The Evolution of the U.S. Family Income-Schooling Relationship and Educational Selectivity*

Christian Belzil†and Jorgen Hansen‡

June 15, 2018

Abstract

We estimate a reduced-form model of schooling on two cohorts of the National Longitudinal Survey of Youth and find that, contrary to conventional wisdom, the effects of family income on college entry and four-year college graduation have practically vanished between the early 1980's and the early 2000's. The effects of AFQT scores also decreased substantially over the same period. Unobserved heterogeneity, which was the most important determinant of college participation and college graduation in the 1980's, has become even more important. Ignoring it would lead to an over-estimation of the drop in college participation income effects and an under-estimation for college graduation. The reduced impacts of family income on both college attendance and four-year college graduation are robust to the allowance for a correlation between family income and unobserved heterogeneity.

*Financial support from Investissement d'Avenir (ANR-11-IDEX-0003/ Labex Ecodec/ANR-11-LABX-0047) and Social Sciences and Humanities Research Council of Canada (grant 435-2017-0129) is gratefully acknowledged. We thank Philippe Belley, Evgenia Dechter, Eleanor Dillon, Lance Lochner, Michael Lovenheim, Arnaud Maurel and Chris Taber for comments and exchanges. The usual disclaimer applies.

†CREST, CNRS, Ecole Polytechnique, IZA and CIRANO

‡Concordia University, CIRANO, CIREQ and IZA

JEL Classification: I2, J1, J3
Keywords: Inequality, Education, Family Income.

1 Introduction

We confront the evolution of the effects of family income and Armed Forces Qualification Test (AFQT) scores on higher education outcomes to the evolution of educational selectivity between the early 1980's and the early 2000's. To do so, we estimate a sequential model of schooling with unobserved heterogeneity on two cohorts of the National Longitudinal Survey of Youth (NLSY).

At the outset, it should be clear that our analysis is not about replication but about modeling the family income-schooling gradient differently than it has been achieved in the recent literature.[1] Our approach differs in two main dimensions.

First, we measure the change in the effect of real income as opposed to the effect of a change in relative income (or inequality) as is done in the existing literature. Differences in schooling outcomes between specific family income quartiles, and compared across cohorts, are imputable to both a change in income distribution as well a change in the effect of real family income which may itself arise from substitution effects induced by changes in the cost of education or in financial aid opportunities. The former may be particularly important if education services are a normal good and the latter may arise if the cost of education moves across cohort or if institutional and non-institutional financial aid opportunities also vary.

Second, we use statistical methods which allow us to control for dynamic selection within each cohort and therefore to compare changes in educational selectivity, induced by movements in the distribution of unobserved heterogeneity, to changes in the effect of individual and family characteristics. The confounding effect of ignoring educational selectivity is a recurrent theme in the paper. Throughout the paper, we use the term unobserved heterogeneity to designate any unmeasured factor such as taste for schooling, monetary or

[1] A review of the literature is found in the next section.

non-monetary costs of education or ability and motivation

However, following the existing literature, we adopt a descriptive (non-structural) approach to the extent that we treat family income and other individual and family characteristics as exogenous regressors. In the paper, we answer the following questions:

- How have the effects of family income, AFQT scores and other characteristics on college participation and four-year college graduation evolved between the early 1980's and the early 2000's?

- Is the evolution of income effects dependent on the level of income and/or AFQT scores?

- How important was unobserved heterogeneity for education in the early 1980's compared to the early 2000's?

- What are the consequences of ignoring dynamic selection when comparing income and AFQT effects across cohorts?

- Do model specifications allowing for a correlation between unobserved heterogeneity and family income deliver similar conclusions as models which ignore it?

The main findings are the following. Contrary to conventional wisdom, the effects of family income on college participation and four-year college graduation have practically vanished. The decrease in magnitude of income effects is general. It depends neither on AFQT score nor on family income level. Thus, we find no evidence that the effects of family income on both college participation and college graduation have increased over the period going from the early 1980's to the mid 2000's.

Similar to the income effects, the effects of AFQT scores on college participation and college completion have decreased significantly. This is especially the case for college attendance where the effects have been reduced by a factor of two to three, depending on the evaluation point. However, the effects remain significant in the early 2000's.

Both in the early 1980's and the early 2000's, unobserved taste for education was by far the most important determinant of the decision to enter college (conditional on high school completion) since no single individual attribute could match its importance. Moreover, the importance of unobserved heterogeneity has increased in relative terms compared to individual and family characteristics such as family income, AFQT scores and other attributes. Indeed, in the early 2000's, unobserved heterogeneity was virtually as important as the entire set of individual and family characteristics for transitions from high school to college. Put differently, the relative contribution of family income to higher education outcomes has dropped because the decrease in the effect of real income has been sufficiently strong to annihilate the impact of the increase in income dispersion observed over the same period.

Our findings therefore suggest that classical educational selectivity based partly on cognitive abilities and/or family income is being gradually replaced by a different form of selectivity based on unmeasured abilities or preferences. Both for college participation and four-year college graduation, ignoring educational selectivity driven by unobservables provides distorted measures of the evolution of the effects of real family income and AFQT scores on higher education outcomes. Ignoring it would lead to an over-estimation of the drop in college participation income effects and an under-estimation of college graduation income effects.

Finally, the vanishing impacts of family income on higher education participation and on four-year college graduation are robust to the allowance for a correlation between unobserved heterogeneity and family income. As is the case with observed AFQT scores, our results suggest that unobserved heterogeneity (at age 16) is positively correlated to family income within each cohort, but that this positive correlation has decreased substantially.

The remaining sections of the paper are structured as follows. In section 2, we present a brief review of the relevant literature. The following section is devoted to the description of the data. In Section 4, we present our econometric model. In the next section, we describe the evolution of the marginal effects of income and AFQT scores. In Section 6, we build on the results of Section

5 to illustrate the implications of ignoring dynamic educational selectivity. In the following section, we investigate to what extent our results are robust to the allowance for a correlation between family income and unobserved heterogeneity. A summary of the results along with some economic interpretation concludes the paper in Section 8.

2 Background Material

The relationship between education inequality and access to financial resources is one of the most contentious issues debated over the past 20 years in the US. As many key determinants of education choices such as parental transfers, borrowing limits, and financial aid are not precisely measured in observational data, it is particularly difficult to obtain clean evidence on the existence of financial barriers to educational achievements, let alone their evolution across cohorts.

For this reason, many economists have estimated reduced-form models of educational choices and used them to evaluate the impact of family income on higher education enrollments. This approach is motivated by the existence of a strong empirical correlation between family income and family resources devoted to education financing.

Cameron and Heckman (1998) estimated an ordered discrete choice model of schooling choices on five different cohorts of US males born between 1907 and 1964 using data from the Occupation Change in a Generation (OCG) and the NLSY79 cohort and report relatively small effects of a 10 percent increase in family income on enrollment and graduation probabilities. They stressed the relative unimportance of family income compared to family human capital indicators and cognitive ability, as measured by Armed Forces Qualification Test (AFQT) scores.[2] Other studies such as Keane and Wolpin (1997), Carneiro and Heckman (2002) and Cameron and Taber (2004) have

[2]As the authors do not report standard errors for the impact of an increase in family income (Table 11, page 314), it is difficult to assess the evolution of the effect of family income over this long period which precedes the period over which income inequality has progressed substantially.

confirmed this finding within diverse frameworks. All of these studies were concerned with cohorts of individuals who made their college participation decision in the early 1980's.

However, the well documented increase in wage inequality taking place between the late 1970's and the early 2000's, coupled with the steady increase in publicly posted tuition costs of four year college (the sticker price), has stimulated interest in the evolution of the effect of family income on educational attainments. There are good reasons for that. In presence of either exogenous borrowing constraints or endogenous constraints driven by various forms of limited commitments, most theoretical models predict that parental transfers (approximated by parental income) can play a role in the decision to invest in higher education.[3]

Based on a comparison of the 1979 cohort of the NLSY with the 1997 cohort, Belley and Lochner (2007) conclude that family income has become a more important determinant of college enrollments in the early 2000's than in the 1980's. To establish their results, the authors essentially regress binary educational outcome indicators, measured at age 21, on relative income measures (the top vs. bottom quartiles), AFQT scores and other regressors measuring individual and family background heterogeneity. They report that differences in mean outcomes between the top and the bottom family income quartiles is higher for the 1997 cohort than it is for the 1979 cohort.

The increasing importance of family income on educational outcomes appears to be widely accepted. Claims about the increasing gap in educational outcomes between low and high income classes are also found in Avery and Kane (2004), Bailey and Dynarski (2011) and Page and Scott-Clayton (2015). In their analysis of the 1979 and 1997 cohorts of the NLSY, Bailey and Dynarski (2011) report not only an increase in educational gaps across low and high income groups but also an increasing gap in four-year college completion after conditioning on post-high school participation.[4]

---

[3] The literature on human capital and liquidity constraints is surveyed in Lochner and Monge-Naranjo (2010).

[4] The increasing effect of parental income on educational outcomes is invoked as one of the main motivations for incorporating credit (liquidity) constraints within structural models

Some recent studies have however offered different perspectives on the evolution of educational inequality. Kinsler and Pavan (2011), who investigated gaps in college quality between different income quartiles, report that the effects of family income on college quality has been stable for average ability students and has even decreased for the more able. Chetty et al (2014), who were primarily interested in the evolution of the inter-generational income correlation, document that education gaps between low and high income US families have been relatively stable and dropped for the most recent cohorts (those born after 1985).[5]

In line with our approach based on measuring the impact of real income (as opposed to relative income), Lovenheim and Reynolds (2011) estimate a multinomial logit model of two-year and four-year enrollments on two samples of high school graduates taken from the 1979 and 1997 cohorts of the NLSY.[6] To measure the effect of real income, they use four income groups defined from the 1997 quartiles which they interact with AFQT terciles. Although the authors conclude against the existence of a steeper income gradient within the 1997 cohort (except perhaps for high ability males), they also recognize that ignoring unobserved heterogeneity may have a substantial impact on their results.

There are two important observations to be made about the existing literature. First, and in line with the vast literature documenting the increase in wage inequality, those who have investigated the evolution of educational inequality have therefore focused mostly on documenting education gaps between various family income quantiles (aside from Lovenheim and Reynolds, 2011). This is surprising as in absence of any actual measure of family resources

of human capital accumulation. This is the case in Lochner and Monge- Naranjo (2010), Johnson (2013), Abbott et al (2016) and Hai and Heckman (2017).

[5]In parallel to the literature on education inequality, several studies concerned with wage inequality have attempted to measure recent changes in the effect of abilities on wages. For instance, Castex and Dechter (2014) have documented a decrease in the effect of AFQT scores on wages using both the 1979 and 1997 cohorts of the NLSY. Beaudry et al (2013) document a decline in the demand for high-skilled workers since 2000 and show that high-skilled workers have moved down the occupational ladder and have begun to perform jobs traditionally performed by lower-skilled workers.

