The Evolution of the US Family Income-Schooling Relationship and Educational Selectivity *

Christian Belzil
CREST, CNRS, Paris Polytechnic Institute and IZA and CIRANO

Jörgen Hansen
Concordia University, CIREQ, CIRANO and IZA

September 9, 2019

*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, Lance Lochner, Michael Lovenheim, Arnaud Maurel, Douglas Webber, and participants at the IAAE 2018 Conference (Montreal) and the ESEM 2019 (Manchester) for comments and exchanges. We also thank Ed Vytlacil and four anonymous referees for their insightful comments. The usual disclaimer applies.

ABSTRACT

We estimate a dynamic model of schooling on two cohorts of the NLSY and find that, contrary to conventional wisdom, the effects of family income on education have practically vanished between the early 1980's and the early 2000's. After conditioning on a cognitive ability measure (AFQT), family background variables and unobserved heterogeneity (allowed to be correlated with observed characteristics), income effects have lost between 30% and 80% of their importance on age-specific grade progression probabilities. A $70,000 family income differential raised college participation by 10 percentage points in the early 1980's. In the early 2000's, a $330,000 income differential had the same impact. The effects of AFQT scores also decreased substantially but did not vanish. Over the same period, the relative importance of unobserved heterogeneity has expanded so much that it has become by far the most important determinant of education.

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

1 Introduction

In recent years, a relatively large number of papers have studied the evolution of education inequalities between high-income and low-income classes in the US. While some have stressed the existence of an increasing gap between top and bottom quartiles (Belley and Lochner, 2007, Lochner and Monge-Naranjo, 2012, Avery and Kane, 2004, Bailey and Dynarski, 2011, and Page and Scott-Clayton, 2015), others have reported a more stable relationship (Kinsler and Pavan, 2011 and Chetty, Hendren, Kline, Saez, and Turner, 2014).[1]

In this paper, we take a different path than most of the existing literature and investigate the evolution of the effects of real family income (as opposed to relative income) on education outcomes. We show that income effects have practically vanished between the 1980's and the early 2000's and that the effects of AFQT scores have also decreased.

To do so, we estimate a dynamic model of schooling with unobserved heterogeneity on two cohorts of the National Longitudinal Survey of Youth (the 1979 cohort and the 1997 cohort) and compare the evolution of the effects of family income and Armed Forces Qualification Test (AFQT) scores on age-specific grade progression probabilities and resulting educational outcomes (years of schooling and college participation).

Our approach differs from the existing literature in three main dimensions. First, we measure the change in the effect of real income as opposed to the evolution of education gaps between high-income and low-income subpopulations. The evolution of the effect of parental income on higher education is impossible to infer solely from results emphasizing the role of relative income because the latter are potentially imputable to both a change in income distribution (creating an income effect) as well as a change in the effect of real family income which could be driven by a change in the cost or in some of its determinants.

While it is widely recognized that income growth has been more important among upper income classes, and that the sticker price of four-year colleges has also increased faster than inflation, other institutional changes affecting higher education decisions may have tempered the effects of income inequality and tuitions and thereby reduced the effect of income.

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

As documented in Page and Scott-Clayton (2015) and in Dynarski and Scott-Clayton (2013), there has been a significant raise in state and federal financial aid programs. Many of those have been designed in order to reduce the cost of college for low income people (Belley et. al., 2014). At the same time, tax regulations have contributed to drive the net price of two-year colleges close to zero (Abel and Deitz, 2014). Also, and as documented in Hoxby (2009), the substantial increase in college enrollments observed between the early 80's and the early 2000's was mostly due to increased capacity at lower quality (and lower tuition) institutions. Finally, many papers in the structural literature have stressed that education choices are largely explained by a non-pecuniary dimension (the consumption value). If its importance has expanded relatively more among lower income sub-populations than among high income families (already endowed with a high taste for education), it could substantially damp the effect of tuition increases. For all these reasons, the evolution of real family income effects remains an open question.

