution. The few studies that estimate the IGE at different quantiles in the US work with small samples, since they consider only a cross-section of individuals at a small age range. As a result, estimates at the tails are prone to being biased and they have arrived at disparate results. In an attempt to overcome these limitations, we use up-to-

---

<sup>20</sup>Note that for a certain level of correlation between parent and son's income, IGE regression estimates increase when the inequality ratio between the sons and parents distributions increases. The Gini Index at disposable income in the US rose from 0.357 in 2000 to 0.380 in 2010 as reported by OECD.

date family income data from the PSID to exploit a greater number of data while still controlling for measurement errors and life cycle bias. We apply quantile regression to the estimation of IGE in the US for the 1980-2010 period and explore the role of the child's amount of education and race as potential conditioning factors in the intergenerational transmission of parental family income. To check the robustness of our results, we carry out a large sensitivity analysis that includes the RIF-OLS unconditional quantile regression.

Our main finding reveals that economic persistence is higher at the tails of the distribution. While our OLS estimate of IGE for the entire pool is 0.47, in line with the literature, using QR we find that 'inheritance' of income varies significantly across the child's adult income distribution. Moreover, the IGE shows a U-shaped relationship with the son's income rank, with maximum values at the tails of the distribution (0.64 at the 10<sup>th</sup> percentile and 0.48 at the 95<sup>th</sup> percentile) and a minimum value -maximum mobility- of 0.37 at the 70<sup>th</sup> percentile. Children at the top and, more importantly, at the bottom of the distribution have been more conditioned by their parental income than those belonging to the 'middle class'.

We believe that these findings may contribute to better target public policies aiming to promote economic mobility. Moreover, they point to education as a relevant factor that influences economic persistence, especially at both tails of the distribution, and to the additional impact of race in mobility at the mid and lower parts of the distribution. For our pooled data, we find that son's education represents between 20% and 50% of the IGE, being particularly important at the tails of the distribution, where a greater share of the intergenerational economic persistence is driven through the different amount of education provided to children. Meanwhile, factors related to race can explain more than 10% of the transmission of parental income, their importance being highest below the 60th percentile of the income distribution.

Finally, and although for most quantiles the hypothesis that all year estimates are equal cannot be statistically rejected at the 5% level - which calls for caution when interpreting these results- there seem to be also different patterns for the IGE evolution at different parts of the distribution. We find that, for all percentiles up to the median (and for the OLS estimate), the trend of IGE decreased in the 80s and 90s and increased slightly in the 00s, while for higher-income percentiles the IGE remained relatively stable all along.

## References

- Aaronson, D. and Mazumder, B. (2008). Intergenerational economic mobility in the United States, 1940 to 2000. *Journal of Human Resources*, 43(1):139–172.
- Adams, J. T. (2012). *The epic of America*. Transaction Publishers.
- Anger, S. and Heineck, G. (2010). Do smart parents raise smart children? The intergenerational transmission of cognitive abilities. *Journal of Population Economics*, 23(3):1105–1132.
- Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist’s companion*. Princeton University Press.
- Bishop, J. A., Liu, H., and Rodríguez, J. G. (2014). Cross-country intergenerational status mobility: Is there a great gatsby curve? *Research on Economic Inequality*, 22:237–249.
- Björklund, A. and Jäntti, M. (2009). Intergenerational income mobility and the role of family background. *Oxford Handbook of Economic Inequality*, Oxford University Press, Oxford.
- Black, S. E. and Devereux, P. J. (2011). Recent developments in intergenerational mobility. *Handbook of Labor Economics*, 4:1487–1541.
- Blanden, J. (2013). Cross-country rankings in intergenerational mobility: a comparison of approaches from economics and sociology. *Journal of Economic Surveys*, 27(1):38–73.
- Blanden, J., Haveman, R., Smeeding, T., and Wilson, K. (2014). Intergenerational mobility in the United States and Great Britain: a comparative study of parent–child pathways. *Review of Income and Wealth*, 60(3):425–449.
- Bowles, S. and Gintis, H. (2002). The inheritance of inequality. *The Journal of Economic Perspectives*, 16(3):3–30.
- Bratberg, E., Nilsen, Ø. A., and Vaage, K. (2007). Trends in intergenerational mobility across off-spring’s earnings distribution in norway. *Industrial Relations: A Journal of Economy and Society*, 46(1):112–129.
- Chadwick, L. and Solon, G. (2002). Intergenerational income mobility among daughters. *The American Economic Review*, 92(1):335–344.



- Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014a). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *National Bureau of Economic Research*.
- Chetty, R., Hendren, N., Kline, P., Saez, E., and Turner, N. (2014b). Is the united states still a land of opportunity? recent trends in intergenerational mobility. *The American Economic Review*, 104(5):141–147.
- Cooper, D. P. (2011). Unlocking the American dream: exploring intergenerational social mobility and the persistence of economic status in the United States.
- Corak, M. (2006). Do poor children become poor adults? Lessons from a cross country comparison of generational earnings mobility. *IZA Discussion Paper*.
- Eide, E. R. and Showalter, M. H. (1999). Factors affecting the transmission of earnings across generations: A quantile regression approach. *Journal of Human Resources*, pages 253–267.
- Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, pages 953–973.
- Fräbendorf, A., Grabka, M. M., and Schwarze, J. (2011). The impact of household capital income on income inequality—a factor decomposition analysis for the UK, Germany and the USA. *The Journal of Economic Inequality*, 9(1):35–56.
- Galton, F. (1886). Regression towards mediocrity in hereditary stature. *Journal of the Anthropological Institute of Great Britain and Ireland*, pages 246–263.
- Gouskova, E., Andreski, P., and Schoeni, R. F. (2010). *Comparing estimates of family income in the Panel Study of Income Dynamics and the March Current Population Survey, 1968-2007*. Survey Research Center, Institute for Social Research, University of Michigan.
- Gouskova, E., Heeringa, S., McGonagle, K., Schoeni, R., and Stafford, F. (2008). Panel Study of Income Dynamics revised longitudinal weights 1993–2005. *Panel Study of Income Dynamics Technical Paper Series, University of Michigan, Retrieved December, 28*.
- Grawe, N. D. (2004). Intergenerational mobility for whom? The experience of high-and low-earning sons in international perspective. *Generational income mobility in North America and Europe*, pages 58–89.

- Hagemann, A. (2016). Cluster-robust bootstrap inference in quantile regression models. *Journal of the American Statistical Association*, forthcoming.
- Hertz, T. (2006). Understanding mobility in America. *Center for American Progress Discussion Paper*.
- Hertz, T. (2007). Trends in the intergenerational elasticity of family income in the United States. *Industrial Relations: A Journal of Economy and Society*, 46(1):22–50.
- Jantti, M., Bratsberg, B., Roed, K., Raaum, O., Naylor, R., Osterbacka, E., Bjorklund, A., and Eriksson, T. (2006). American exceptionalism in a new light: A comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States. *IZA Discussion Paper No. 1938*. Available at SSRN: <http://ssrn.com/abstract=878675>.
- Koenker, R. (2005). *Quantile regression*. Number 38. Cambridge University Press.
- Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, pages 33–50.
- Lee, C.-I. and Solon, G. (2009). Trends in intergenerational income mobility. *The Review of Economics and Statistics*, 91(4):766–772.
- Levine, L. (2012). An analysis of the distribution of wealth across households, 1989-2010. *Congressional Research Service*.
- Lucas, R. E. and Kerr, S. P. (2013). Intergenerational income immobility in finland: contrasting roles for parental earnings and family income. *Journal of Population Economics*, 26(3):1057–1094.
- Mayer, S. E. and Lopoo, L. M. (2005). Has the intergenerational transmission of economic status changed? *Journal of Human Resources*, 40(1):169–185.
- Mazumder, B. (2005). Fortunate sons: New estimates of intergenerational mobility in the united states using social security earnings data. *Review of Economics and Statistics*, 87(2):235–255.
- Mitnik, P., Bryant, V., Weber, M., and Grusky, D. (2015). New estimates of intergenerational mobility using administrative data. Technical report, SOI Working Paper, Statistics of Income Division, Internal Revenue Service.
- Nybom, M. and Stuhler, J. (2016). Heterogeneous income profiles and lifecycle bias in intergenerational mobility estimation. *Journal of Human Resources*, 51(1):239–268.
- Schoeni, R. F. and Wiemers, E. E. (2015). The implications of selective attrition for estimates of intergenerational elasticity of family income. *The Journal of Economic Inequality*, 13(3):351–372.