[6]Lovenheim and Reynolds (2011) use a restricted access version of the NLSY.

devoted to higher education, differences in real income are much more likely to approximate access to financial resources than relative income measures.[7]

A second feature is the absence of any discussion of the effect of educational (dynamic) selection. The sensitivity of the evolution of the marginal effects of family income and AFQT scores to potential changes in educational selectivity therefore remains undocumented. This latter dimension is particularly important when evaluating the effect of income on the probability of continuing to college or on the probability of completing college, as the correlation between family income (or any other characteristic) and unobserved heterogeneity becomes more and more negative with grade completion as long as educational attainments are affected by neglected heterogeneity. Indeed, a large body of papers using structural dynamic methods have pointed out that unobserved heterogeneity in preferences for schooling was more important than any other observed characteristic in the NLSY79 (Keane and Wolpin, 1997). On top of this, if the relative importance of unobserved heterogeneity changes across cohorts, any comparison between marginal effects of income obtained from OLS across different cohorts becomes completely uninformative (even if income is an exogenous regressor).[8]

Evaluating the importance of dynamic selection requires to model the educational process as a sequential process in which unobserved heterogeneity plays an explicit role. This is precisely what we achieve in this paper. As is done in the previous literature on the effects of family income on actual grade attainments, we also adopt a descriptive (non-structural) approach and treat family income and other individual and family characteristics as exogenous regressors. In order to facilitate comparison with existing papers that ignore dynamic selection, we try to incorporate the same regressors used in papers performing comparisons between the 1979 and the 1997 cohorts of the

[7]Bils and Klenow (2000) provide evidence in favor of the claim that education is a normal good. Heckman and Mosso (2014) surveys the empirical literature on family income and child development.

[8]Evaluating the importance of changes in educational selectivity is currently raising much interest in the literature (see Ashworth et al (2017), Castro and Coen-Pirani (2016) and Dillon and Veramendi (2018)).

NLSY (Belley and Lochner (2007), Kinsler and Pavan (2011) and Bailey and Dynarski (2011)).[9]

3 Data

Our analysis is based on data from two cohorts of the National Longitudinal Survey of Youth, NLSY79 and NLSY97. The NLSY79 is a nationally representative sample of 12,686 young men and women who were 14-22 years old when they were first surveyed in 1979 while the NLSY97 consists of a nationally representative sample of 8,984 youths who were 12-16 years old as of late December 1996. For both NLSY cohorts, there are detailed information on family background and income as well as on individual scholastic ability (measured by AFQT scores). Interviews are ongoing for both cohorts and conducted on a annual or biannual basis.

Because we are primarily interested in the effect of family income on college decisions, we remove all respondents who are older than 18 at the time of the first survey. In the end, we retain only respondents born between 1961 and 1964 in the NLSY79 and respondents born between 1980 and 1983 in the NLSY97. Our selection criteria in this regard therefore closely resemble those used by Belley and Lochner (2007) and Kinsler and Pavan (2011).

Further, we limit our sample to respondents of each respective cross-sectional sample and exclude those with missing information on included observed characteristics such as family income, AFQT scores, mother's education, family stability (whether the individual report having been raised within a nuclear family or not), number of siblings, age of the mother at birth, area of residence (urban vs. rural), and ethnic background. Given our focus on college attendance and graduation, we also require individuals to be observed for the first five surveys. After these exclusions, we obtain samples of 2,180 individuals for

[9]However, studies using both the 1979 and 1997 cohorts of the NLSY do not necessarily use the same regressors. For instance, Lovenheim and Reynold (2011) control for both father's and mother's education and split males and females while Belley and Lochner (2007) ignore father's education and group males and females together.

the 1979 cohort, and 2,259 individuals for the 1997 cohort.[10]

Following Belley and Lochner (2007) and Kinsler and Pavan (2011), we use information on family income for each individual at ages 16 and 17, if available, and construct an average income measure. If income is only available for one of the years, the average income is replaced by that income. If no income information is available for these ages, we consider income at earlier ages if available in order to minimize the number of individuals dropped because of missing income. For both cohorts, we express income in year 2000 dollars using the CPI for all urban consumers. Like the earlier literature, we use AFQT scores to control for cognitive ability. The scores, provided by Altonji et al (2012), were adjusted to improve comparability across cohorts.

For each individual, we measure schooling attainment as highest grade completed by age 26. However, for individuals who drop out of the sample before age 26, we use their highest reported education in their last interview. Measuring schooling as of age 26 constitutes a major difference with Belley and Lochner (2007) who analyze schooling attainments by age 21.

Summary statistics are found in Table 1. First, we note that over a 20 year period, family income has grown by 23 percent on average (from $53,969 to $66,491). This corresponds to a 1 percent growth rate per year. These annual growth rates match aggregate measures provided by the Bureau of Labor Statistics.[11]

Other studies, including Kinsler and Pavan (2011), Lovenheim and Reynolds (2011), Castex and Dechter (2014) and Nielsen (2015) which are also comparing the 1979 and 1997 cohorts report similar family income growth.[12]

As is well known, income dispersion has increased even more over this period. In our sample, the standard deviation of family income increased from $33,210 in the 1979 sample to 59,500 in the 1997 sample. Not surprisingly, and as documented in Table 1, family income is higher than average among

[10] A detailed description of exclsusions

[11] According to the Bureau of Labor Statistics (variable MEFAINUSA672N), median household income grew by 20.6 percent between 1980 and 2000.

[12] Castex and Dexter (2014) report changes in the logarithm of income but their sample data also discloses a growth in real income levels which is comparable to ours.

college entrants and especially college graduates. Among observed characteristics, mother's education (going from 11.7 to 13.2 years) and the intact family indicator (going from 0.786 to 0.583) are those that have changed the most.

Unlike family income, the average AFQT score has remained more or less stable across cohorts. Average AFQT scores of those who have graduated from college exceed both average AFQT among college participants and average AFQT in the population. There seems to be a slight decrease in AFQT scores of college participants and a more important one (from 196 to 189) among four-year college graduates. In line with the recent evolution of college selectivity described in Hoxby (2009), this suggests that college has become globally less selective.

To obtain a clearer picture of the relationship between education and family income, we compute average schooling attainments statistics for the first, second, third and fourth income quartiles in the 1979 cohort and compare them to the corresponding quartiles in the 1997 cohort. We consider two statistics: the proportion of college attendants (those reporting 13 years of schooling or more) and the fraction of the population graduating from a four-year college (those reporting 16 years of schooling or more).

These statistics, found in the first and third columns of Table 2 incorporate both the effects of the overall increase in income and the increase in income dispersion as the income thresholds defining the third and fourth quartiles are those that have increased notably.

In order to remove the effects caused by the increase in income dispersion at high income levels, we also report average schooling attainments in 1997 evaluated at the 1979 quartiles. These numbers, presented in column two of Table 2, allow us to remove the impact of the larger income increase arising at the upper income quartile.

There are five observations that should be retained after examining the first three columns of Table 2. First, and regardless of whether schooling attainments of the 1997 cohort are evaluated at the 1979 income distribution (column two) or in terms of the 1997 thresholds (column three), there is evidence that college attendance and college graduations have increased at all income quartiles.

A second feature emerging when measuring schooling outcomes using each cohort's own income distribution is that the largest increases in graduation and enrollment rates are observed for those in the third quartile while the lowest increases are for the fourth quartile.

A third notable feature is that when analyzing differences between cohorts while keeping income intervals of the 1979 cohort, the first, second and third income quartiles experience increases in attendance of about 14 percentage points, while the highest income quartile experiences an increase of seven percentage points. For college graduation, the fourth quartile also experience the smallest increase (five percentage points) while the third quartile experiences the largest (13 percentage points).

A fourth observation relates to the effect of the increase in income dispersion. Because the change in income distribution has been realized mostly at the third and fourth quartiles, both enrollment and graduation rates for respondents in those quartiles, using the income thresholds for the younger cohort, exceed the enrollment and graduation rates of those in the third and fourth quartiles computed using the 1979 income thresholds.

Finally, a striking feature resulting from the comparison of the differences in attendance and graduation rates between successive quartiles (when the 1979 quartiles are used for both cohorts) is a relative stability (or small differences) at lower quartiles and a convergences in enrollment and graduation rates between the top income quartiles. For instance, the differences in enrollment rates between the top two quartiles dropped from 17 percentage points in 1979 (0.496 vs 0.673) to 10 percentage points in 1997 (0.653 vs 0.746). For college graduation, a similar pattern is observed. As these numbers have been obtained while holding income intervals fixed, they suggest that the marginal effect of real income tend to decrease, at least at relatively high income levels.

Similar to the analysis of income, we split the population into four AFQT quartiles and compute participation and graduation rates for each category. Unlike family income, the distribution of AFQT scores has remained relatively stable, so we only compute outcomes using each cohort's own thresholds.

Those are found in the last two columns of Table 2.

Similar to the observations for income, both participation and graduation have increased for all AFQT levels. The change in attendance rates per quartile are equal to 0.13 (quartile one), 0.2 (quartile two), 0.17 (quartile three) and 0.05 (quartile four). For graduation rates, the changes are equal to 0.045 (quartile one), 0.165 (quartile two), 0.17 (quartile three), and 0.06 (quartile four). The largest changes in college participation and attendance have therefore been observed among those belonging to the second and third quartiles.

Before discussing the econometric model, we report OLS estimates of the effects of real income for each cohort in Table 3a. To ease comparison with the previous literature on the effect of relative income on higher education, we include a similar set of regressors in the regression equations. The most striking result is the decrease in importance of family income. The marginal effect of a $10,000 increase in real income on college attendance, equal to 0.019 in the 1979 cohort, is only equal to 0.007 in the 1997 cohort. Similarly, the income effect on college graduation, which is equal to 0.016 in the 1979 cohort, drops to 0.008 in the 1997 cohort. While these results may seem at odds with those reported in the literature, it should be noted that changes in the income interquartile range caused by an increase in earnings dispersion are likely to induce a much larger impact on schooling than an increase in real income.

In Table 3b, we present estimates from OLS regressions of college attendance and college graduation indicators on AFQT and income quartiles instead of the linear measures in Table 3a. The remaining regressors are identical to those in Table 3a. The estimates suggest that the effect of income in quartiles two and three have increased over time, especially at the college graduation level. For example, the estimate of the second income quartile for college graduation is 0.01 in 1979 (and not signicant) and 0.046 (with a t-value of 1.84) in 1997. For the third quartile, the corresponding estimates are 0.034 for 1979 and 0.090 for 1997 and only the last one is signicant. For the highest income levels, we do not find any support for signicant differences across the cohorts, unlike Belley and Lochner (2007), neither at college attendance nor at college graduation.[13] However, the income estimates are largest for this income category for both cohorts and both outcomes.

4 Model

As we are interested primarily in the effects of income and AFQT scores on schooling outcomes, we do not find it necessary to estimate a multi-state dynamic programming model like Keane and Wolpin (1997). Instead, we build on ideas governing the literature on reduced-form models of schooling such as Cameron and Heckman (2001) and Ashworth et al (2017) and in which inter-temporal utilities are represented by linear (in the parameters) functions and in which unobserved heterogeneity plays a key role. Our model uses a reduced-form representation of the grade attainment process but also retains the impact of aging by modeling explicitly the time needed to achieve any given grade completion as well as grade completed by age 16 (the period at which we start modeling education). It therefore help captures potential discontinuities in the schooling accumulation process, an issue that has raised much interest in recent years.[14] However, and in order to remain as close as possible to the literature on the family income-schooling relationship, we do not distinguish between those who attend school part-time and work part-time and those who work full-time and return to school subsequently.