A second difference is that we allow income effects to vary with age within each cohort using 3 different intervals; from 16 to 18, from 19 to 21, and from 22 to 25.[2] There are good reasons for making a distinction between age-specific income effects. At age 16, individuals progress toward high school graduation and are fully dependent of their parents. At age 19, individuals are likely to still be dependent of their parents but start experiencing early college transitions that are more costly. At age 22, grade transitions are less frequent and individuals are likely to be less dependent.[3] As the degree of dependence on parental income may therefore change substantially as young individuals are aging, it is conceivable that the overall relationship between family income and schooling is driven by some age-specific dependence. On top of this, the parent-child relationship may have evolved over the period of time elapsed between the 79 and the 97 cohorts, and it is particularly important to know if changes in income effect across cohorts have been spread uniformly between late teenage years and early adulthood.

Finally, a third difference which turns out to be fundamental, is that we control for dynamic selection within each cohort and therefore separate changes in the effects of real income, AFQT and family characteristics from

[2]In our terminology, and to take an example, a transition at age 16 refers to a grade change (or lack thereof) from 16 to 17. In our data, the last transition is from 25 to 26.

[3]Note that according to US federal financial aid regulations (FAFSA), an individual becomes automatically independent from the parents only at age 24.

potential movements in the distribution of unobservables. Note that these last two dimensions of our analysis (the effect of aging on grade transitions and the incorporation of unobserved heterogeneity) may not be tackled with standard linear regression models.

Throughout the paper, we use the term "unobserved heterogeneity" to designate any unmeasured factor such as taste for schooling, monetary or non-monetary costs of education or ability and motivation, which remains significant after controlling for AFQT scores, non-cognitive skill measures such as the Rotter Locus of Control (measuring the internal control of an individual) and the Rosenberg Self-Esteem Scale (measuring the degree of approval toward oneself), and other characteristics. Although those non-cognitive skill measures are observed in the NLSY79 only, we develop an approach that makes use of a predicted non-cognitive skill component using individual and family characteristics observed in both cohorts. We then allow for unobserved heterogeneity (unobserved type probabilities) to depend on predicted non-cognitive skill measures. These measures being themselves correlated with family background variables, family income and AFQT's, we need not assume orthogonality between unobserved heterogeneity and measured characteristics such as income, AFQT scores and other background variables.

Documenting the importance of dynamic selection is neither a technicality nor a fantasy. First, if choices are truly sequential and unobserved heterogeneity is important within a given cohort, ignoring it is likely to bias income and AFQT effects measured within a given cohort. On top of this, if the relative importance of unobserved heterogeneity changes across cohorts, any comparison between marginal effects of income obtained from OLS estimates across different cohorts becomes completely uninformative (even if income is an exogenous regressor).[4]

The incidence of changes in educational selectivity has also attracted attention among those involved in education policies. According to the National Center for Education statistics, the yearly flow of individuals graduating with a Bachelor's degree has fluctuated between 1.5 and 2.0 million per year between 2005 and 2018. These numbers are far in excess of those for

[4]The notion of dynamic selection plays a key role in many areas of economics and econometrics. Its statistical implications are discussed in Lancaster (1990) in the context of duration and transition data analysis and in Cameron and Heckman (1998) within an optimal schooling model. We provide an intuitive illustration in Section 2.

the early 80's.[5] Because this large increase in enrollments has been observed mostly at lower quality institutions, this has most likely driven important changes in the educational selection process and contributed to a reduction in higher education selectivity. It is therefore important to take this into account when measuring the evolution of income effects.

Following the existing literature, we adopt a nonstructural (or a semi-structural) approach to the extent that we treat family income and other individual and family characteristics as exogenous regressors. This means that we can actually measure the reduced-form effect of income on grade attainment at any specific age of interest and compare it directly with results obtained from regression methods and which have been reported in the existing literature.