- Solon, G. (1992). Intergenerational income mobility in the United States. *The American Economic Review*, pages 393–408.
- SSP Research Group (2012). International Social Survey Programme: Social Inequality IV - ISSP 2009.
- Sullivan, L., Meschede, T., Dietrich, L., Shapiro, T., Traub, A., Ruetschlin, C., and Draut, T. (2015). The racial wealth gap: Why policy matters. *Institute for Assets and Social Policy, Brandeis University and Demos*.
- Tejada, C. A. O., Bertoldi, A. D., Carraro, A., Ribeiro, F. G., Motta, J. V. d. S., Barros, F. C., Horta, B. L., and Barros, A. J. (2015). Poor dad, poor child? an investigation of intergenerational income mobility in the 1982 birth cohort in pelotas, rio grande do sul state, brazil. *Cadernos de Saúde Pública*, 31(6):1225–1233.
- Torche, F. (2013). Education and the intergenerational transmission of advantage in the US. *WP available at SSRN 2335007*.
- Zimmerman, D. J. (1992). Regression toward mediocrity in economic stature. *The American Economic Review*, pages 409–429.

## 6 Figures and tables

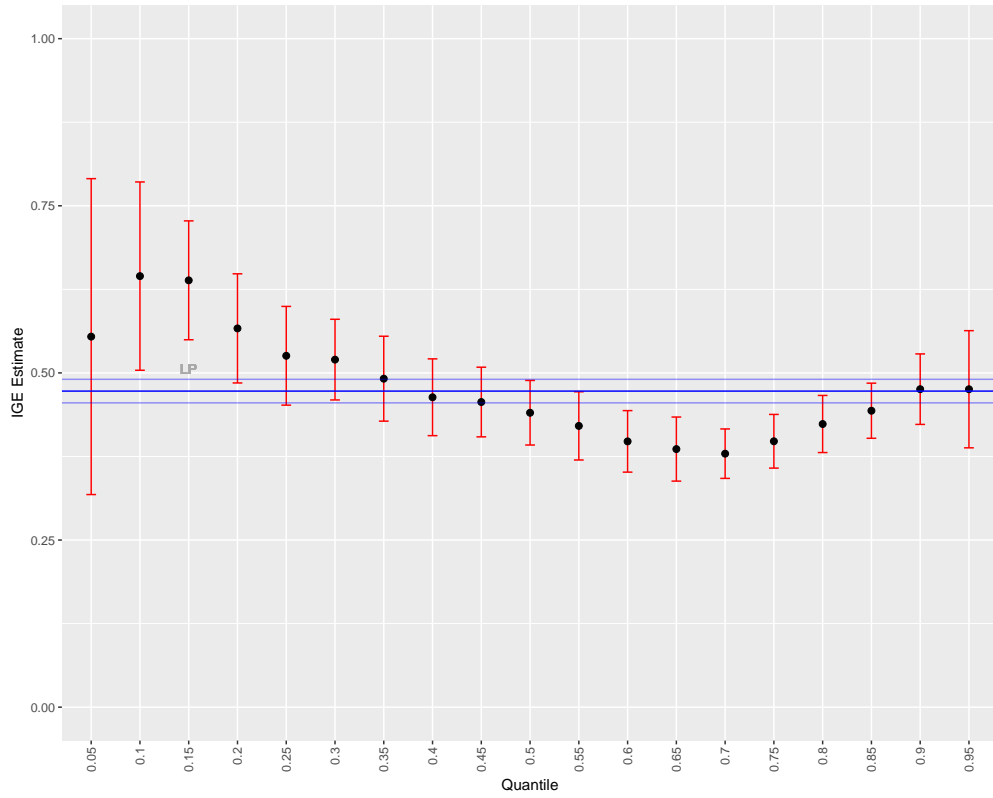
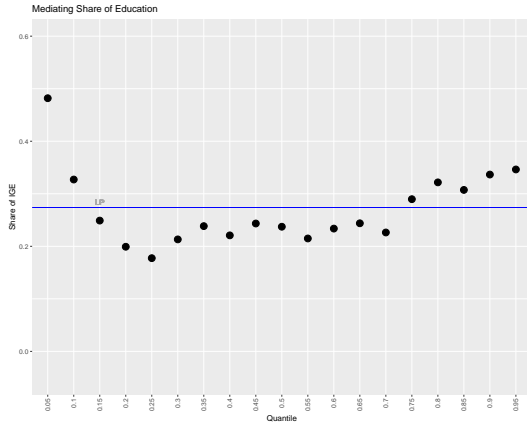
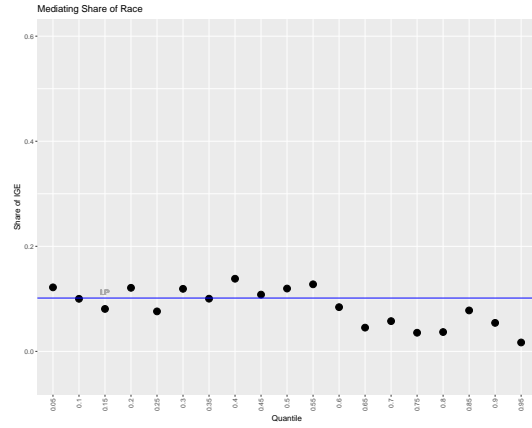


Figure 1: IGE pool estimates for the baseline model (25084 observations). Confidence intervals plotted in red for conditional quantile estimates obtained with QR estimation and in blue for the linear projection estimates (LP) obtained with OLS estimation.



(a) Education impact ('Years of education')



(b) Race impact ('White race' dummy)

Figure 2: Impact of Education and Race as a share of IGE. Pooled estimates.