[13]Our estimates differ somewhat from the ones reported in Belley and Lochner (2007). For the younger cohort, the main reason for this difference is that we measure schooling as of age 26 whereas Belley and Lochner (2007) measure schooling at age 21. When we consider education at age 21 for the 1997 cohort, our estimates (available in the supplementary file, Tables S-3a and S-3b) are very similar to Belley and Lochner (2007). However, this is not the case for the older cohort where we are unable to produce estimates resembling Belley and Lochner (2007).

[14]Another option would be to build a grade transition model (essentially a hazard model) and ignore the distinction between calendar time and grade progression. However, as demonstrated in Cameron and Heckman (1998), the hazard rate representation of the education accumulation process implies a form myopic behavior which seriously undermines its credibility to approximate economic behavior. For instance, it would ignore any impact school interruptions may have on grade completion.

4.1 Choice Probabilities

We assume that the decision process starts at age 16. The choice variable is denoted $d(a)$, where $d(a) = g$ when an individual decides to invest in an additional grade attainment and where $d(a) = n$ when choosing the alternative option. We denote accumulated schooling by age a as $G(a)$ and accumulated non-school periods as $N(a)$. At each possible value of age a, each individual therefore chooses between accumulating an additional grade or involves in other activities.

To avoid the estimation of an excessively large number of parameters, we allow the parameters of the model to vary between lower schooling levels $(G(a) < 12)$ and college level $(G(a) \geq 12)$. The first level captures the effect of regressors and unobserved factors on progression within high school (until high school graduation). The second level is key as it allows us to evaluate the impact of observed and unobserved heterogeneity on the decision to enter college, conditional on high school graduation until four-year college graduation or beyond.

The school choice probabilities are defined as follows

$$
\begin{aligned}
&Pr\left(d(a) = g | G(a) < 12\right) = \\
&\qquad \Lambda\left(\beta_{0i}^{h} + \beta_{g}^{h} G(a) + \beta_{n}^{h} N(a) + P^{h}\left(AFQT_i, Inc_i; \beta_{p}^{h}\right) + \beta_{x}^{h} X_i\right) \\
&Pr\left(d(a) = g | G(a) \geq 12\right) = \\
&\qquad \Lambda\left(\beta_{0i}^{c} + \beta_{g}^{c} G(a) + \beta_{n}^{c} N(a) + P^{c}\left(AFQT_i, Inc_i; \beta_{p}^{c}\right) + \beta_{x}^{c} X_i\right)
\end{aligned}
$$

where the superscripts h and c are used to denote the set of parameters governing transitions within high school (h), and transitions into college and potentially leading to four-year college graduation (c) and where $\Lambda(.)$ denotes the logistic distribution function equal to $\exp(.)/(1 + \exp(.))$. The terms $P^h(.)$ and $P^c(.)$ are polynomials incorporating AFQT scores and family income (Inc_i) as well as their squares and an interaction, X_i is a vector of individual and family characteristics and β_g^h, β_n^h, β_p^h , β_x^h and β_g^c, β_n^c, β_p^c, β_x^c are parameters to be estimated. Finally, β_{0i}^{h} and β_{0i}^{c} are individual specific terms representing unobserved factors affecting schooling attainment. This may incorporate unobserved taste for education or any other non-cognitive element relevant to the schooling accumulation process.

The probabilities of involving in non-school activities, $\Pr\left(d\left(a\right)=n \mid N\left(a\right), G\left(a\right)<12\right)$ and $\Pr\left(d\left(a\right)=n \mid N\left(a\right), G\left(a\right)\geq 12\right)$ follow automatically from the definition of the school choice probabilities. Given our definitions of schooling and non-schooling choices, accumulated schooling by age $a+1$ is equal to

$$G\left(a+1\right)=G\left(a\right)+I\left(d\left(a\right)=g\right)$$

where $I(.)$ is the identity function.

To close the model, we assume that each individual is endowed with a vector of heterogeneity term $\left\{\beta_{0i}^{h},\beta_{0i}^{c}\right\}$ which are arbitrarily correlated. To capture the endogeneity of the initial condition, the distribution of unobserved heterogeneity is conditional on the initial grade level, $G(16)$.

As is common in the literature, we approximate its distribution by a multivariate histogram with M support points. Each type m is therefore endowed with a vector $\left\{\beta_{0m}^{h},\beta_{0m}^{c}\right\}$ with proportion $p_m(.)$ and where

$$p_m\left(G_{16}\right)=\frac{\exp\left(\tilde{p}_m+\tilde{p}_{mG}\cdot G\left(16\right)\right)}{1+\sum_{j=1}^{M-1}\exp\left(\tilde{p}_j+\tilde{p}_{jG}\cdot G\left(16\right)\right)}$$

The dependence of $p_m\left(G_{16}\right)$ on initial grade attainment (at age 16) allows us to capture the correlation between unobserved heterogeneity and grade progression realized before age 16 (Keane and Wolpin (1997)).

4.2 The Likelihood Function

We estimate the model by mixed likelihood techniques. The history of each individual i is contained in a vector

$$\left\{G_{16},\, d_i\left(a=16\right),\, d_i\left(a=17\right), ..., \, d_i\left(a=26\right)\right\}$$

and the likelihood function for observation i is equal to

$$L_i(.) = \sum_{m=1}^{M} p_m\left(G_{16}\right) \cdot \prod_{a=16}^{A} \left[\Pr\left(d_i\left(a\right) = g \mid \beta_{0m}^{h}, \beta_{0m}^{c}\right)\right]^{I(d_a=g)} \left[\Pr\left(d_i\left(a\right) = n \mid \beta_{0m}^{h}, \beta_{0m}^{c}\right)\right]^{I(d_a=n)}$$

where $I\left(.\right)$ is the identity function and A is the age at which the respondent is last observed in our sample (at most 26).

The likelihood of the sample data is formed by the product of each individual contribution $L_i(.)$. Estimates are obtained using Fortran routines.

5 Changes in the Effects of Family Income and AFQT scores

To start with, we estimated our core model on the 1979 and 1997 cohorts separately. We investigated unobserved heterogeneity specifications with two and three types and found strong support for the existence of two types but not for three. Results obtained for three types are practically identical to those obtained with only two. The importance of unobserved heterogeneity will be addressed in details below.

As for most non-linear models, the parameter estimates do not raise immediate interests. For this reason, our presentation is based mostly on various statistics obtained (or simulated) after estimating the model. The parameter estimates are found in Table A1 in a supplementary file.

Our discussion will be centered upon various estimates of the marginal effects of family income and AFQT scores on the proportion of individuals participating in higher education (those who completed grade 13 or more), and the proportion graduating from a four-year college (completing grade 16 or more). To obtain marginal effects that may be easily compared with OLS estimates found in the literature, we simulate educational choices between age 16 and 26 for a large number of individuals (10,000) reflecting the unobserved type distribution as well as random shocks. We then compute marginal effects using highest grade completed in the last period (at age 26).

To obtain a broad picture of the marginal effects of family income, we evaluate them at four different levels corresponding to the average income within the first, second, third and fourth income quartiles of each cohort. At each of these points, we compute the effect of an increase in $10,000. For the 1979 cohort, the four income benchmark levels are $17,434, $39,487, $60,269 and $99,184, respectively. For the 1997 cohort, they are $13,937, $46,726, $69,631 and $141,855. Expressed in terms the standard deviation of the income distribution, a $10,000 increase is equivalent to around 33 percent of the standard deviation for the 1979 cohort and less (about 17 percent) for the 1997 cohort.

We proceed similarly with AFQT scores and calculate the marginal effects at four different points corresponding to average AFQT scores within each quartile. To obtain an AFQT increment comparable to a $10,000 income increase, we evaluate the effects of a 10 point increase, which correspond to about 33 percent of the standard deviation in the distribution of AFQT scores for both 1979 and 1997 as this distribution has remained more or less stable between the early 1980's and the early 2000's. These four AFQT evaluation values are: 126.5 (quartile one), 162.1 (quartile two), 183.9 (quartile three) and 203.7 (quartile four) for the 1979 cohort. For the 1997 cohort, the corresponding values are 128.5, 166.0, 185.3 and 204.6.

5.1 The Marginal Effects of Family Income

The marginal effects have been computed conditional on unobserved heterogeneity (for each type) and have also been averaged over the heterogeneity distribution. To evaluate the impact of ignoring dynamic selection, we also computed the marginal effects using the parameter estimates that ignore unobserved heterogeneity. The impact of ignoring dynamic selection will however be discussed in a subsequent section, but for the moment, we ignore it. The marginal effects of a $10,000 increase in family income are displayed in the top panel of Table 4. Overall, there is overwhelming evidence that the effects of real family income on higher education attainment have decreased between the early 1980's and the early 2000's. This is true for both college participation and four-year college graduation. We now examine them in details.

5.1.1 College Attendance

The income effects for college participation for the 1979 cohort, averaged over the distribution of unobserved heterogeneity, are equal to 0.018 when evaluated at the first quartile and 0.013, 0.012 and 0.009 for quartiles two, three and four, respectively. Except for the first quartile, where the marginal effect is statistically significant, the effects are borderline significant with a t-ratio around 1.7 for the second and third quartile and not significant for the fourth quartile. For the 1997 cohort, the income effects are 0.007 (at quartile one), 0.004 (at quartile two), 0.003 (at quartile three) and 0.002 (at quartile four). The effects obtained for the 1979 cohort are therefore two to six times larger than the corresponding effects for the 1997 cohort but more importantly, none of the 1997 income effects are significantly different from zero. To fix ideas, when evaluated at the average income level, a $10,000 increase in family income would raise the population proportion enrolling in college by 1.4 percentage points in the early 1980's but only by 0.3 percentage points in the early 2000's.

In the literature, it is customary to document the evolution of income effects at different ability levels. Our model is also well suited to achieve this sort of analysis as it includes interactions between income and AFQT and allows these interactions to depend on grade level. To investigate this issue, we computed income effects at the average level of income while letting AFQT vary. As we did for income levels, we selected the average AFQT score within each specific quartile.

The results, presented in the lower panel of Table 4, reveal that the highest income effect (0.033) for the 1979 cohort is found at the lowest AFQT quartile and that the effects drop with AFQT to reach about 0.016 for quartile two, 0.014 for quartile three and 0.012 for quartile four. For the 1997 cohort, the income effects again decrease with AFQT scores (from 0.006 at first quartile to 0.001 at the top quartile) but none of these estimates are significant. This confirms the decline in income effects observed at different income levels but it also indicates that, as of the early 2000's, there was no longer any relationship between income effects and AFQT scores in determining college participation.

5.1.2 College graduation

The marginal effects for college graduation are found in the top panel of Table 5. As was the case for college participation, the income effects for college graduation have decreased substantially between the early 1980's and the early 2000's, even though income tends to have a larger impact on graduation than on participation. For the 1979 cohort, the average (over types) income effects are equal to 0.013 (quartile one), 0.021 (quartile two), 0.027 (quartile three) and 0.019 (quartile four). All of them are significant.