In the paper, we answer the following questions:

- Within each cohort, do the effects of family income and AFQT scores on grade progression differ between pre-college transitions (between 16 and 18), college transitions (between 19 and 21) and belated college transitions (beyond age 22)?
- How do the effects of income and AFQT on grade progression probabilities translate in terms of educational outcomes such as grade completed and college participation?
- How have the effects of income, AFQT scores and unobserved heterogeneity evolved between the early 1980's and the early 2000's?
- What would be the impact of ignoring dynamic selection on within-cohort income effect estimates and on their evolution?

The main findings are the following. Contrary to conventional wisdom, our results indicate that the effects of family income on grade progression, highest grade completed and college participation have practically vanished. Between the early 1980's and the early 2000's, income effects have lost between 30% and 80% of their importance on grade transition probabilities, depending on age at which those are measured.

[5] Source: 120 years of American Education: A Statistical Portrait, National Center for Education Statistics, 1993.

In the 79 cohort, a $10,000 difference in income raised the probability of grade progression at age 19 by 0.016, highest grade attainment (at age 25) by 0.073 year and the population proportion of college enrollees by 1.4 percentage point.

In the 97 cohort, the same real income difference raised grade progression probability by 0.004, highest grade completed at age 25 by 0.027 year and college participation by 0.3 percentage point. As a result, it would have taken a $140,000 differential in family income to generate one extra year of education in the early 1980's but a $300,000 differential in the early 2000's. Put differently, a $70,000 family income differential raised college participation by 10 percentage points in the early 1980's. In the early 2000's, a $330,000 income differential had the same impact on college participation.

Despite the growing frequencies of belated grade progression (those taking between 22 and 25) between the 80's and the early 2000's, transitions taking place beyond 22 appear to be entirely disconnected from family income in both cohorts.

Either in the early 1980's or in the early 2000's, AFQT scores have been the most important observed determinant of grade progression and education outcome in the NLSY. However, and aside from pre-college transitions, the effects of AFQT scores on grade transitions have also decreased between the 1980's and the early 2000's. A 1/3 standard deviation increase in AFQT raised grade progression probability at age 20 by 0.050 in the 79 cohort but by only 0.016 in the 97 cohort. The 79 cohort marginal effects translated into increases of 0.16 year of schooling at age 20 and 0.324 at age 25. In the early 2000's, the corresponding AFQT marginal effects had dropped to 0.084 (at age 20) and 0.143 (at age 25). For belated college transitions, the AFQT effects dropped from 0.035 to 0.008. All in all, the effects of AFQT scores on schooling attainments measured at any point between age 19 and 25 have been divided by more than 2 between the early 1980's and the early 2000's.

It is important to note that the vanishing impacts of parental income and AFQT scores are neither the by-product of the allowance for a correlation between unobserved heterogeneity and family income, nor dependent on extreme income levels. Results obtained with orthogonal unobserved heterogeneity are similar and others obtained when removing the top and bottom 1% income levels also disclose the same patterns.

Our findings suggest an interesting phenomenon. Classical educational selection based partly on cognitive abilities and/or family income (present in the 79 cohort) is being gradually dominated by a different form of selection which reserves a more important role to non-cognitive abilities and/or preferences and a lesser role to cognitive ability and family income.[6] Put differently, the evolution of income and AFQT score effects indicate that college expansion has benefitted mostly lower income and lower ability individuals. Indeed, the decrease in the effect of real income on higher education has most likely been sufficiently strong to counteract the increase in income dispersion and thereby prevent an increasing education gap between high and low family income families.

Not accounting for dynamic selection would inflate income effects on both grade progression probabilities and schooling attainments. Because unobserved heterogeneity is found to be relatively more important in the early 2000's, the inflation would be more serious for the 97 cohort. While ignoring unobserved heterogeneity would still reveal decreasing income effects, it would do so from inaccurate within-cohort income effect estimates. Moreover, ignoring dynamic selection would hide the fact that income effects have practically vanished.