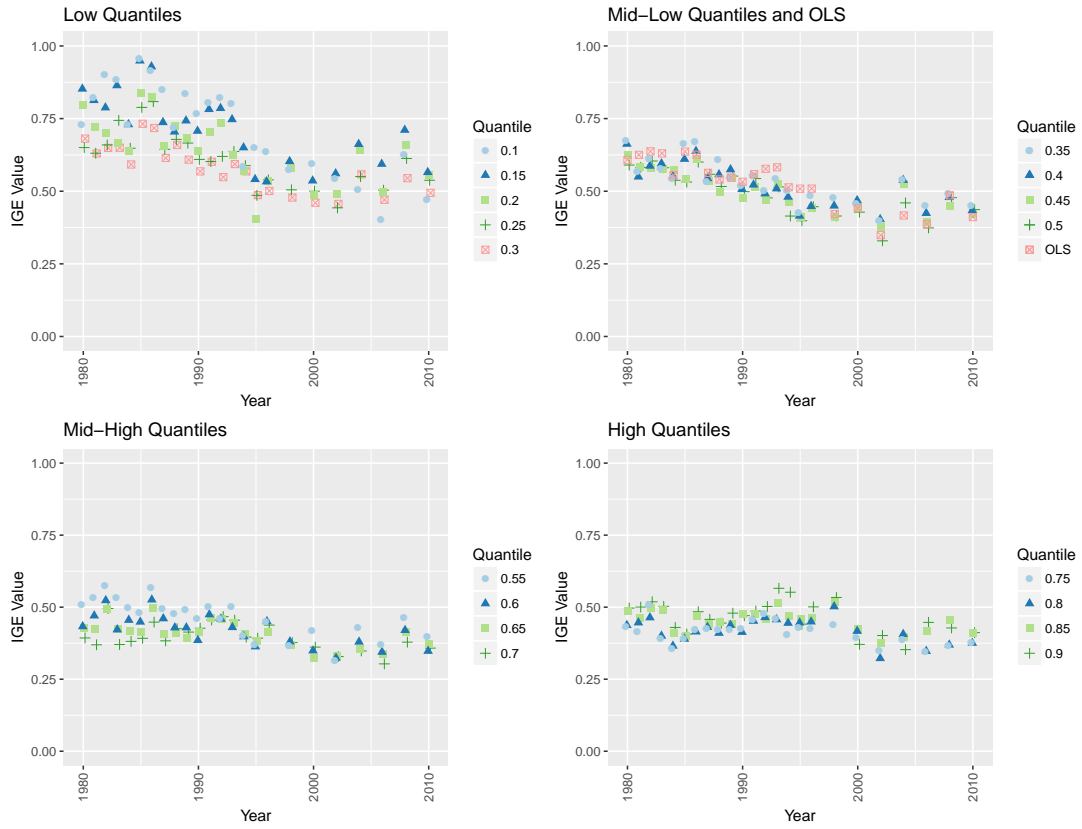


Figure 3: IGE Trend 1980-2010

Year	Obs.	Son's Age		Son's Income		Dad's Age		Dads's Income		Race Share (%)		
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Whites	Blacks	Other
1980	457	26.79	1.43	10.72	0.73	44.92	6.34	10.94	0.63	61.49	36.98	1.53
1981	551	27.32	1.67	10.66	0.84	44.57	5.99	10.97	0.63	62.25	35.93	1.81
1982	629	27.81	1.93	10.61	0.86	44.59	6.17	10.97	0.63	61.84	36.09	2.07
1983	680	28.43	2.20	10.67	0.86	44.84	6.15	11.00	0.63	62.50	35.15	2.35
1984	776	28.90	2.50	10.70	0.85	44.76	6.11	11.01	0.64	63.14	34.28	2.58
1985	850	29.28	2.80	10.70	0.87	44.80	6.10	11.02	0.64	65.41	33.76	0.35
1986	921	29.73	3.09	10.75	0.83	44.94	6.22	11.03	0.63	65.58	33.66	0.33
1987	993	30.22	3.33	10.76	0.87	44.95	6.17	11.05	0.63	65.86	33.33	0.40
1988	1043	30.82	3.60	10.81	0.83	44.95	6.17	11.05	0.63	65.77	33.37	0.19
1989	1095	31.37	3.84	10.81	0.89	44.83	6.17	11.06	0.64	67.03	32.05	0.27
1990	1158	31.81	4.10	10.80	0.86	44.66	5.98	11.05	0.65	66.67	32.12	0.52
1991	1228	32.37	4.34	10.79	0.88	44.70	6.05	11.04	0.66	65.07	30.62	0.81
1992	1247	33.00	4.68	10.84	0.94	44.60	6.06	11.07	0.65	65.76	30.07	0.64
1993	1326	33.54	4.87	10.82	0.96	44.41	6.00	11.07	0.66	63.20	27.98	0.90
1994	1317	33.99	5.06	10.89	0.89	44.31	5.97	11.09	0.65	67.81	29.46	1.75
1995	1339	34.42	5.40	10.90	0.90	44.28	6.02	11.09	0.65	68.33	28.83	1.87
1996	950	35.11	5.81	11.01	0.83	44.08	5.75	11.20	0.63	76.74	20.84	1.16
1998	1014	35.95	6.59	11.13	0.82	43.77	5.74	11.21	0.62	75.84	21.20	2.76
2000	1077	36.68	7.10	11.13	0.86	43.60	5.80	11.19	0.65	76.04	21.08	2.60
2002	1136	37.40	7.74	11.13	0.80	43.33	5.69	11.20	0.67	75.88	21.74	2.20
2004	1215	37.62	8.40	11.10	0.88	43.18	5.62	11.20	0.68	75.56	23.05	1.32
2006	1293	37.76	9.02	11.06	0.92	43.02	5.36	11.21	0.68	73.86	24.36	1.39
2008	1381	38.09	9.47	11.04	0.92	42.99	5.24	11.21	0.70	73.35	24.76	1.59
2010	1408	38.24	9.91	10.94	0.96	42.95	5.30	11.22	0.73	72.30	25.36	2.06

Table 1: Descriptive Statistics

Quantile	Baseline		Model w/ Education		Model w/ Edu + Race		Impact on IGE (Share of Baseline)		
	IGE	(SE)	IGE	(SE)	IGE	(SE)	Edu	Race	Edu + Race
OLS	0.473	0.018	0.343	0.017	0.295	0.017	0.274	0.101	0.375
0.05	0.554	0.223	0.287	0.286	0.220	0.233	0.482	0.122	0.604
0.1	0.645	0.163	0.434	0.129	0.369	0.165	0.327	0.100	0.427
0.15	0.639	0.084	0.480	0.055	0.428	0.132	0.249	0.081	0.330
0.2	0.567	0.095	0.454	0.086	0.385	0.097	0.199	0.121	0.320
0.25	0.526	0.065	0.432	0.051	0.392	0.048	0.177	0.076	0.254
0.3	0.520	0.058	0.409	0.057	0.347	0.062	0.213	0.119	0.332
0.35	0.491	0.055	0.374	0.078	0.325	0.050	0.238	0.100	0.339
0.4	0.464	0.053	0.361	0.071	0.297	0.054	0.221	0.138	0.359
0.45	0.456	0.051	0.345	0.056	0.296	0.073	0.243	0.108	0.352
0.5	0.440	0.051	0.336	0.024	0.283	0.049	0.237	0.120	0.357
0.55	0.421	0.055	0.330	0.045	0.277	0.038	0.215	0.128	0.343
0.6	0.398	0.055	0.305	0.025	0.271	0.051	0.234	0.084	0.318
0.65	0.386	0.043	0.292	0.042	0.274	0.053	0.244	0.045	0.289
0.7	0.379	0.038	0.293	0.051	0.272	0.033	0.226	0.058	0.284
0.75	0.398	0.046	0.283	0.052	0.268	0.039	0.290	0.036	0.325
0.8	0.424	0.037	0.287	0.038	0.272	0.041	0.322	0.037	0.359
0.85	0.443	0.046	0.307	0.039	0.273	0.047	0.307	0.078	0.385
0.9	0.476	0.059	0.316	0.044	0.290	0.058	0.337	0.054	0.391
0.95	0.476	0.092	0.311	0.104	0.303	0.074	0.346	0.017	0.363

Table 2: Pooled Regression Estimates

Quantile	Wald Statistic	P-Value
OLS	78.449	0.000
0.1	24.418	0.381
0.15	29.747	0.157
0.2	56.663	0.000
0.25	31.562	0.110
0.3	22.796	0.473
0.35	39.172	0.019
0.4	24.231	0.391
0.45	30.784	0.128
0.5	31.960	0.101
0.55	40.646	0.013
0.6	22.881	0.468
0.65	18.219	0.746
0.7	28.760	0.188
0.75	21.712	0.538
0.8	34.355	0.060
0.85	18.144	0.750
0.9	23.005	0.460

Table 3: Wald statistics for equality of IGE coefficients across the 1980-2010 period for each quantile and for the linear projection