By the early 2000's, the income effects had dropped at all income levels. The quartile specific income effects are 0.011 (quartile one), 0.007 (quartile two), 0.008 (quartile three) and 0.005 (quartile four). When evaluated at average income, it is also clear that income effects have been substantially reduced as the 1997 cohort average effect (0.008) is three times smaller than the 1979 effect (0.025). However, and except among those in the first income quartile, none of the estimates for the 1997 cohort are significantly different from zero.

When evaluated at different AFQT levels, the 1979 college graduation income effects disclose an inverse U shape. They are low at the bottom and at the top of the AFQT distribution (0.0002 and 0.006 respectively) and higher at the second (0.021) and third (0.027) quartiles. However, by the early 2000's, the relationship between income effects and AFQT scores had become flat except for the first two AFQT quartiles for which the effects are around 0.01 (but only significant for the first quartile). For the rest of the AFQT distribution, income effects are insignificant.

To summarize, our estimates essentially indicate that between the early 1980's and the early 2000's, the effects of family income on college participation and four-year college graduation have practically vanished. Our results also indicate that the decrease in income effects depends neither on AFQT scores nor on family income level.

5.2 The Marginal Effects of AFQT scores

We now turn to the marginal effects of AFQT scores that are presented in Table 6. First, it is clear that the effects of a 10 point increase (which corresponds to an increase of one third of a standard deviation) are higher than the effects of a $10,000 increase. For the early 1980's, the quartile specific marginal effects on college participation (shown in the top panel of Table 6) are equal to 0.088 (quartile one), 0.046 (quartile two), 0.062 (quartile three) and 0.075 (quartile four). When evaluated at the average AFQT score, the estimate is equal to 0.048. For the early 2000's, the corresponding marginal effects were equal to 0.029, 0.020, 0.021, and 0.028, respectively for the quartiles and 0.022 when evaluated at the average AFQT. Similar to the findings for income effects, the effects of AFQT scores on college participation are substantially lower for the younger cohort.

The effects of AFQT on college graduation is however not as uniform as the corresponding effects of family income. When evaluated at the average AFQT scores, the marginal effects on college graduation have decreased from 0.103 in the early 1980's to 0.075 in the early 2000's. However, the overall decrease coexists with a significant increase at the first quartile (from 0.001 to 0.037).

To summarize, the marginal effects of AFQT on both college participation and college graduation have declined substantially. However, one key difference between the effects of income and of AFQT is that the latter have remained statistically significant.

At this stage, it seems that over the period going from the early 1980's until the mid 2000's, higher education has become less selective. In particular, both family income and cognitive abilities (as measured by AFQT scores), which played key roles in the 1970's and 1980's, have gradually become less important.[15] These results raise an obvious question; what happened to the

[15] This thesis is also advanced in Hoxby (2009).

role of unobserved heterogeneity over that period? This will be adressed in Section 6.

6 Changes in Educational Selectivity

In order to comprehend the sources of changes in the marginal effects of family income and AFQT scores, it is necessary to evaluate the importance of unobserved heterogeneity within each cohort and to quantify its evolution.

First, and as documented in Table 7, there are large differences between type one individuals (low schooling) and type two individuals (high schooling type) in both cohorts. About 85 percent of type two individuals are predicted to attend college in the 1979 cohort while only this proportion is only 25 percent for type one individuals. For the 1997 cohort, 15 percent of type one individuals attend college while 96 percent of type two do so. The differences in college graduation across types are even more significant (0.2 percent vs. 44 percent in the 1979 cohort, and 0.1 percent vs. 60 percent for the 1997 cohort). It is also interesting to see that the proportion of low schooling type, which was equal to 51 percent in the early 1980's, has dropped to 34 percent in the early 2000's. These estimates point toward an increasing importance of unobserved heterogeneity.

A different way to gauge the relative importance of unobserved versus observed heterogeneity within cohort (across grade level) or across cohort (for a given grade level) is to compare the standard deviations of $X_i\beta_g$ and $\theta_{i,g}$ for each cohort-level combinations. Those are also presented in Table 7.

The results suggest that at lower schooling levels (in high school), schooling accumulation is primarily explained by observed characteristics and unobserved heterogeneity plays a minor role as its standard deviation (1.4) is half of the standard deviation of observed characteristics in 1979 (2.7) and one third in 1997 (1.2 over 3.3).

However, the importance of unobserved heterogeneity increases substantially after high school completion. This is particularly true in the early 2000's, where the standard deviation of the unobserved heterogeneity term

(1.5) practically matches that of the observed heterogeneity index (1.7). The gain in relative importance of unobserved heterogeneity at college level is essentially explained by a decrease in the explanatory power of individual and family characteristics as indicated by the drop in the standard deviation from 2.1 (in the 1979 cohort) to 1.7 (in the early 2000's).

To further illustrate the changing roles of observed and unobserved heterogeneity, we also ran separate regressions of simulated schooling on individual, observed characteristics as well as on unobserved heterogeneity. This allows to obtain a ranking of the relative importance of each determinant, and in particular, to see which characteristics have become more or less important over time. The results are reported in Table 8.

When the relative importance of each observed characteristic is evaluated on a one-by-one basis, the dominant role of unoberved heterogeneity becomes even more striking. In terms of college attendance, unobserved heterogeneity was twice as important as AFQT scores in the early 1980's and about six times as important in the early 2000's. It was also five times more important than family income in the early 1980's and about 30 times more important in the early 2000's. The third most important component was mothers' education in both cohorts but its impact on college participation has decreased substantially.

Unobserved heterogeneity also dominates all regressors explaining four-year college graduation. AFQT scores remain the second most important determinant either in the early 1980's and in the early 2000's. One striking difference with college participation is that the relative importance of mother's education (the 3rd most important determinant) has remained stable or increased sligthly over that period and that the family composition indicator recording family stability has become the 4th most important determinant (after mother's education but before family income).

Despite the increase in income dispersion taking place over that period, the relative contribution of family income has therefore dropped for both college participation and college completion. This is explained by the decrease in the effect of real income (already documented in Section 5) which is sufficiently strong to annihilate the impact of an increase in income dispersion.

There are two important observations to be made. First, both in the early 1980's and the early 2000's, unobserved taste (or cost) for education was by far the most important determinant of the decision to enter college (conditional on high school completion) since no single individual attribute can match its importance.

Secondly, the importance of unobserved heterogeneity has increased in relative terms compared to individual and family characteristics such as family income, AFQT scores and other attributes. Indeed, in the early 2000's, unobserved heterogeneity had become virtually as important as the entire set of individual and family characteristics for transitions from high school to college.

To summarize, our findings suggest that classical educational selectivity based on cognitive abilities and/or family income is being gradually replaced by different forms of selectivity that are based on unmeasured abilities, preferences or costs.

6.1 Impact of Ignoring Dynamic Selection

Evaluating the implication of ignoring unobserved taste for schooling is a crucial issue. The consequences of ignoring dynamic selection on estimates of the effect of income on college participation and college graduation are however difficult to anticipate as the bias depends not only on the correlation between unobserved taste for education and income induced by dynamic selection but also on the covariances between family income and other characteristics. The former is complicated by the fact that our model allows for differentiated effect of unobserved heterogeneity by schooling level while the latter is obscured by the possibility that any given observed characteristic may affect schooling differently at lower and higher education levels.

To evaluate the impact of ignoring dynamic selection, we compare marginal effect estimates of income and AFQT scores obtained with and without unobserved heterogeneity. To simplify our analysis, we compare marginal effects obtained at the average income level (in Tables 4 and 5) and average AFQT (in Table 6).

With respect to college participation, ignoring unobserved heterogeneity primarily translates into an over-estimation of the 1979 income effect as the estimate, equal to 0.023, is twice as large as the estimate with unobserved heterogeneity which is equal 0.014 and borderline insignificant. For the 1997 cohort, estimates with and without unobserved heterogeneity indicate that the effect of income is practically zero. This obviously translates into a much stronger decrease in income effects when ignoring dynamic selection than when it is accounted for.

For college graduation, the effect of ignoring unobserved heterogeneity is opposite. Without it, the model discloses stable marginal effects (0.014 in 1979 and 0.013 in 1997) when indeed, estimates that account for dynamic selection have indicated a serious drop in marginal effects from 0.025 to practically zero.

Ignoring unobserved heterogeneity also has implications for the evolution of AFQT effects on college participation and college graduation. First, our findings suggest that ignoring dynamic selection would lead us to exaggerate the decrease in the impact of AFQT scores on college participation. Without unobserved heterogeneity, one would infer that the marginal effect of AFQT has been divided by four as it drops from 0.094 (in the early 1980's) to 0.025 (in the early 2000's).In the model with unobserved heterogeneity, the AFQT effects dropped from 0.048 to 0.022. This exageration is mostly explained by the very high marginal effect obtained in the 1979 cohort (0.094) when ignoring unobserved heterogeneity.

With respect to four-year college graduation, ignoring unobserved heterogeneity leads to even more misleading conclusions. Without it, one would conclude in favor of an increase in the effect of AFQT as the 1979 marginal effect (0.069) is much below its 1997 counterpart (0.113). With unobserved heterogeneity, and as already noted, the AFQT effect dropped from 0.103 (in 1979) to 0.075 (in 1997).

To conclude this section, and both for college participation and college graduation, ignoring educational selectivity driven by unobservables provides distorted measures of the evolution of the effects of real family income and

AFQT scores on higher education outcomes.

7 Robustness Checks and Alternative Modeling

One criticism that may be adressed to our approach is that it ignores the effect that parental income might have had on unobserved heterogeneity if the latter is influenced by parental investments taking place during childhood or early teenage (before age 16). In turn, if the relationship between income and unobserved tastes and/or abilities has changed across cohorts, our estimates may turn out to be misleading.

To motivate what follows, we first make use of AFQT scores which act as observed measures of cognitive abilities and evaluate the evolution of the correlation between family income and cognitive skills.

To do so, we regressed AFQT scores on parental income and the same background variables used in previous sections. The results, available upon request, point to two important facts. First, and in line with the existence of a positive correlation between income and parental investments, there is a positive effect of income on AFQT scores as the 1979 cohort marginal effect of income on AFQT is equal to 0.13 and its 1997 cohort pendant is 0.05. However, a second important feature is that the impact of family income on AFQT scores decreased substantially as the effect obtained for the 1997 cohort is between two and three times smaller than its 1979 counterpart. This change over time is consistent with findings in Nielsen (2015).

Although it is not certain that these features of AFQT scores may be extrapolated to unobserved heterogeneity, it seems relevant to evaluate the evolution of income effects under alternative assumptions because neither AFQT scores nor other parental background variables are likely to capture all dimensions in which parental income may matter. Note that our model allows for the unobserved type distribution to depend on initial schooling. Although it does not force orthogonality between unobserved heterogeneity and family income, it is only admitted indirectly through the correlation between initial schooling (at age 16) and family income. For this reason, we now discuss two alternative modeling strategies.