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 following section, we describe the evolution of the marginal effects of income and AFQT scores. In Section 6, we illustrate the implications of ignoring dynamic educational selectivity. In the final section, we present a summary of the results along with some economic interpretation and avenues for future research.

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

[6]Interestingly, Deming (2017) analyzes changes in the wage returns to skills using the NLSY79 and NLSY97 and finds that labor market return to social skills was much greater in the 2000s than in the mid 1980s and 1990s.

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.

In a seminal piece, 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.[7] Other studies such as Keane and Wolpin (1997), Carneiro and Heckman (2002) and Cameron and Taber (2004) have 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.[8]

Based on a comparison of the 1979 cohort of the NLSY with the 1997

[7] 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

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

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 (quartile indicators), 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 are higher for the 1997 cohort than the 1979 cohort. 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).[9]

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 have been stable for average ability students and have even decreased for the more able. Chetty et. al. (2014), who were primarily interested in the evolution of the intergenerational 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).[10]

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.[11] To measure the effect of real income, they use four income groups defined from the 1997 quartiles which they interact with AFQT terciles.

[9]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 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)

[10]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 highskilled workers have moved down the occupational ladder and have begun to perform jobs traditionally performed by lower-skilled workers.

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

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.[12]

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 devoted to higher education, differences in real income are much more likely to approximate access to financial resources than relative income measures.[13]

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.

In intuitive terms, the implications of dynamic selection may be illustrated as follows. Consider two types of individuals; one poorly endowed with a characteristic favoring education (for instance, an individual coming from a low-income family) and another type endowed with a high level of the same characteristic (coming from a high income family), and assume the existence of one unobserved factor favoring education and distributed independently from income at the beginning of the accumulation process.

As we move toward realized high education levels (as we consider more

[12] In parallel, a substantial number of papers have tried to estimate the effect of parental resources on educational outcomes using quasi-experimental methods. The literature is voluminous and diverse. Recent examples include Bulman et. al. (2016) who estimate the effects of unearned household income in late childhood from more than one million state lottery wins, Loken (2007) who uses oil revenues in Norway and Lovenheim and Reynolds (2013) who analyze the impact of changes in housing prices. The effects reported in the literature vary greatly across studies but most importantly, and because of the very specific sources of identification used across studies, the recent evolution of income effects has been ignored.

[13] Studies using both the 1979 and 1997 cohorts of the NLSY do not necessarily use the same regressors. For instance, Lovenheim and Reynolds (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.

and more selected sub-populations), an individual coming from a low-income family is more likely to be endowed with a high level of the unobserved factor favoring education than another individual coming from a high income family. This automatically creates a negative correlation between unobserved heterogeneity and family income, which becomes more important as we condition on increasingly high education levels. For instance, this may be particularly serious when evaluating the effect of income on the probability of continuing to college or on the probability of completing college if the sub-population of high school graduates is not representative of the original population.[14]

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).[15]

To summarize, it is impossible to estimate the effects of income, AFQT and other characteristics (along with their evolution) without accounting for the role of unobserved heterogeneity. Evaluating the importance of dynamic selection therefore requires to model grade progression as a sequential process in which unobserved heterogeneity plays an explicit role. This is precisely what we achieve in this paper.

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 infor-

[14] In a context where there is more than one observed characteristic, the argument would be more complicated but the implications of dynamic selection would be as serious.

[15] Evaluating the importance of changes in educational selectivity is currently raising much interest in the literature (see Ashworth et al., 2017, Dillon and Veramendi, 2018, and Bound, Lovenheim and Turner, 2010).

mation 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. The NLSY is one of the most commonly used data set in the US. While the surveys have been constructed to preserve symmetry across cohorts, attrition appears to be slightly more important in the 97 cohort.[16]

One objective guiding us is to use selection criteria that are the closest possible to those used in the literature. Although the NLSY allows for the possibility of constructing larger samples in which minorities are oversampled, we limit ourselves to a random sample. We do so not only to focus on a representative sample, but also because using sample weights within non-linear econometric models with unobserved heterogeneity would be far from obvious.[17]

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), Kinsler and Pavan (2011) and others.