	Data	Income Variable	Sample size (obs.)	IGE estimate	Impact of education	Impact of race	IGE trend
Solon (1992)	PSID	Log earnings averaged 5 years for parental income (1967-71) ; year 1984 for sons	290	0.41			
Zimmerman (1992)	NLS	Log earnings	192	Circa 0.4			
Eide and Showalter (1999) [parental earnings]	PSID	Log of average of three years of father's earnings (1967-69) and 7 years of son earnings (1984-91).	469	0.34	OLS Decrease in Income Elasticity of 29.4% (To 0.24).		
Eide and Showalter (1999) [parental income]	PSID	Log of average of three years of father's income (1967-69) and 7 years of son earnings (1984-91).	612	0.45	OLS Decrease in Income Elasticity of 26.7% (To 0.33).		
Grawe (2004)	PSID	Father earnings observed from 1967 to 1971, averaged if there are at least three observations; children earnings observed from 1978 - 81, included in the sample if there are at least three observations out of five.	354	0.47			
Hertz (2006)	PSID	Log of average family income per person. Children observed in the 1995, 1996, 1997, 1999 and 2001 surveys. Parents averaged in the 1968-72 surveys( 4 year average). Mean ages 37 and 38 respectively for parents and children.	4,004	0.51		Controlling for race reduces IGE from 0.515 to 0.429. That's a 16.7%	
Lee and Solon (2009)	PSID	Log of son family income controlling for life cycle on the years 1977-2000. Parental income averaged for three years (children aged 15-17).	11,230	0.44 (Avg)			No trend for the 1978-2000 period
Cooper (2011)	PSID	A sample of male heads. Average labor income of parents and sons who report at least 3 years of income at ages 35-50, from the years 1967 to 2007.	1,424	0.42	OLS Decrease IGE of 35% (To 0.27).		
Torche (2013)	NLSY-79	Log of Family income For adult children, she uses an average of family income over the 1996-2002 period. Parental income is the total household income during 1978, as reported by the parents in the first NLSY79 interview wave.	2,178	0.37	OLS Decrease of IGE of 54% (To 0.172), controlling for level of education.	OLS Decrease of IGE to 0.323 (13.63%) including race and a (non statistically significant) rural area control.	
Blanden et al. (2014)	PSID	Log averaged earnings for male children born between 1960 and 1970 measured at ages 30-34, with at least one observation. Parental income is averaged when the children was 10-16 with at least one observation.	647	0.38	48.1% of IGE explained by Education (Pathway decomposition method)		
Palomino, Marrero and Rodríguez (2015)	PSID	Log of family income controlling for life cycle on the years 1978-2000. Parental income averaged for seven years (children age 13-19).	25,084	0.47	OLS Decrease of IGE of 27.43% (to 0.34)	Additional OLS Decrease of IGE to 0.43 (10.1%).	Decreasing trend 1980-2000 period, turned increasing in 2002-2010.

Table 4: Review of OLS IGE estimates for the U.S. in the literature



	IGE Estimate (10th, 25th, 50th, 75th and 90th percentiles)	Impact of Education (% decrease controlling for years of education at the 10th, 25th, 50th, 75th and 90th percentiles)	Impact of race (% decrease controlling for race at the 10th, 25th, 50th, 75th and 90th percentiles)	IGE trend at different quantiles
Eide and Showalter (1999) [parental earnings]	0.47; 0.35; 0.37; 0.35; 0.17	30; 26; 35; 34; 12 (*)		
Eide and Showalter (1999) [parental income]	0.67; 0.49; 0.44; 0.35; 0.26	27; 35; 30; 26; 19 (*)		
Grawe (2004)	0.35; 0.494; 0.54; 0.457; 0.40			
Cooper (2011)	0.52; 0.49; 0.46; 0.41; 0.38	35; 32; 31; 33; 53 (*)		
Palomino, Marrero and Rodríguez (2015)	0.64; 0.53; 0.44; 0.40; 0.48	33; 18; 24; 29; 34	12; 8; 12; 4; 1	In the 1980-2010 period, no trend for the mid-high and high percentiles. For the mid and mid low percentiles, decrease of IGE in the 80s and 90s and slight increase in the 2000s.

(\*) Own calculation using the authors' reported results

Table 5: Review of Quantile Regression IGE estimates for the US in the literature

## A1 Online Appendix

### A1.1 Sensitivity analysis of database choices

#### *Sample choice*

The quality of the PSID database has often been assessed by comparing different distributions from this database with their equivalent in other sources. For instance, Gouskova et al. (2010) have compared estimates of family income between the PSID and the March Current Population Survey (CPS) for the entire history of the PSID (1968-2007). They find that the distributions are in close agreement throughout the 39-year history of the PSID, above all in the range between the 5<sup>th</sup> and 95<sup>th</sup> percentiles. To keep the maximum possible number of observations, we use the 'core' sample of the PSID, conformed by two independent probability samples: the first one is an equal probability sample of households based on a stratified multistage selection of the civilian non-institutional population of the U.S. (drawn by the Survey Research Center, SRC); the second one is a national sample of low-income households (drawn by the Survey of Economic Opportunity, SEO). The combination of both is also a probability sample, but selection probabilities are unequal and, therefore, population weighting is needed in the estimation of intergenerational income elasticity. These weights, designed to compensate for unequal selection probabilities and differential attrition, are supplied by the PSID.<sup>A1</sup> Despite the fact that some studies have previously considered only the SRC sample (Solon (1992); Lee and Solon (2009)), it is interesting to note that Solon ((Solon, 1992, p. 404)) found that his results were comparable when using the full core sample with weights and that Hertz (2007) has shown that, in terms of the evolution of the variance of family income, the combination of the SEO and the SRC samples resembles the much larger Current Population Survey (CPS) more than each of the samples alone.

Nevertheless, we have checked the robustness of our estimates when using only the SRC sample. For the pooled estimation, using only the SRC sample of the PSID yields an OLS estimate of 0.45 (Table A1), which is slightly lower than our OLS estimate of 0.47. Quantile regression estimates of the IGE still present a clear U-shaped relation with the son's position at the income distribution (Figure A1)

---

<sup>A1</sup>On the construction and revision of the PSID weights for the whole core sample see Gouskova et al. (2008). A representative sample of 2,043 Latino (Mexican, Cuban, and Puerto Rican) households was added to the PSID data in 1990. However, this sample missed out Asians, and because of this crucial shortcoming, and a lack of sufficient funding, the Latino sample was dropped after 1995. To avoid longitudinal inconsistencies, we have not considered this Latino sample and therefore, one must be aware that our PSID sample could not reflect the recent changes in the composition of the American population. Also, a recent study finds that attrition in the PSID affects specially the lower part of the distribution, and that IGE might be downward biased at the lower quantiles (Schoeni and Wiemers (2015)). This effect, although less important when -like in our analysis- sampling weights are used, could made the U-shaped Quantile-IGE relation even steeper at the low part of the distribution.

showing very similar estimates.

### Outliers

To test the sensitivity of the estimates to the choice of outliers, we have changed the data selection choice and kept all valid income observations except for negative values, instead of our baseline criteria for outliers proposed by Lee and Solon (2009). As expected, the inclusion of more extreme values affects significantly the OLS estimation, which for the whole pooled sample rises from 0.47 to 0.55 (Table A1). However, quantile estimation is quite robust to the inclusion of these outliers, except at the extreme quantiles, where IGE estimates are slightly higher, making the U-shape more pronounced (Figure A2).