7.1 Alternative 1 - Income Effect in Type probabilities

In order to evaluate the robustness of our result, we first propose the following approach. We retain the effect that income may have on transitions from high school to college but impose a null effect of family income on schooling accumulation before high school completion (after conditioning on unobserved heterogeneity). At the same time, we allow the distribution of unobserved heterogeneity (essentially the type probabilities) to depend on family income. The null effect restriction at high school level along with the existence of persistent unobserved heterogeneity (which depends on family income) allows us to separate the income effect at college level from the correlation between income and persistent unobserved heterogeneity. The assumption of a null effect of income during high school transitions may be motivated by the very low level of financial resources needed to graduate from high school.[16]

In terms of our econometric model introduced in Section 4, this boils down to specifying the type probabilities as follows:

$$p_m\left(G_{16},\, Inc\right) = \frac{\exp\left(\tilde{p}_m + \tilde{p}_{mG}\cdot G\left(16\right) + \tilde{p}_{mI}\cdot Inc\right)}{1+\sum_{j=1}^{M-1}\exp\left(\tilde{p}_j + \tilde{p}_{jG}\cdot G\left(16\right) + \tilde{p}_{mI}\cdot Inc\right)}$$

This approach allows us to achieve two main things. First, we can evaluate the evolution of the correlation between unobserved heterogeneity and family income across cohorts. Second, and most importantly, it is possible to evaluate marginal effects of income while holding unobserved heterogeneity constant even if the latter is correlated with income.

To help comparison with the results obtained when regressing AFQT scores on family income, we also report the correlations between unobserved heterogeneity and income for each cohort.

There are two main results to retain. First, as described in the top part

[16] The null effect assumption may also be motivated by the fact that in our core model, the marginal effects of income on the conditional grade progression probabilities are virtually equal to 0 until high school graduation

of Table 11, after removing the correlation between unobserved heterogeneity and family income that may exist at age 16, both participation and graduation income effects decrease and they are practically equal to zero for both outcomes and for both cohorts.

Secondly, and as expected, persistent unobserved heterogeneity is correlated positively with family income in both the early 1980's and the early 2000's. Specifically, for this model specification, the correlation equals 0.279 for the older cohort and 0.257 for the younger cohort. This is coherent with the hypothesis that abilities and tastes other than cognitive abilities (AFQT scores) are affected by parental income.

7.2 Alternative 2 - Extrapolating the effect of Income on AFQT to Unobserved Heterogeneity

The second approach is based on the observed correlation between AFQT scores and family income. We first divided the population into two groups: those above the mean AFQT score and those below. We then estimated the probability of being of the high AFQT type as a function of initial education (at age 16), income and mother's education. We then retained the parameters and used them in the expression of the probability of being a type one (low schooling) individual when re-estimating our core model. This approach therefore allows us to estimate the effects of income on higher education without assuming a null effect at lower transitions while still imposing a non-zero correlation between income and unobserved heterogeneity.

For the 1979 cohort, the type one probability is given by

$$p_m\left(G_{16},\, Inc\right) = \frac{\exp\left(10.7 - 0.8 \cdot G_{16} - 0.13 \cdot Inc - 0.27 \cdot MED\right)}{1 + \exp\left(10.7 - 0.8 \cdot G_{16} - 0.13 \cdot Inc - 0.27 \cdot MED\right)}$$

and for the 1997 cohort, it is given by

$$p_1 \left(G_{16},\, Inc, MED \right) = \frac{\exp \left(11.3 - 0.9 \cdot G_{16} - 0.03 \cdot Inc - 0.23 \cdot MED \right)}{1 + \exp \left(11.3 - 0.9 \cdot G_{16} - 0.03 \cdot Inc - 0.23 \cdot MED \right)}$$

The results, also reported in Table 11, are quite coherent with those reported with our first approach. First, the correlation between income and unobserved heterogeneity is positive in both cohorts and the 1979 correlation (equal to 0.225) is much larger than its 1997 counterpart (0.127).

Second, as was the case with our first approach, the participation income effects measured in the 1979 and 1997 cohorts (0.004 and 0.002 respectively) are not significantly different from zero. Thirdly, and in line with results presented in Section 5, the college graduation income effect seem to have vanished. The 1979 cohort effect, equal to 0.013 and statistically significant, should be compared to the 1997 effect which is equal to 0.003 and not significant.

To summarize, whether or not unobserved heterogeneity driving educational selectivity is allowed to be correlated with family income, we find no evidence suggesting that family income has become more important. As is the case with observed AFQT scores, our results suggest that unobserved heterogeneity is positively correlated within each cohort and is therefore in line with standard views regarding parental investments and parental income. However, this positive correlation appears to have decreased substantially.

8 Summary and Conclusions

We find that the effects of family income on both college participation and college graduation have practically vanished over the period going from the early 1980's to the mid 2000's. The reason for this change is that the decrease in the effect of real income has been sufficiently strong to annihilate the impact of the increase in income dispersion observed over this period. Further, the role of AFQT scores has also been decreasing unambiguously. Both in the early 1980's and in the early 2000's, unobserved taste for education was by far the most important determinant of the decision to enter higher education but its importance has increased in relative terms compared to individual and family characteristics. Indeed, by the early 2000's, unobserved heterogeneity was as important as all the other included individual and family characteristics combined for transitions from high school to college.

Because dynamic selection induces a strong composition effect on the subpopulation of college graduates, the consequences of ignoring dynamic selection appear to be serious when making inference about the effect of individual characteristics on both college participation and college graduation. Ignoring it would lead to an exageration of the decrease in the importance of income on college participation and a serious under-estimation of the decrease in college graduation income effects.

Finally, it is interesting to note that the vanishing impacts of family income on higher education participation and on four-year college graduation are robust to the allowance of a correlation between family income and unobserved heterogeneity. Our results suggest that unobserved heterogeneity is positively correlated with family income within each cohort, just as AFQT scores are also correlated with income, but that both correlations have decreased substantially.

To appreciate our main results, it is useful to put them in perspective with the existing literature. As noted earlier, this is not necessarily easy to do as the vast majority of papers have concentrated on measuring gaps in enrollments between children of different family income quartiles. By doing so, Belley and Lochner (2007), Bailey and Dynarski (2011) and others ignore the distinction between income effects induced by change in income distribution and changes in the marginal effect of income potentially due to increases (or decreases) in the total cost of education. They conclude in favor of an increasing gap in education between top and bottom quartiles of the distribution but as already noticed, they ignore the role of unobserved heterogeneity.

From our perspective, the parameter of interest is the effect of real income since it is a natural proxy of the capacity of any family to devote resources to education. Our main findings are therefore not necessarily incompatible with increasing education gaps across income quantiles. However, it is also interesting to note that Kinsler and Pavan (2011), who investigate gaps in college quality between different quartiles, conclude that the effects of family income on college quality have been stable for average ability students but have decreased for the more able. Chetty et al (2014) also document that gaps between low and high income families have been relatively stable and even dropped more recently. Despite all this, our paper remains unique in that it documents changes in real income effects, as opposed to relative income measures, and discloses an increase in the relative importance of educational selectivity.

Before closing, it should be noted that our main results are coherent with the evolution of some key variables affecting higher education decisions. While there has been an increase in the sticker price of four year colleges, there has also been a spectacular raise in state and federal financial aid programs such as documented in Page and Scott-Clayton (2015) and Dynarski and Scott-Clayton (2013). As documented in Belley et al (2014), many of these programs have been designed in order to reduce the costs of college for low income people. On top of this, tax regulations have contributed to reduce the cost of college ceteris paribus, and the net price of two-year colleges is close to zero.[17] For example, this may have induced many individuals to initiate their college education at community colleges and subsequently transfer to regular four year colleges.[18] When coupled with the increased capacity at lower quality institutions (Hoxby (2009)), all of these may explain the decreasing impact of real family income on educational attainment.

Finally, to close our discussion, one should note that we have focussed on educational outcomes measured as of age 26. Earlier papers such as Belley and Lochner (2007) used information up to age 21 and also ignored four-year college graduation. By measuring outcomes beyond age 25, our analysis incorporates individuals who have discontinuous college attendance patterns and individuals who attend college part-time and therefore graduate beyond age 21

[17] See Abel and Deitz (2014) for a descriptive analysis of the evolution of the cost of higher education

[18] This hypothesis is examined both theoretically and empirically in Trachter (2015) and in Lovenheim and Reynolds (2011).

or 22. One avenue for future research would be to evaluate to what extent discontinuous higher education trajectories may depend on family income. This would however require a finer analysis that would distinguish between different combinantions of school attendance and labor supply. This issue is adressed in Ashworth et al (2017).

Author Affiliations

Christian Belzil: CREST, CNRS, Ecole Polytechnique, IZA and CIRANO
Jorgen Hansen: Concordia University, CIRANO, CIREQ and IZA

References

[1] Abel, Jaison.R. and Richard Deitz, "Do the Benefits of College Still Outweigh the Costs?," Current Issues in Economics and Finance, New York Federal Reserve Board, 20 (2014), 1-11.

[2] Abbot, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante, "Education Policy and Intergenerational Transfers in Equilibrium," NBER Working Paper No. w18782, 2016.

[3] Altonji, Joseph G., Prashant Bharadwaj, and Fabian Lange, "Changes in the Characteristics of American Youth: Implications for Adult Outcomes," Journal of Labor Economics, 30 (2012), 783-828.

[4] Ashworth, Jared, V. Joseph Hotz, Arnaud Maurel, and Tyler Ransom, "Changes across Cohorts in Wage Returns to Schooling and Early Work Experiences," Working paper, Duke University, 2017.

[5] Avery, Christopher and Thomas .J. Kane, "Student Perceptions of College Opportunities: The Boston COACH Program," in College choices: The Economics of where to go, when to go, and how to pay for it, Caroline M. Hoxby, ed. (Chicago, IL: University of Chicago Press, 2004), 355-394.

[6] Bailey, Martha J. and Susan M. Dynarski, "Inequality in postsecondary education," in Whither Opportunity?: Rising Inequality, Schools, and Children's Life Chances, Greg J. Duncan and Richard J. Murnane, eds. (New York: Russell Sage, 2011), 117-132.

[7] Beaudry, Paul., David A. Green, and Benjamin M. Sand, "The Great Reversal in the Demand for Skill and Cognitive Tasks," NBER Working Paper No. w18901, 2013.

[8] Belley, Philippe and Lance Lochner, "The Changing Role of Family Income and Ability in Determining Educational Achievement," Journal of Human Capital, 1 (2007), 37-89.

[9] Belley, Philippe, Marc Frenette, and Lance Lochner, "Post-secondary Attendance by Parental Income in the U.S. and Canada: Do Financial Aid Policies Explain the Differences?," Canadian Journal of Economics, 47 (2014), 664-696.

[10] Belzil, Christian, "Testing the Specication of the Mincer Wage Equation," Annals of Economics and Statistics, 92 (2008), 427 - 451.

[11] Bils, Mark and Peter .J. Klenow, "Does Schooling Cause Growth?," American Economic Review, 90 (2000), 1160-1183.