Further, we 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 that our model deals with grade progression, we also require individual grade transitions to be observed for the first six surveys. After these exclusions, we obtain samples of 2,151 individuals for the 1979 cohort, and 2,651 individuals for the 1997 cohort.

Following Belley and Lochner (2007), Kinsler and Pavan (2011) and others, 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 ear-

[16] An in-depth comparison between the 79 and 97 cohorts is found in Nielsen (2015).

[17] There seems to be no well-recognized (standard) method for dealing with sample weights within a dynamic discrete choice model with unobserved heterogeneity.

lier 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.

As is common in the earlier literature, we use AFQT scores to control for cognitive ability. AFQT scores are an average of 4 components of the Armed Services Vocational Aptitude Battery (ASVAB) and should therefore contain a lesser measurement error level than each component introduced separately.[18] We use scores provided by Altonji et. al. (2012) which are adjusted to improve comparability across cohorts. To take into account differences in AFQT's that could be explained by differences in education and age when it was measured, we regress AFQT scores on age and education and use the standardized value of the residual as the cognitive ability indicator.

For each individual, we measure schooling attainment as indicated by the highest grade completed by each given age, and do so between age 16 until age 26. Measuring schooling until grade transitions from age 25 to 26 constitutes a major difference with most of the papers found in the literature (such as Belley and Lochner, 2007) who investigated the determinants of schooling attainments by age 21 (focussing on college participation).

To motivate our approach, it is informative to examine the fraction of the population experiencing a grade progression between age 16 and 25 and for both cohorts (in Figure 1). First, and not surprisingly, the proportion of the population experiencing a grade progression tends to be higher in the 1997 cohort than in the 79 cohort (except at age 16). Second, it is interesting to note that for the 1997 cohort, the proportion remains between 10% and 20% between age 22 and 25. For the 79 cohort, the proportion is smaller but remain significant.

To gage the representativeness of our sample, we report a table in which the average values of some regressors in the overall NLSY core sample may be compared with the averages in our sample. This table is denominated Table S1 in a supplementary file. A detailed description of the number of exclusions induced by our selection criteria is provided in Table S2 also found in supplement. Overall, our sample is quite comparable to the pre-selection NLSY in terms of all observed characteristics.

Some of the most important characteristics of our samples are found

[18]The AFQT components are Arithmetic Reasoning, Word Knowledge, Paragraph Comprehension, and Mathematics Knowledge.

in Table 1 (devoted to summary statistics). First, we note that over a 20 year period, family income has grown by about 21 percent (from $54,155 to $65,572). This corresponds to a 1 percent growth rate per year. These annual growth rates match aggregate measures provided by the Bureau of Labor Statistics.[19] 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.[20]

As is well known, income dispersion has increased even more over this period. In our sample, the standard deviation of family income increased from $33,308 in the 1979 sample to $58,136 in the 1997 sample. Not surprisingly, family income is higher than average among college entrants.

To obtain a clearer picture of the relationship between education and family income, we also compute average schooling attainments (highest grade completed by age 25) for the first, second, third and fourth income quartiles in the 1979 cohort and compare them to the corresponding quartiles in the 1997 cohort. These statistics incorporate both the effects of a change in income effects as well as the effects of an increase in income dispersion due largely to changes in income thresholds defining the third and fourth quartiles.

There are 2 main observations to be made after examining highest grade completed and income quartiles. First, schooling attainments have increased at all income quartiles. A second feature emerging is that the largest increase has been observed for those in the third quartile (from 13.3 to 14.1 years). Those in the first, second and fourth quartiles have experienced a practically identical increase of about 0.4 year.

Information about AFQT scores are found in the second panel of Table 1. 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 (from 185.8 to 182.2). In line with the recent evolution of college selectivity described in Hoxby (2009), this suggests that college has become globally less selective on cognitive skills.

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

[20] 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.