### Permanent income and life cycle

In principle, income elasticity estimates need the permanent income component of parents and children. Unfortunately, it is usually not possible to have income data over the whole life cycle of individuals, so typically there is a bias in IGE estimation due to the life cycle bias and transitory shocks. Solon (1992), Zimmerman (1992)) or Mitnik et al. (2015) have proposed to average several years of parental income to proxy ‘permanent’ income and to reduce the effect of temporary shocks. A shorter averaged period of parental income is a worse proxy of permanent income and one should expect a lower value of the IGE in this case, due to the effects of transitory shocks that produce an ‘attenuation bias’ in the estimates (Mazumder (2005)). Our sensitivity analysis confirms this prediction, with an OLS value of the IGE of only 0.37 when we average up to 3 years of parental income when the child was between 15 and 17 years old (a mean of 2.76 years), instead of our choice of up to 7 years when the child was between 13 and 19 years old (mean 6.24 years). Also, when using only 3 years of parental income, QR estimates for the IGE are smaller across the entire income distribution, diverging especially at the top quantiles, where the attenuation bias caused by transitory shocks seems to be higher. When we increase the number of years averaged to up to 9 years in the 12-20 range of the child age (mean 8.12 years), the estimates do not change significantly from our baseline choice (Table A1 and Figure A3).<sup>A2</sup> It is also important to test the life-cycle trajectories across the period analyzed. Given that Equation 3 estimates the age control covariates for the whole pooled sample, there could be an estimation bias in the trend due to changing life income trajectories across cohorts. To analyze this possibility, we

---

<sup>A2</sup>Nybom and Stuhler (2016) show that without observations of parent and son incomes over the full lifespan of the individual, the ‘true’ lifetime IGE might always be underestimated. However, recent studies with large administrative databases find, respectively, that the attenuation bias is greatly reduced after five (Chetty et al. (2014b)) and nine (Mitnik et al. (2015)) years of income are averaged. In our case, the data show almost no change in the estimates when averaging more than 7 years of parental income. Also note that Lee and Solon (2009) average yearly parental family income when the child was between 15 and 17 years old (three years)

have calculated the average income at each age for cohorts born in the 50s, 60s, and 70s and have checked that the life cycle trajectory is similar for the three groups of cohorts (Figure A4).

#### The choice of the education and race variables

In addition to the main education variable used in the analysis (years of education), to test for possible non-linearities of the effect of education on income we also run our main pool analysis using a categorical education variable. The results were coincident, with only slight differences in the extreme quantiles of the distribution (see Figures 2a and A5a).<sup>A3</sup>

In the case of race, when we use a white race dummy in the main analysis, we also considered using a black race dummy, thus assimilating white to the other non-black race observations, and did a robustness check with that choice. Given that less than 3% of the observations report races other than black or white each year (see Table 1), the results of the impact of race in IGE for the whole pool were quite similar with either choice (see Figures 2b and A5b). Among the scarce non-black and non-white race observations, 19.8% of them were "Spanish-American, Latino Origin", 13.9% "Asian, Pacific Islander", 9.4% "American Indian, Aleut or Skimo" and 56.9% were coded "Other". Given the slight prevalence of the hispanics among this group, and the fact that hispanic socioeconomic demographics are far more similar to blacks than to whites (Sullivan et al. (2015)), we opted to keep our white race dummy all across our study.<sup>A4</sup>

### **A1.2 Sensitivity analysis of conditional and unconditional quantile regressions**

Both OLS and QR estimates are obtained based on the distribution of the outcome variable  $Y$  conditional to the distribution of the explanatory set of  $X$  variables. The law of iterated expectations permits to obtain an 'unconditional' expectation as the weighted average of the conditional expectations of the distribution. Thus, in general, the OLS estimated coefficients at the expectation (mean) of the conditional distribution can be used to recover the estimates of the 'unconditional' distribution. However, this strategy does not apply to quantile regression (Angrist and Pischke (2008)). Obtaining the unconditional quantiles estimates when there is no 'law of iterated quantiles' is a challenging and

---

<sup>A3</sup>We used the following categories: less than primary (less than 8 years of education completed), primary (8 years), some secondary education completed (more than 8 and less than 12 years), secondary (12 years), post-secondary (more than 12 and less than 16 years), tertiary (16 years) and post tertiary (17 years of education or more).

<sup>A4</sup>Note that the "Spanish-American" coding in the PSID disappears from 1985 and only reappears in 1990, while the "Asian, Pacific Islander" and the "American Indian, Aleut or Skimo" codings both appear only from 1985. That explains the high proportion of answers coded "Other". Also remember that we are only using the SRC and Census samples of the PSID, and not the so-called, "Latino Sample".

“work in progress” task for statisticians and econometricians.

So far, the most successful ‘unconditional quantile regression’ approach among practitioners is the method proposed by (Firpo et al. (2009)). They propose a two-step process which uses the ‘influence function’ concept to estimate the impact of different X covariates at different unconditional quantiles.

<sup>A5</sup> Unfortunately, the presence of age controlling variables in our main specification makes this method not directly applicable in our case. These age controlling variables are included in the model to correct for measurement error, control life cycle bias and make all income observations comparable. Therefore, not conditioning the estimation to these controls would give us unreliable estimates of IGE.

Nevertheless, to check the robustness of our results, we will compare in our sensitivity analysis (see the online appendix) conditional and unconditional estimations of IGE with an age-restricted subsample.

We find the results in the conditional and unconditional regressions to be overall quite similar.

As explained in Section 2 in the main text, we cannot directly apply the unconditional quantile regression proposed by Firpo et al. (2009), to our main specification (2). Nevertheless, we wanted to check whether the estimation at the unconditional quantiles would differ significantly from the conditional quantile regression we apply. In order to compare both methods, we created a subsample of our pooled observations keeping only pairs of observations when both the son and parent were between 35 and 45 years old (note that parental age and income are always measured when the son was 16).

This way, we can remove the age control covariates from the model and apply unconditional quantile regression. This subsample still has 4583 observations, not sufficient to undergo yearly estimations –as we have done with the full sample– but enough to estimate IGE and the role of education and race.<sup>A6</sup>

For the unconditional estimates, the results for the IGE are in Table A2 and plotted in Figure A6. We observe that the results are quite similar to those obtained with the full sample using conditional quantile regression in an age restricted sample. Also, the restricted sample estimates are not significantly different from the ones of the baseline specification with the whole sample, confirming that the age controls included in our main analysis are effectively controlling for the life cycle. With respect to the effect of education and the importance of race in the IGE at different quantiles, unconditional estimation provides again similar measures (Figure A7).

---

<sup>A5</sup>The first step estimates the ‘Recentered Influence Functions’ (RIF) for observations below and above each unconditional quantile; the second step regresses these RIF values on the X covariates.

<sup>A6</sup>We have used the RIF-OLS function provided by Nicole Fortin (available at <http://faculty.arts.ubc.ca/nfortin/datahead.html>).

### A1.3 Education and race impact on IGE between 1980 and 2010

The share of the IGE captured by the 'Years of education' variable at each year and quantile of interest is displayed in Figure A8. At the lower quantiles, the share of the IGE attributed by education shows an increasing pattern in the 80s and 90s, and a stable level in the 2000s. At the mid-low quantiles, median and at the OLS estimation, the role of education remains fairly steady along the whole period, only slightly decreasing in the mid 90s, and increasing again from the mid 00s. For the mid-high quantiles, the pattern seems to be increasing in the 80s, slightly decreasing in the 90s, and slowly increasing again in the 2000s. At the high quantiles -like in the mid-low and mid-high groups of quantiles- there is a fall in the importance of education in IGE in the late 90s and followed by an increase in the 2000s. The importance of race in the intergenerational transmission of income (Figure A9) decreased during the 80s and 90s all across the distribution and at the mean (OLS estimation). At the lower quantiles, where race had a stronger role in the beginning of the 80s, this reduction was more pronounced. In the quantiles around the median (mid-low and mid-high figures) this decreased occurred mainly in the 90s. From the early 2000s, however, this trend was reversed and the impact of race on IGE shows a similarly increasing trend all across the income distribution. Remarkably, this increase offsets a great part of the decrease that took place in the previous two decades.<sup>A7</sup>

---

<sup>A7</sup>The 1991 Civil Rights Act against discrimination may have contributed to this declining importance of race as a conditioning factor in the transmission of parental income, and to the decrease of the income 'white premium'. As with the 'years of education variable' (see Footnote 17), the impact of race on IGE might be encompassing other factors correlated with race. Chetty et al. (2014a) point out that it is demographical segregation and the level of public goods in an area what has a greater impact on mobility. They find mobility seems to be lower for people living in these areas regardless of the actual race of the individual. In any case, we believe the upturn we find in the role of race from the early 2000s makes this topic deserving of a detailed analysis in future research.