[12] Cameron Stephen V. and James .J. Heckman, "Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males," Journal of Political Economy, 106 (1998), 262-333.

[13] Cameron Stephen V. and James .J. Heckman, "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," Journal of Political Economy, 109 (2001), 455-499.

[14] Cameron Sthephen V. and Christopher Taber, "Estimation of Educational Borrowing Constraints using Returns to Schooling," Journal of Political Economy, 112 (2004), 132-182.

[15] Carneiro Pedro and James .J. Heckman, "The Evidence on Credit Constraints in Post-Secondary Schooling," Economic Journal, 112 (2002), 705-734.

[16] Castro, Rui and Daniele Coen-Pirani, "Explaining the Evolution of Educational Attainment in the United Statesy," American Economic Journal: Macroeconomics, 8 (2016), 77-112.

[17] Castex, Gonzalo and Evgenia K. Dechter, "The Changing Role of Education and Ability in Wage Determination," Journal of Labor Economics, 32 (2014), 685-710.

[18] Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner, "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility," American Economic Review 104 (2014), 141-147.

[19] Dynarski, Susan M. and Judith Scott-Clayton, "Financial Aid Policy: Lessons from Research," Future of Children, 23 (2013), 67-91.

[20] Hai, Rong and James .J. Heckman, "Inequality in Human Capital and Endogenous Credit Constraints," Review of Economic Dynamics, 25 (2017), 4-36.

[21] Heckman James J. and Stefano Mosso, "The Economics of Human Development and Social Mobility," Annual Review of Economics, 6 (2014), 689-733.

[22] Hoxby, Caroline, "The Changing Selectivity of American Colleges," Journal of Economic Perspective, 23 (2009), 95-118.

[23] Johnson, Matthew T., "Borrowing Constraints, College Enrollment, and Delayed Entry," Journal of Labor Economics, 31 (2013), 669-725.

[24] Keane, Michael P. and Kenneth I. Wolpin, "The Career Decisions of Young Men," Journal of Political Economy, 105 (1997), 473-522.

[25] Kinsler, Josh and Ronni Pavan, "Family Income and Higher Education Choices: The Importance of Accounting for College Quality," Journal of Human Capital, 5 (2011), 453-477.

[26] Lovenheim, Michael and C. Lockwood Reynolds, "Changes in Postsecondary Choices by Ability and Income: Evidence from the National Longitudinal Surveys of Youth," Journal of Human Capital, 5 (2011), 70-109.

[27] Nielsen, Eric R., "The Income-Achievement Gap and Adult Outcome Inequality," Finance and Economics Discussion Series 2015-041. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.041, 2015.

[28] Page, Lindsay C. and Judith Scott-Clayton, "Improving College Access in the United States: Barriers and Policy Responses," NBER Working Paper No. w21781, 2015.

[29] Trachter, Nicholas, "Stepping Stone and Option Value in a Model of Postsecondary Education," Quantitative Economics, 6 (2015), 223-256.

Table 1: Summary Statistics in 1979 and 1997

	1979 cohort	1997 cohort
	Parental Income	
Average	$53,969	$66,491
Std dev	$33,210	$59,500
Average among college attendants	$58,662	$66,928
Average among college graduates	$73,961	$86,231
	AFQT	
Average	168.7	170.8
Std dev	30.3	29.9
Average among college entrants	177.3	174.3
Average among college graduates	196.1	188.8
Male	0.492	0.498
Mother's education	11.7	13.2
Intact Family	0.786	0.583
Urban	0.762	0.734
Number of Siblings	3.1	2.2
Black	0.116	0.135
Hispanics	0.079	0.107
Mother's age at birth	26.3	25.8
Sample size	2,180	2,259

Table 2: Educational attainment by income and AFQT quartiles in 1979 and 1997.

	Income Distribution 1979	Income Distribution 1979	Income Distribution 1997		
	1979 cohort	1997 cohort	1997 cohort	1979 cohort	1997 cohort
	Family Income			AFQT	
Quartile	College Attendance				
First	0.279	0.422	0.421	0.14	0.27
Second	0.383	0.506	0.511	0.348	0.541
Third	0.496	0.653	0.672	0.537	0.714
Fourth	0.673	0.746	0.78	0.815	0.868
	College Graduation				
First	0.079	0.164	0.163	0.016	0.062
Second	0.16	0.257	0.26	0.088	0.25
Third	0.237	0.372	0.385	0.235	0.402
Fourth	0.409	0.466	0.505	0.549	0.609

Note:
1979 income: $28,884 (25%); $49,994 (50%); $71,210 (75%)
1997 income: $28,417 (25%); $53,760 (50%); $85,252 (75%)
1979 AFQT: 148.1 (25%); 173.7 (50%); 193.7 (75%)
1997 AFQT: 152.4 (25%); 175.7 (50%); 193.0 (75%)

Table 3a: Marginal Effects of Observed Characteristics on Schooling Attainment Obtained from OLS

	1979 cohort		1997 cohort	
Variable	**Estimate**	**T-statistic**	**Estimate**	**T-statistic**
		College Attendance		
Mother's education	0.030	7.20	0.024	6.11
Intact Family	0.010	0.41	0.132	6.76
Rural	0.012	0.57	-0.021	-1.04
Number of Siblings	-0.020	-4.27	0.006	0.73
AFQT	0.074	20.55	0.064	19.26
Family income	0.018	5.94	0.006	3.70
Black	0.250	7.96	0.158	5.72
Hispanic	0.237	6.57	0.031	1.02
Mother's age at birth	0.005	3.34	0.005	2.63
Male	-0.008	-0.48	-0.106	-6.01
		College Graduation		
Mother's education	0.023	6.26	0.024	6.39
Intact Family	0.007	0.33	0.129	6.77
Rural	-0.015	-0.82	-0.031	-1.55
Number of Siblings	-0.011	-2.65	0.017	2.02
AFQT	0.051	16.47	0.051	15.54
Family income	0.017	6.19	0.006	3.73
Black	0.106	3.89	0.096	3.55
Hispanic	0.108	3.46	-0.006	-0.19
Mother's age at birth	0.005	3.67	0.007	3.91
Male	0.016	1.06	-0.088	-5.13

Note: Income is measured in $10,000s and adjusted using CPI-U.

Table 3b: Marginal Effects of Observed Characteristics on Schooling Attainment Obtained from OLS

Variable	1979 cohort Estimate	1979 cohort T-statistic	1997 cohort Estimate	1997 cohort T-statistic
		College Attendance		
Mother's education	0.034	8.14	0.024	6.23
Intact Family	0.023	0.96	0.113	5.65
Rural	0.018	0.86	-0.026	-1.29
Number of Siblings	-0.023	-4.87	0.006	0.73
AFQT - Q2	0.168	6.22	0.218	8.38
AFQT - Q3	0.325	11.61	0.364	13.48
AFQT - Q4	0.568	19.59	0.489	17.46
Family income - Q2	0.023	0.88	0.032	1.28
Family income - Q3	0.068	2.43	0.108	4.00
Family income - Q4	0.154	5.31	0.130	4.64
Black	0.200	6.25	0.144	5.14
Hispanic	0.223	6.19	0.037	1.23
Mother's age at birth	0.005	2.50	0.005	2.50
Male	-0.026	-1.44	-0.112	-6.22
		College Graduation		
Mother's education	0.026	7.18	0.024	6.35
Intact Family	0.023	1.15	0.112	5.60
Rural	-0.007	-0.39	-0.034	-1.72
Number of Siblings	-0.013	-3.25	0.016	1.90
AFQT - Q2	0.020	0.91	0.127	5.08
AFQT - Q3	0.131	5.46	0.250	9.62
AFQT - Q4	0.411	16.44	0.422	15.63
Family income - Q2	0.010	0.45	0.046	1.84
Family income - Q3	0.034	1.48	0.090	3.46
Family income - Q4	0.129	5.16	0.125	4.46
Black	0.046	1.70	0.087	3.22
Hispanic	0.091	2.94	0.003	0.10
Mother's age at birth	0.005	5.00	0.007	3.73
Male	-0.007	-0.47	-0.097	-5.63

Table 4: Marginal Effects of Family Income on College Attendance

	1979 cohort		1997 cohort	
Evaluated at:	**M. E.**	**T-statistic**	**M. E.**	**T-statistic**
Mean income				
in Q1	0.018	2.59	0.007	1.02
in Q2	0.013	1.84	0.004	0.52
in Q3	0.012	1.68	0.003	0.39
in Q4	0.009	1.31	0.002	0.26
Overall	0.014	1.92	0.003	0.38
Mean income				
and at AFQT - Q1	0.033	5.37	0.006	0.92
and at AFQT - Q2	0.016	1.93	0.003	0.49
and at AFQT - Q3	0.014	2.03	0.002	0.22
and at AFQT - Q4	0.012	2.03	0.001	0.20

Note: In each case, income is increased with $10,000.
The marginal effects are evaluated at the average values of observed characteristics presented in Table 1.

Table 5: Marginal Effects of Family Income on College Graduation

	1979 cohort		1997 cohort	
Evaluated at:	**M. E.**	**T-statistic**	**M. E.**	**T-statistic**
Mean income				
in Q1	0.013	4.74	0.011	1.83
in Q2	0.021	5.46	0.007	1.16
in Q3	0.027	5.66	0.008	1.30
in Q4	0.019	3.21	0.005	0.68
Overall	0.025	5.52	0.008	1.30
Mean income				
and at AFQT - Q1	0.0002	0.58	0.009	2.40
and at AFQT - Q2	0.021	5.92	0.010	1.54
and at AFQT - Q3	0.027	4.24	0.005	0.66
and at AFQT - Q4	0.006	0.83	0.0003	0.04

Note: In each case, income is increased with $10,000.
The marginal effects are evaluated at the average values of observed characteristics presented in Table 1.

Table 6: Marginal Effects of AFQT

	1979 cohort		1997 cohort	
Evaluated at:	**M. E.**	**T-statistic**	**M. E.**	**T-statistic**
Mean AFQT		**College Attendance**		
in Q1	0.088	13.77	0.029	4.05
in Q2	0.046	6.59	0.020	2.83
in Q3	0.062	9.23	0.021	3.04
in Q4	0.075	13.08	0.028	4.31
Overall	0.048	6.83	0.022	3.22
Mean AFQT		**College Graduation**		
in Q1	0.001	2.86	0.037	9.42
in Q2	0.079	19.40	0.075	11.50
in Q3	0.095	14.62	0.056	7.94
in Q4	0.024	3.38	0.018	2.52
Overall	0.103	20.43	0.075	11.20

Note: In each case, the AFQT score is increased with 10 units.
The marginal effects are evaluated at the average values of observed characteristics presented in Table 1.

Table 7: The Explanatory Power of Observed and Unobserved Heterogeneity

Variable	1979 cohort	1997 cohort
Grades < 12		
Std dev of UH terms	1.44	1.23
Std dev of OH terms	2.73	3.35
Grades 12 or more		
Std dev of UH terms	1.60	1.52
Std dev of OH terms	2.11	1.74
Proportion Type 1	0.514	0.339
Expected schooling		
Type 1		
Attend College	0.252	0.148
College Graduate	0.002	0.001
Type 2		
Attend College	0.849	0.958
College Graduate	0.437	0.604

Note: UH represents Unobserved Heterogeneity while OH is used for Observed Heterogeneity.