Finally, the 3rd panel is devoted to family characteristics. Among observed characteristics, mother's education, increasing from 11.7 to 13.2 years, and the intact family indicator, going from 0.788 to 0.581, are those that have changed the most. There is also a relatively important decrease in number of siblings going from 3.1 to 2.3.

Before discussing our econometric model, it is interesting to evaluate the impact of real income on highest grade completed obtained from standard OLS regressions and compare those with estimates measuring differences in schooling by income quartiles. As mentioned earlier, the literature has focus almost exclusively on inter-quartile education differences on relatively early education outcomes (around age 21).

In Table 2, we report OLS estimates of regressions of highest grade completed by age 25 on both real income and relative income (income quartiles) as well as individual characteristics. The most striking result is by far the decrease in the effect of real income. In the 1979 cohort, the effect of a $10,000 difference in family income was about 0.12 year of schooling. In the 1997 cohort, it has been divided by 3 to reach 0.04.

Interestingly, this spectacular decrease is not incompatible with the existence of some increase in education attainments differential between specific quartiles. In reference to the first quartile (Q1), the 3rd quartile (Q3) differential appears to be the only one which has increased as it moved from 0.495 in the 79 cohort to 0.652 in the 97 cohort. Indeed, it is also noticeable that education differences between those in the 3rd and 2nd quartiles also increased as the Q3-Q2 difference grew from 0.23 year of schooling (0.495-0.257) in the 79 cohort to 0.43 year (0.652-0.222) in the 97 cohort.

However, and at the same time, it is interesting to note that the schooling attainment differences between the highest and the lowest quartiles has decreased, going from 0.994 (in the 79 cohort) to 0.821 (in the 97 cohort). Similarly, the education gap between the second and the first quartile has practically not moved.

To summarize, OLS estimates disclose an interesting paradox. That is the existence of a simultaneous increase in differences in average schooling attainments between some income quartiles (but not all) and a decrease in the effect of real income on schooling attainments. Obviously, these differences do not account for dynamic selection. If the grade accumulation process is affected by unobserved heterogeneity and its dispersion is sufficiently high, income effect estimates may change drastically.

4 A Model of Educational Attainments

To construct our model, we build on ideas governing the literature on reduced-form models of schooling such as Cameron and Heckman (2001) and Ashworth et. al. (2018) and in which intertemporal 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 allows for potential discontinuities in the schooling accumulation process, an issue that has raised much interest in recent years. [21]

4.1 Choice Probabilities

We assume that the decision process starts at age 16. The choice variable is denoted $d(a)$, where $d(a) = 1$ when an individual decides to invest in an additional grade attainment at age a and where $d(a) = 0$ 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, every individual therefore chooses between accumulating an additional grade or to involve in other activities. When $d(a) = 1$, it follows that $G(a + 1) = G(a) + 1$.

To avoid the estimation of an excessively large number of parameters, we allow the parameters of the model to vary with age but according to 3 different levels only. The first level captures the effect of regressors and unobserved factors on progression between 16 and 18. The second one covers age 19 until 21, while the third one deals with transitions from age 22 to 25 (when financial aid applicants are more likely to become independent from their parents).

Given age, we allow grade progression to depend on accumulated grade level. To do so, we define an indicator denoted $G^{co}(a)$ equal to 1 when an individual has completed high school and 0 if not. It allows us to capture

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

changes in the cost of education when reaching post-high school education.

The school choice probabilities are defined as follows.