## A1.4 Online Appendix Figures and Tables

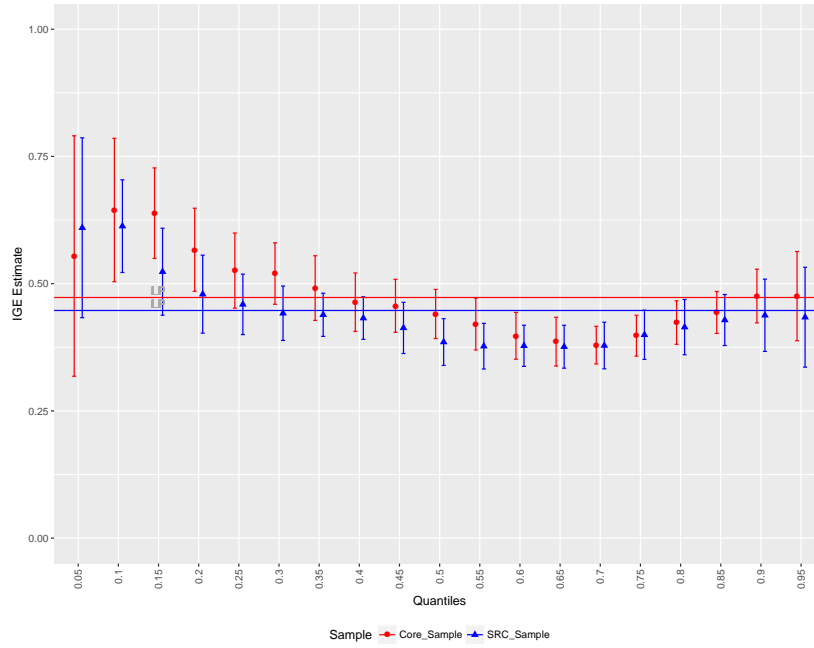


Figure A1: Sensitivity of pooled IGE estimates to PSID sample used. Baseline choice: 25084 observations. SRC sample: 16239 observations. Includes quantile estimates with standard errors and linear projections (LP) obtained with OLS. Standard errors for the linear projection have not been plotted for clarity but are available in Table A1.



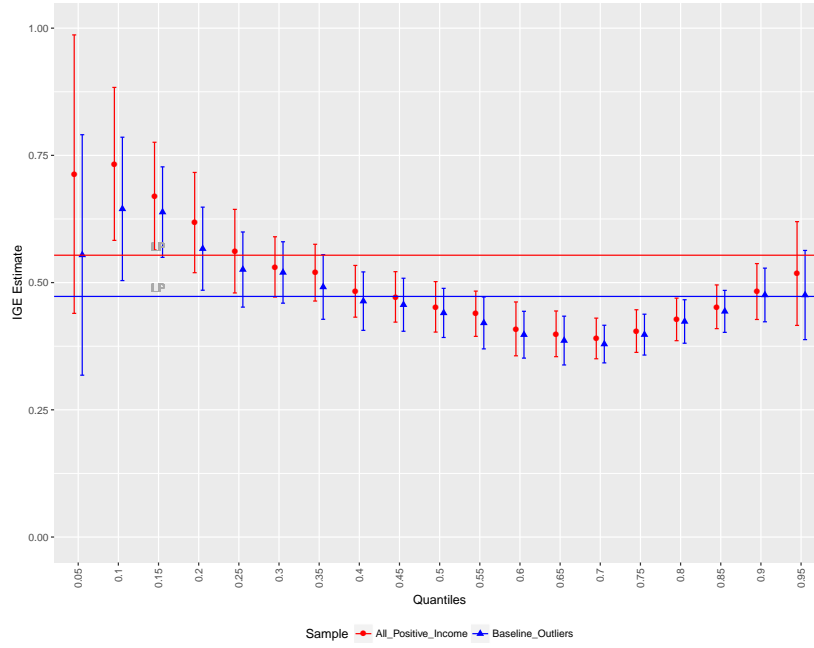


Figure A2: Sensitivity of pooled IGE estimates to the outlier thresholds. Baseline choice (excluding income under 100 and over 150000 USD dollars of 1967): 25084 observations; Excluding only negative income: 25271 observations. Includes quantile estimates with standard errors and linear projections (LP) obtained with OLS. Standard errors for the linear projection have not been plotted for clarity but are available in Table A1.

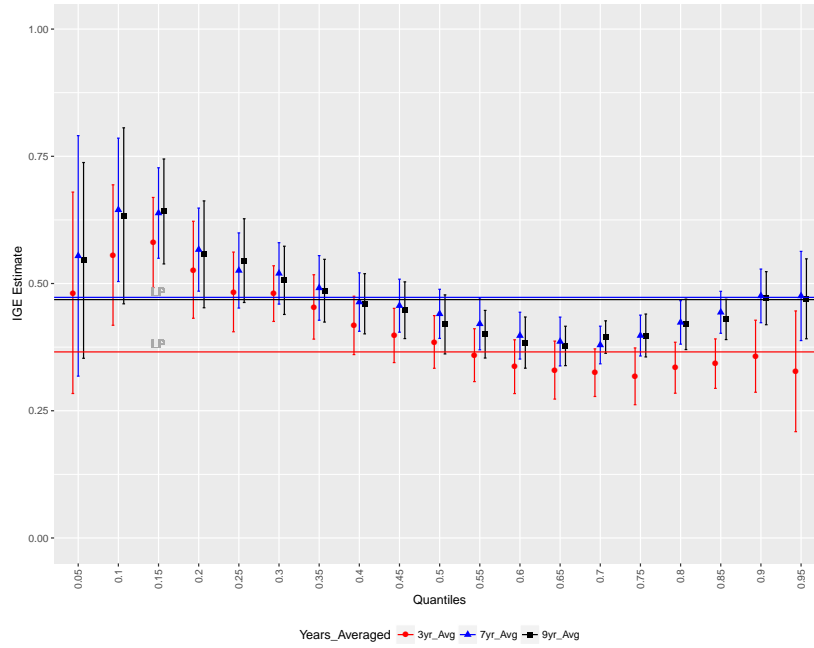


Figure A3: Sensitivity of pooled IGE estimates to the number of years averaged. 25245 observations when averaging up to 3 years of parental income (mean 2.76 years). 25084 observations when averaging up to 7 years income (mean 6.24 years) and 23577 observations when averaging up to 9 years of parental income (mean 8.13 years). Includes quantile estimates with standard errors and linear projections (LP) obtained with OLS. Standard errors for the linear projection have not been plotted for clarity but are available in Table A1.