Table 8: Relative Contributions to Explained Variations in Schooling Attainment

Variable	1979 cohort R-squares	1997 cohort R-squares
	College Attendance	
Mother's education	0.075	0.03
Intact Family	0.016	0.019
AFQT	0.231	0.103
Family income	0.074	0.021
Black	0.009	0.002
Hispanic	0.005	0.004
Unobserved Heterogeneity	0.358	0.680
All of the above	0.560	0.721
	College Graduation	
Mother's education	0.085	0.088
Intact Family	0.012	0.057
AFQT	0.181	0.193
Family income	0.072	0.048
Black	0.010	0.004
Hispanic	0.005	0.008
Unobserved Heterogeneity	0.282	0.339
All of the above	0.451	0.500

Note: Individual R-squares from regressions of predicted college attendance and graduation indicators.

Table 9: Marginal Effects of Family Income on College Attendance and Graduation - Model without Unobserved Heterogeneity

	1979 cohort		1997 cohort	
Evaluated at:	**M. E.**	**T-statistic**	**M. E.**	**T-statistic**
Mean income		**College Attendance**		
in Q1	0.036	5.19	0.006	1.40
in Q2	0.029	4.46	0.005	1.23
in Q3	0.024	3.86	0.003	0.95
in Q4	0.016	2.84	0.002	0.70
Overall	0.023	3.67	0.005	1.36
Mean income		**College Attendance**		
and at AFQT - Q1	0.029	4.24	0.017	2.79
and at AFQT - Q2	0.027	4.04	0.005	1.25
and at AFQT - Q3	0.015	2.99	0.001	0.51
and at AFQT - Q4	0.007	3.06	0.0001	0.08
Mean income		**College Graduation**		
in Q1	0.008	3.40	0.014	2.12
in Q2	0.009	2.89	0.013	2.02
in Q3	0.014	3.71	0.013	1.98
in Q4	0.015	3.20	0.009	1.32
Overall	0.014	4.07	0.013	1.91
Mean income		**College Graduation**		
and at AFQT - Q1	0.001	1.37	0.01	2.55
and at AFQT - Q2	0.008	2.99	0.013	2.04
and at AFQT - Q3	0.028	4.96	0.01	1.37
and at AFQT - Q4	0.041	5.91	0.004	0.76

Note: In each case, income is increased with $10,000.
The marginal effects are evaluated at the average values of observed characteristics presented in Table 1.

Table 10: Marginal Effects of AFQT scores on College Attendance and Graduation - Model without Unobserved Heterogeneity

	1979 cohort		1997 cohort	
Evaluated at:	**M. E.**	**T-statistic**	**M. E.**	**T-statistic**
Mean AFQT	**College Attendance**			
in Q1	0.079	11.56	0.06	10.12
in Q2	0.103	15.88	0.031	8.68
in Q3	0.072	16.15	0.018	8.28
in Q4	0.022	12.09	0.005	4.91
Overall	0.094	15.69	0.025	7.57
Mean AFQT	**College Graduation**			
in Q1	0.002	3.18	0.036	8.27
in Q2	0.041	12.67	0.102	15.23
in Q3	0.163	26.51	0.140	20.25
in Q4	0.231	36.06	0.106	20.46
Overall	0.069	16.94	0.113	16.34

Note: In each case, the AFQT score is increased with 10 units.
The marginal effects are evaluated at the average values of observed characteristics presented in Table 1.

Table 11: Results from Alternative Model Specifications

	1979 cohort		1997 cohort	
	Marginal Effects of Family Income			
Model specification	**M. E.**	**T-statistic**	**M. E.**	**T-statistic**
Alternative 1 - Income in type proportion				
College Attendance	0.002	0.34	-0.0001	-0.01
College Graduation	0.005	1.13	-0.0004	-0.06
Alternative 2 - Fixed proportions				
College Attendance	0.004	0.61	0.002	0.37
College Graduation	0.013	2.84	0.003	0.48

Model specification	**Correlation between income and unobserved heterogeneity**	
Alternative 1 - Income in type proportion	0.279	0.257
Alternative 2 - Fixed proportions	0.225	0.127

Note: See discussion in Section 7 for details on the two alternative model specifications. In each case, income is increased with $10,000. The marginal effects are evaluated at the average values of observed characteristics presented in Table 1.

Table S1a: Parameter Estimates for Model with two-type Unobserved Heterogeneity Distribution

	1979 cohort			1997 cohort		
Transitions - Grades up to 12	Estimate	std. err.	T-statistic	Estimate	std. err.	T-statistic
g1i_1	-0.751	0.231	-3.25	1.836	1.269	1.45
g1i_2	1.061	0.082	12.97	-4.958	1.263	-3.93
g1x1 (mothers ed)	0.086	0.025	3.47	0.072	0.016	4.40
g1x2 (nuclear)	0.443	0.120	3.70	0.461	0.072	6.43
g1x3 (urban)	0.354	0.108	3.27	0.163	0.088	1.85
g1x4 (siblings)	-0.025	0.024	-1.06	0.015	0.030	0.49
g1x5 (afqt)	0.534	0.078	6.86	-0.155	0.148	-1.05
g1x6 (income)	0.294	0.105	2.79	-0.005	0.046	-0.10
g1x7 (black)	1.124	0.166	6.75	0.463	0.089	5.19
g1x8 (hisp)	0.236	0.186	1.27	0.218	0.112	1.94
g1x9 (mothers age)	0.017	0.007	2.21	0.018	0.008	2.24
g1x10 (school periods)	-0.593	0.058	-10.19	-0.072	0.036	-2.01
g1x11 (non-school periods)	-1.130	0.045	-25.07	-1.581	0.048	-33.01
g1x12 (male)	-0.234	0.093	-2.52	-0.228	0.072	-3.18
g1x13 (afqt^2)	-0.007	0.003	-2.40	0.010	0.005	2.17
g1x14 (income^2)	-0.010	0.004	-2.83	-0.003	0.001	-4.38
g1x15 (afqt*income)	-0.004	0.006	-0.68	0.006	0.003	2.28
Transitions - Grades 13 and up						
g2i_1	-8.674	1.671	-5.19	-0.527	0.533	-0.99
g2i_2	1.163	0.039	29.90	1.060	0.042	25.49
g2x1 (mothers ed)	0.141	0.024	5.86	0.054	0.013	4.23
g2x2 (nuclear)	-0.017	0.130	-0.13	0.227	0.066	3.44
g2x3 (rural)	0.007	0.089	0.08	-0.190	0.062	-3.04
g2x4 (siblings)	-0.132	0.024	-5.60	0.035	0.035	0.99
g2x5 (afqt)	0.816	0.203	4.02	0.207	0.078	2.66
g2x6 (income)	0.431	0.099	4.35	0.092	0.036	2.58
g2x7 (black)	1.569	0.192	8.17	0.366	0.095	3.85
g2x8 (hisp)	1.064	0.201	5.29	-0.085	0.102	-0.84
g2x9 (mothers age)	0.036	0.007	5.25	0.009	0.007	1.40
g2x10 (school periods)	-0.628	0.049	-12.84	-0.442	0.023	-19.16
g2x11 (non-school periods)	-0.582	0.026	-22.51	-0.482	0.016	-30.92
g2x12 (male)	0.047	0.057	0.83	-0.244	0.057	-4.31
g2x13 (afqt^2)	-0.009	0.006	-1.55	0.000	0.003	-0.09
g2x14 (income^2)	-0.005	0.002	-2.05	-0.001	0.000	-1.10
g2x15 (afqt*income)	-0.016	0.006	-2.82	-0.004	0.002	-1.83
Type 1 probability parameters						
Constant	5.574	0.716	7.79	3.397	0.902	3.77
Age 16 education	-0.591	0.079	-7.48	-0.411	0.096	-4.29
Average Log-likelihood value		-3.25			-4.23	
Akaike Information Criteria		14,242			19,183	

Table S1b: Parameter Estimates for Model with no Unobserved Heterogeneity

	1979 cohort			1997 cohort		
Transitions - Grades up to 12	Estimate	std. err.	T-statistic	Estimate	std. err.	T-statistic
g1i_1	2.152	1.009	2.13	1.776	0.615	2.89
g1x1 (mothers ed)	0.052	0.017	3.06	0.073	0.015	4.79
g1x2 (nuclear)	0.264	0.084	3.13	0.443	0.074	5.97
g1x3 (urban)	0.319	0.092	3.48	0.170	0.062	2.72
g1x4 (siblings)	-0.043	0.018	-2.44	0.019	0.031	0.62
g1x5 (afqt)	0.116	0.122	0.95	-0.155	0.086	-1.80
g1x6 (income)	0.019	0.074	0.26	-0.004	0.038	-0.09
g1x7 (black)	0.857	0.124	6.93	0.474	0.100	4.72
g1x8 (hisp)	0.220	0.133	1.65	0.237	0.114	2.08
g1x9 (mothers age)	0.022	0.006	3.42	0.018	0.007	2.47
g1x10 (school periods)	-0.346	0.038	-9.12	-0.067	0.034	-1.99
g1x11 (non-school periods)	-1.350	0.039	-34.23	-1.590	0.045	-35.33
g1x12 (male)	-0.177	0.074	-2.40	-0.217	0.063	-3.45
g1x13 (afqt^2)	0.003	0.004	0.82	0.010	0.003	3.54
g1x14 (income^2)	-0.009	0.003	-2.70	-0.003	0.001	-4.28
g1x15 (afqt*income)	0.011	0.005	2.32	0.006	0.002	2.64
Transitions - Grades 13 and up						
g2i_1	-2.569	1.320	-1.95	-0.552	0.517	-1.07
g2x1 (mothers ed)	0.099	0.012	8.49	0.063	0.009	7.33
g2x2 (nuclear)	-0.064	0.082	-0.79	0.240	0.045	5.28
g2x3 (rural)	-0.007	0.054	-0.12	-0.129	0.043	-2.98
g2x4 (siblings)	-0.085	0.015	-5.54	0.038	0.020	1.95
g2x5 (afqt)	0.053	0.151	0.35	0.006	0.061	0.10
g2x6 (income)	0.153	0.065	2.35	0.087	0.024	3.57
g2x7 (black)	0.710	0.078	9.10	0.348	0.056	6.26
g2x8 (hisp)	0.642	0.102	6.27	0.011	0.030	0.36
g2x9 (mothers age)	0.019	0.004	4.38	0.014	0.005	3.04
g2x10 (school periods)	-0.161	0.017	-9.61	-0.144	0.012	-12.33
g2x11 (non-school periods)	-0.753	0.019	-40.19	-0.593	0.013	-46.92
g2x12 (male)	0.003	0.048	0.06	-0.244	0.040	-6.18
g2x13 (afqt^2)	0.007	0.004	1.62	0.005	0.002	2.81
g2x14 (income^2)	-0.002	0.001	-1.67	0.000	0.000	-1.19
g2x15 (afqt*income)	-0.004	0.004	-0.99	-0.004	0.001	-2.70
Average Log-likelihood value		-3.41			-4.37	
Akaike Information Criteria		14,932			19,808	