From age 16 to 18:

$$
\begin{aligned}
\Pr(d(a) &= 1 \mid N(a), G(a), X_i, birthyear, AFQT_i, Fam.Income) \\
&= \Lambda(\beta_{0i}^{16} + \beta_x^{16} \cdot X_i + \beta_{by}^{16} \cdot birthyear_i + \beta_n^{16} \cdot N(a) + \\
&\quad \beta_{AF}^{16} \cdot AFQT_i + \beta_{FI}^{16} \cdot Fam.Income + \beta_{FI,2}^{16} \cdot Fam.Income^2)
\end{aligned}
$$

From age 19 to 21:

$$
\begin{aligned}
\Pr(d(a) &= 1 \mid N(a), G(a), X_i, birthyear, AFQT_i, Fam.Income) \\
&= \Lambda(\beta_{0i}^{19} + \beta_x^{19} \cdot X_i + \beta_{by}^{19} \cdot birthyear_i + \beta_{co}^{19} \cdot G^{co}(a) + \\
&\quad \beta_n^{19} \cdot N(a) + \beta_{AF}^{19} \cdot AFQT_i + \beta_{FI}^{19} \cdot Fam.Income + \beta_{FI,2}^{19} \cdot Fam.Income^2)
\end{aligned}
$$

From age 22 onwards:

$$
\begin{aligned}
\Pr(d(a) &= 1 \mid N(a), G(a), X_i, birthyear, AFQT_i, Fam.Income) \\
&= \Lambda(\beta_{0i}^{22} + \beta_x^{22} \cdot X_i + \beta_{by}^{22} \cdot birthyear_i + \beta_{co}^{22} \cdot G^{co}(a) + \\
&\quad \beta_n^{22} \cdot N(a) + \beta_{AF}^{22} \cdot AFQT_i + \beta_{FI}^{22} \cdot Fam.Income + \beta_{FI,2}^{22} \cdot Fam.Income^2)
\end{aligned}
$$

where

$$
\begin{aligned}
\beta_{0i}^{19} &= \delta_0^{19} + \delta_1^{19} \cdot \beta_{0i}^{16} \\
\beta_{0i}^{22} &= \delta_0^{22} + \delta_1^{22} \cdot \beta_{0i}^{16}
\end{aligned}
$$

and where X_i is a vector of individual and family characteristics (already defined in Section 3), $birthyear_i$ is a set of binary indicators for the year of birth, and $\Lambda(.)$ denotes the logistic distribution function.

Aside from the distribution of unobserved heterogeneity, the parameters to be estimated are $\beta_x^{16}, \beta_{by}^{16}, \beta_n^{16}, \beta_{AF}^{16}, \beta_{FI}^{16}$ and $\beta_{FI,2}^{16}$ as well as both their age 19-21 pendants (with superscript 19) and their age 22-25 pendants (with superscript 22).

4.2 Unobserved Heterogeneity and the Initial Condition

To close the model, we now provide more details about unobserved heterogeneity. In line with Bajari, Fox and Ryan (2007) and Train (2008), we adopt a fixed mass points approach by choosing grid points covering the entire range of possible values and estimate all associated frequencies (type probabilities) while adapting this approach so to incorporate the endogeneity of initial conditions. To start with, we assume M support points.[22] Each type m is endowed with a vector $\{\beta_{0m}^{16}, \beta_{0m}^{19}, \beta_{0m}^{22}\}$ and impose a positive correlation by forcing both δ_1^{19} and δ_1^{22} to be non-negative.

In an ideal setting, we would access measures of non-cognitive skills for both cohorts and evaluate changes in its effect on grade attainment as we do for AFQT scores. Unobserved heterogeneity would then be interpreted as a residual factor prevailing after conditioning on cognitive and non-cognitive factors.

This solution is however impossible to implement. There exist some non-cognitive skill measures in the NLSY but those are only available in the 1979 cohort. Those are the Rotter Locus of Control index and the Rosenberg self-esteem score. The Rotter Locus of Control is designed to measure the extent to which individuals believe they have control over outcomes through self-motivation or self-determination (internal control) as opposed to the extent that the environment controls their lives (external control). The Rosenberg Self-Esteem Scale is a 10-item scale measuring the self-evaluation that an individual makes. It describes a degree of approval toward oneself.