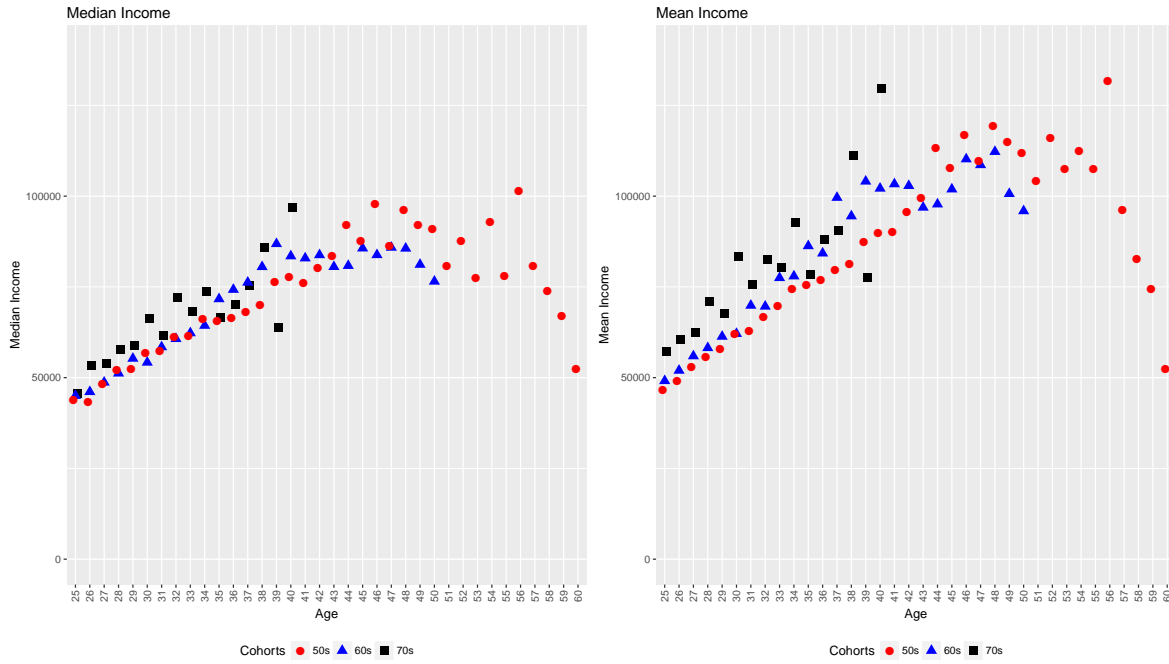
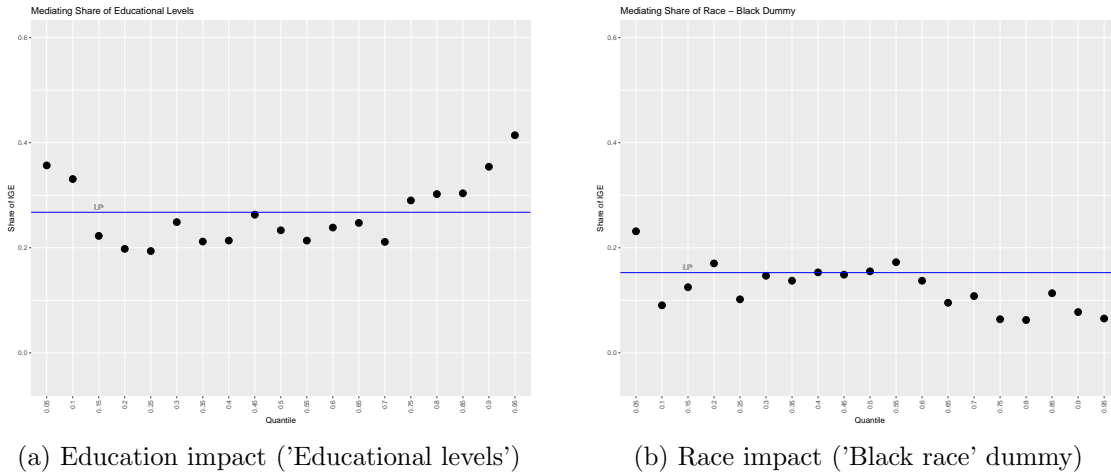


Figure A4: Life cycle trajectories for cohorts by decades



(a) Education impact ('Educational levels')

(b) Race impact ('Black race' dummy)

Figure A5: Impact of Education and Race as a share of IGE. Pooled estimates.

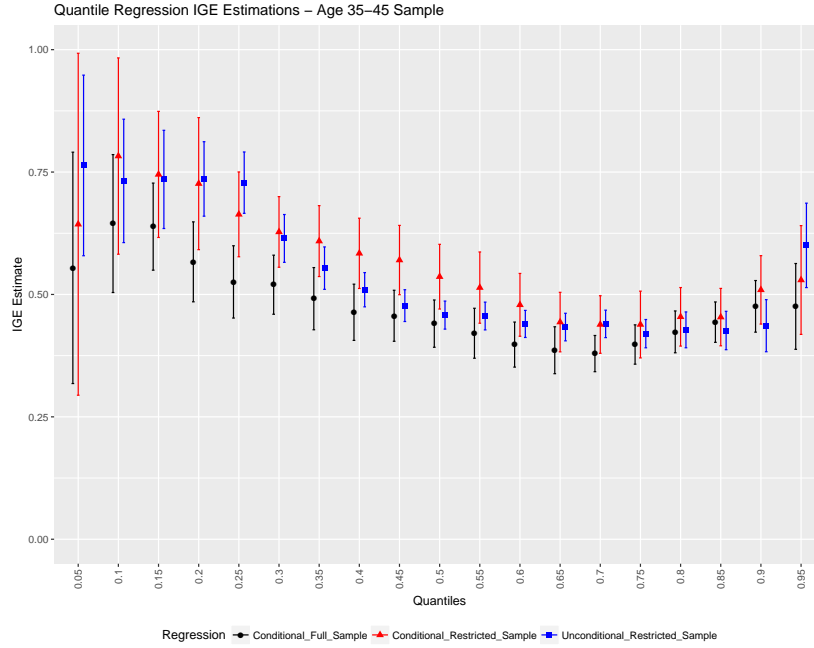


Figure A6: Conditional and Unconditional Estimates of IGE. Baseline model: 25084 observations of individuals between 25 and 65 years old, with age controls in the specification. Age restricted model: 4583 observations of individuals only between 35 and 45 years of age, no age controls in the specification

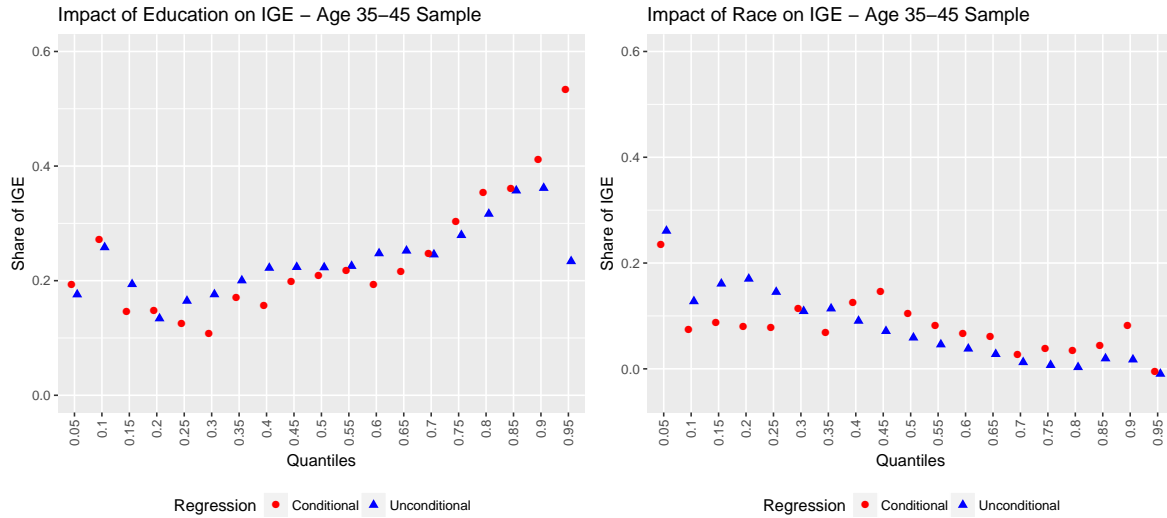


Figure A7: Conditional and Unconditional Estimates of Education and Race impact on IGE. Age restricted model: 4583 observations of individuals only between 35 and 45 years of age, no age controls in the specification