Table S1c: Parameter Estimates for Alternative Model with two-type Unobserved Heterogeneity Distribution
Alternative 1 - Income in Type Proportion

	1979 cohort			1997 cohort		
Transitions - Grades up to 12	Estimate	std. err.	T-statistic	Estimate	std. err.	T-statistic
g1i_1	-0.991	0.303	-3.27	0.758	1.226	0.62
g1i_2	1.083	0.117	9.29	0.970	0.088	11.03
g1x1 (mothers ed)	0.071	0.021	3.34	0.062	0.016	3.87
g1x2 (nuclear)	0.408	0.101	4.05	0.480	0.072	6.69
g1x3 (urban)	0.274	0.106	2.57	0.231	0.086	2.69
g1x4 (siblings)	-0.022	0.026	-0.84	0.002	0.040	0.06
g1x5 (afqt)	0.578	0.082	7.05	0.171	0.064	2.68
g1x7 (black)	0.918	0.141	6.49	0.417	0.108	3.85
g1x8 (hisp)	0.058	0.168	0.34	0.174	0.101	1.73
g1x9 (mothers age)	0.020	0.007	2.79	0.015	0.010	1.46
g1x10 (school periods)	-0.466	0.045	-10.30	-0.270	0.065	-4.18
g1x11 (non-school periods)	-1.177	0.043	-27.33	-1.334	0.050	-26.54
g1x12 (male)	-0.150	0.062	-2.44	-0.098	0.104	-0.95
g1x13 (afqt^2)	-0.010	0.003	-3.58	0.000	0.002	0.02
Transitions - Grades 13 and up						
g2i_1	-7.153	1.598	-4.48	-2.746	0.241	-11.40
g2i_2	1.234	0.036	34.57	1.126	0.052	21.85
g2x1 (mothers ed)	0.127	0.020	6.22	0.066	0.011	6.19
g2x2 (nuclear)	0.014	0.116	0.12	0.309	0.065	4.77
g2x3 (rural)	-0.074	0.078	-0.95	-0.157	0.045	-3.52
g2x4 (siblings)	-0.132	0.033	-4.02	0.048	0.021	2.24
g2x5 (afqt)	0.856	0.173	4.96	0.262	0.052	5.07
g2x6 (income)	0.110	0.103	1.06	0.050	0.022	2.33
g2x7 (black)	1.296	0.209	6.21	0.384	0.083	4.63
g2x8 (hisp)	0.978	0.231	4.24	-0.036	0.137	-0.27
g2x9 (mothers age)	0.030	0.008	3.81	0.015	0.007	2.33
g2x10 (school periods)	-0.668	0.038	-17.67	-0.361	0.020	-18.26
g2x11 (non-school periods)	-0.565	0.023	-24.35	-0.510	0.015	-34.10
g2x12 (male)	0.045	0.100	0.46	-0.236	0.050	-4.72
g2x13 (afqt^2)	-0.012	0.005	-2.52	-0.002	0.002	-1.00
g2x14 (income^2)	-0.001	0.002	-0.63	0.000	0.000	0.16
g2x15 (afqt*income)	-0.004	0.006	-0.76	-0.003	0.001	-2.40
Type 1 probability parameters						
Constant	5.641	0.742	7.60	10.281	0.809	12.71
Initial education	-0.485	0.078	-6.25	-1.059	0.088	-12.03
Income	-0.166	0.020	-8.44	-0.108	0.012	-8.80
Average Log-likelihood value		-3.24			-4.23	
Akaike Information Criteria		14,194			19,179	

Table S1d: Parameter Estimates for Alternative Model with two-type Unobserved Heterogeneity Distribution
Alternative 2 - Fixed Proportions

	1979 cohort			**1997 cohort**		
Transitions - Grades up to 12	Estimate	std. err.	T-statistic	Estimate	std. err.	T-statistic
g1i_1	0.998	0.932	1.07	0.815	1.176	0.69
g1i_2	1.134	0.100	11.37	0.894	0.069	12.90
g1x1 (mothers ed)	-0.008	0.020	-0.39	0.019	0.018	1.09
g1x2 (nuclear)	0.386	0.098	3.94	0.431	0.077	5.60
g1x3 (urban)	0.274	0.099	2.77	0.234	0.050	4.69
g1x4 (siblings)	-0.021	0.021	-1.01	0.011	0.043	0.24
g1x5 (afqt)	0.428	0.123	3.49	0.202	0.105	1.92
g1x6 (income)	0.080	0.061	1.31	0.025	0.048	0.51
g1x7 (black)	1.045	0.146	7.15	0.428	0.140	3.06
g1x8 (hisp)	0.081	0.153	0.53	0.205	0.126	1.63
g1x9 (mothers age)	0.012	0.007	1.75	0.018	0.009	1.97
g1x10 (school periods)	-0.494	0.044	-11.20	-0.254	0.042	-6.06
g1x11 (non-school periods)	-1.118	0.042	-26.81	-1.364	0.048	-28.23
g1x12 (male)	-0.183	0.084	-2.17	-0.134	0.019	-6.88
g1x13 (afqt^2)	-0.006	0.004	-1.51	-0.002	0.003	-0.50
g1x14 (income^2)	-0.012	0.004	-2.77	-0.003	0.001	-3.62
g1x15 (afqt*income)	0.008	0.005	1.58	0.004	0.003	1.15
Transitions - Grades 13 and up						
g2i_1	-6.443	1.388	-4.64	-2.307	0.012	-196.92
g2i_2	1.291	0.047	27.72	1.132	0.041	27.91
g2x1 (mothers ed)	0.048	0.015	3.14	0.031	0.011	2.84
g2x2 (nuclear)	0.125	0.086	1.46	0.310	0.060	5.19
g2x3 (rural)	-0.041	0.073	-0.57	-0.131	0.070	-1.88
g2x4 (siblings)	-0.093	0.021	-4.45	0.043	0.025	1.74
g2x5 (afqt)	0.614	0.158	3.89	0.274	0.043	6.31
g2x6 (income)	0.079	0.049	1.61	0.088	0.041	2.14
g2x7 (black)	1.073	0.147	7.29	0.377	0.052	7.27
g2x8 (hisp)	0.764	0.194	3.94	0.018	0.084	0.21
g2x9 (mothers age)	0.027	0.006	4.63	0.017	0.006	2.85
g2x10 (school periods)	-0.498	0.026	-18.98	-0.383	0.017	-22.44
g2x11 (non-school periods)	-0.633	0.021	-30.47	-0.502	0.014	-35.80
g2x12 (male)	-0.006	0.055	-0.11	-0.224	0.051	-4.42
g2x13 (afqt^2)	-0.007	0.005	-1.55	-0.002	0.002	-1.07
g2x14 (income^2)	-0.003	0.002	-1.72	0.000	0.000	-0.86
g2x15 (afqt*income)	0.000	0.003	0.09	-0.004	0.002	-1.84
Average Log-likelihood value		-3.23			-4.19	
Akaike Information Criteria		14,151			18,998	

Table S2: Details on Sample selections

Remaining individuals after:	NLSY 1979	NLSY 1997
Selecting cross-section and individuals between 14 and 17 in 1979 or 1997	2,987	4,506
Removing those with missing or no family income	2,840	3,943
Removing those with missing information on mother's age at birth	2,532	3,753
Removing those who reported family income instead of their parents		3,668
Removing those with missing on nuclear family, urban residence and AFQT	2,449	2,384
Removing those who did not participate at least 6 years in the panel	2,274	2,318
Removing those with missing information on mather's education	2,180	2,259

Table S-3a: Marginal Effects of Observed Characteristics on Schooling Attainment Obtained from OLS

	1979 cohort		1997 cohort	
Variable	**Estimate**	**T-statistic**	**Estimate**	**T-statistic**
		College Attendance		
Mother's education	0.032	7.84	0.024	6.13
Intact Family	0.009	0.37	0.140	7.14
Rural	0.019	0.89	-0.020	-1.01
Number of Siblings	-0.019	-4.12	0.002	0.28
AFQT	0.073	20.41	0.064	19.12
Family income	0.019	6.23	0.007	4.43
Black	0.269	8.62	0.166	5.98
Hispanic	0.259	7.24	-0.010	-0.33
Mother's age at birth	0.006	3.77	0.005	2.77
Male	-0.012	-0.69	-0.090	-5.08
		College Graduation		
Mother's education	0.018	5.15	0.020	5.54
Intact Family	0.010	0.51	0.104	5.76
Rural	-0.008	-0.48	-0.024	-1.29
Number of Siblings	-0.011	-2.74	0.018	2.21
AFQT	0.041	14.11	0.043	13.97
Family income	0.016	6.47	0.008	5.28
Black	0.094	3.68	0.078	3.05
Hispanic	0.086	2.93	-0.029	-1.03
Mother's age at birth	0.004	3.20	0.005	2.66
Male	0.001	0.07	-0.085	-5.24

Note: Income is measured in $10,000s and adjusted using CPI-U. Further, the AFQT scores were divided by 10. Eduational attainment is measured at age 21 or at the age during the latest interview, whichever is highest.

Table S-3b: Marginal Effects of Observed Characteristics on Schooling Attainment Obtained from OLS

	1979 cohort		1997 cohort	
Variable	**Estimate**	**T-statistic**	**Estimate**	**T-statistic**
		College Attendance		
Mother's education	0.037	8.96	0.024	6.24
Intact Family	0.023	0.98	0.118	5.83
Rural	0.025	1.17	-0.025	-1.25
Number of Siblings	-0.022	-4.69	0.002	0.20
AFQT - Q2	0.157	5.96	0.207	8.04
AFQT - Q3	0.306	10.91	0.356	13.33
AFQT - Q4	0.564	19.39	0.490	17.36
Family income - Q2	0.010	0.39	0.040	1.56
Family income - Q3	0.068	2.49	0.125	4.68
Family income - Q4	0.151	5.21	0.160	5.61
Black	0.215	6.81	0.152	5.49
Hispanic	0.243	6.74	-0.003	-0.09
Mother's age at birth	0.006	4.04	0.005	2.86
Male	-0.031	-1.76	-0.097	-5.45
		College Graduation		
Mother's education	0.021	6.34	0.020	5.48
Intact Family	0.028	1.49	0.091	4.89
Rural	-0.002	-0.12	-0.027	-1.45
Number of Siblings	-0.013	-3.33	0.016	2.02
AFQT - Q2	0.001	0.07	0.100	4.24
AFQT - Q3	0.085	3.77	0.208	8.52
AFQT - Q4	0.339	14.57	0.364	14.06
Family income - Q2	-0.004	-0.18	0.027	1.18
Family income - Q3	0.015	0.70	0.078	3.19
Family income - Q4	0.114	4.87	0.140	5.34
Black	0.037	1.47	0.069	2.72
Hispanic	0.068	2.37	-0.022	-0.77
Mother's age at birth	0.004	3.28	0.005	2.63
Male	-0.020	-1.41	-0.093	-5.73

Note: Eduational attainment is measured at age 21 or at the age during the latest interview, whichever is highest.