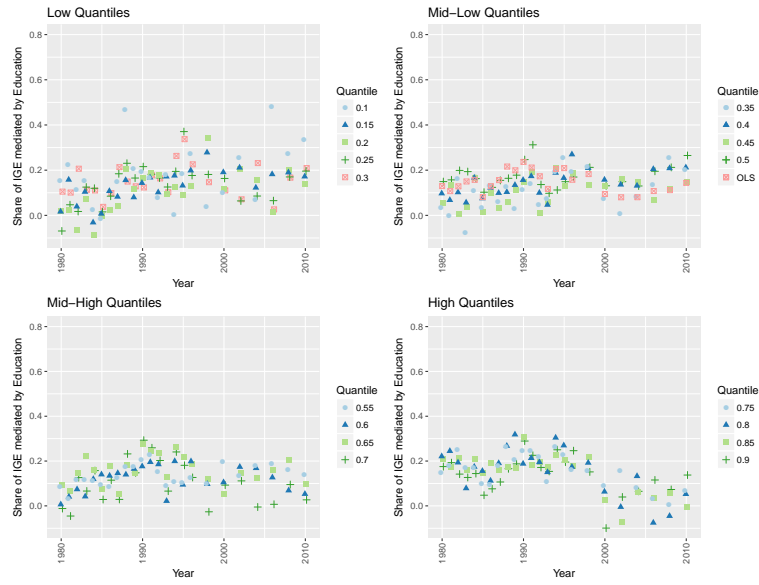


Figure A8: Share of IGE decrease when controlling for Education. Trend 1980-2010

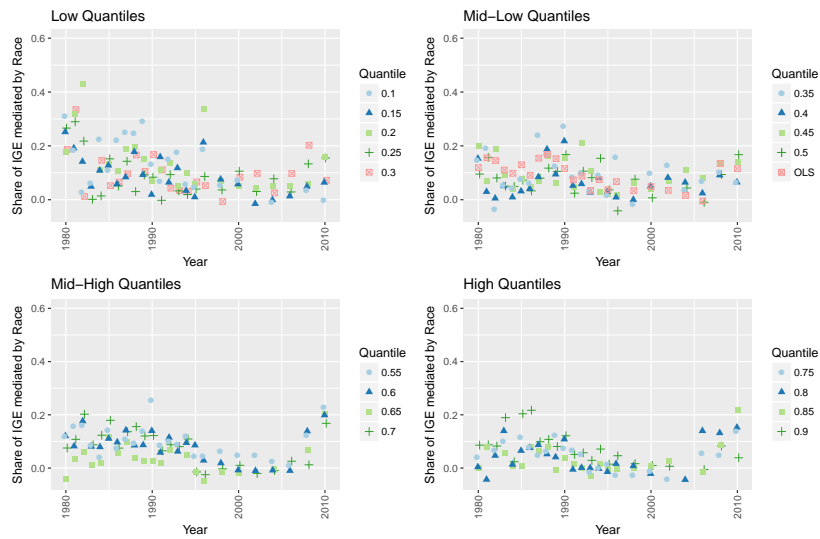


Figure A9: Share of IGE decrease when controlling for Race. Trend 1980-2010

Quantile	Baseline Estimates		SRC Sample		3 yr. Avg		9 yr. Avg		Exclude Neg. Income	
	IGE	(SE)	IGE	(SE)	IGE	(SE)	IGE	(SE)	IGE	(SE)
OLS	0.473	0.018	0.447	0.021	0.366	0.015	0.468	0.019	0.554	0.019
0.05	0.554	0.223	0.610	0.185	0.482	0.216	0.545	0.197	0.713	0.282
0.1	0.645	0.163	0.613	0.096	0.556	0.135	0.633	0.181	0.733	0.151
0.15	0.639	0.084	0.523	0.076	0.581	0.088	0.642	0.102	0.670	0.100
0.2	0.567	0.095	0.479	0.076	0.527	0.083	0.557	0.096	0.618	0.089
0.25	0.526	0.065	0.459	0.052	0.483	0.071	0.545	0.091	0.562	0.079
0.3	0.520	0.058	0.442	0.047	0.480	0.060	0.506	0.071	0.531	0.068
0.35	0.491	0.055	0.439	0.041	0.454	0.059	0.486	0.061	0.520	0.054
0.4	0.464	0.053	0.432	0.048	0.418	0.056	0.460	0.048	0.483	0.052
0.45	0.456	0.051	0.413	0.047	0.398	0.046	0.448	0.059	0.472	0.041
0.5	0.440	0.051	0.385	0.048	0.385	0.050	0.420	0.051	0.452	0.053
0.55	0.421	0.055	0.377	0.045	0.359	0.052	0.400	0.048	0.439	0.055
0.6	0.398	0.055	0.378	0.042	0.337	0.047	0.384	0.051	0.409	0.050
0.65	0.386	0.043	0.376	0.034	0.330	0.048	0.377	0.041	0.399	0.044
0.7	0.379	0.038	0.378	0.035	0.325	0.045	0.395	0.031	0.390	0.035
0.75	0.398	0.046	0.399	0.054	0.318	0.052	0.398	0.040	0.405	0.049
0.8	0.424	0.037	0.415	0.056	0.335	0.048	0.420	0.046	0.428	0.041
0.85	0.443	0.046	0.429	0.051	0.343	0.055	0.429	0.041	0.452	0.049
0.9	0.476	0.059	0.438	0.058	0.357	0.058	0.471	0.050	0.482	0.051
0.95	0.476	0.092	0.434	0.125	0.328	0.108	0.470	0.085	0.518	0.102

Table A1: Sensitivity Analysis: IGE estimates for different data options. Baseline options has 25084 observations; SRC Sample option has 16239 observations; Up to 3 years averaged for parental income has 25245 observations; Up to 9 years averaged has 23577 observations ; The option of outliers excluding only negative observations has 25271 observations.

Quantile	Conditional Regression		Unconditional Regression	
	IGE	(SE)	IGE	(SE)
0.05	0.643	0.349	0.764	0.184
0.1	0.783	0.200	0.732	0.126
0.15	0.745	0.129	0.735	0.100
0.2	0.726	0.135	0.736	0.076
0.25	0.664	0.087	0.728	0.063
0.3	0.628	0.072	0.614	0.049
0.35	0.609	0.072	0.554	0.043
0.4	0.584	0.072	0.510	0.035
0.45	0.570	0.071	0.477	0.032
0.5	0.536	0.066	0.458	0.029
0.55	0.514	0.073	0.456	0.028
0.6	0.479	0.064	0.440	0.028
0.65	0.444	0.061	0.433	0.028
0.7	0.439	0.059	0.440	0.028
0.75	0.439	0.068	0.420	0.029
0.8	0.454	0.060	0.428	0.037
0.85	0.454	0.059	0.426	0.039
0.9	0.509	0.070	0.436	0.053
0.95	0.530	0.111	0.600	0.086

Table A2: Conditional and Unconditional Estimates. Baseline (Age Restricted) Model. The age restricted model does not include age controls in the specification, and includes only individuals between 35 and 45 years old. Total sample of 4583 observations