To circumvent orthogonality conditions between unobserved heterogeneity and observed characteristics (family income, AFQT scores, and others), and in order to make use of the availability of non-cognitive measures in the 79 cohort, we proceed as follows. First, we obtain a predicted Rotter scale (denoted $\hat{L}_i$) and a predicted Rosenberg scale (denoted $\hat{R}_i$) by regressing them on all observed characteristics and AFQT scores using the 79 cohort. We then use the parameters to construct predicted values for both the 79 and the 97 cohorts, assuming a stable relationship between the measures and parental background variables. This leads us to a vector of predicted values

[22] To choose them, we first estimate the model without unobserved heterogeneity and select an equal number of points above and below the estimated intercept for β_0^{16}.

for each cohort.[23]

We then build a distribution of unobserved heterogeneity incorporating those values.[24] While it would have been possible to use the actual scores for the 79 cohort, this would have introduced an asymmetry between cohorts. Instead, we preferred using predicted Rotter and Rosenberg scores in both cohorts. To capture both the correlation between unobserved heterogeneity and individual characteristics (including AFQT scores) and the endogeneity of grade attainment by age 16 (the initial condition), the type proportions, $p_m(.)$, are allowed to depend on the initial grade level, $G(16)$, and on predicted values of the Locus of Control index and the Rosenberg scale.

We write them as follows

$$p_m(G_{16}, \hat{L}_i, \hat{R}_i) = \frac{\exp(\tilde{p}_m + \tilde{p}_{mG} \cdot G(16) + \tilde{p}_{mL} \cdot \hat{L}_i + \tilde{p}_{mR} \cdot \hat{R}_i)}{1 + \sum_{j=2}^{M} \exp(\tilde{p}_j + \tilde{p}_{jG} \cdot G(16) + \tilde{p}_{jL} \cdot \hat{L}_i + \tilde{p}_{jR} \cdot \hat{R}_i)}$$

While this approach is based on the maintained assumption that Rotter Locus of Control index and the Rosenberg self-esteem score are stable across cohort (and that their correlation with characteristics are also stable), the flexibility of the parameters $\tilde{p}_{mL}$ and $\tilde{p}_{mR}$ which can vary across cohorts, allows us to measure variations in the relative importance of unobserved heterogeneity. We can also compare results obtained assuming orthogonality with those obtained with the latter approach in order to examine if our results are affected or not by the allowance for correlated effects.

4.3 The Likelihood Function

We estimate the model by mixed likelihood techniques. Each individual $i's$ grade attainment progression history is contained in the following vector

$$\{G(16)_i, d_i(a = 16), d_i(a = 17), ...d_i(a = 25)\}$$

and the likelihood function for observation i is equal to

[23] The results of those OLS regressions are found in Table S4 in a supplementary file.

[24] A more ambitious alternative option (which we follow in a companion paper) would be to make use of the ASVAB measurement components in both cohorts, along with the components of the Rotter index and the Rosenberg index, and estimate the stationary (across cohort) distribution of cognitive and non-cognitive factors while allowing those to depend on observed characteristics.

$$L_i(.) = \sum_{m=1}^{M} p_m(.) \cdot \prod_{a=16}^{25} (\Pr(d_{ia} = 1 \mid type\ m))^{I(d_{ia}=1)} \cdot (\Pr(d_{ia} = 0 \mid type\ m))^{I(d_{ia}=0)}$$

where $I(.)$ is the identity function. The likelihood of the sample data is formed by the product of each individual contribution.

5 Changes in the Effects of Family Income and AFQT Scores

To estimate the model, we defined the heterogeneity distribution over 10 support points ranging from -2.0 to 0.0 with intervals of 0.2 for the 79 cohort and ranging from -2.5 to -0.5 with intervals of 0.2 for the 97 cohort. As each type probability depends on 4 parameters (an intercept plus parameters associated to grade attained by age 16, the predicted Rosenberg scale and the predicted Rotter scale), we therefore characterize the heterogeneity distribution with 46 parameters (36 for the type probabilities and 10 fixed support points). The distribution of unobserved heterogeneity and the evolution of educational selectivity will be summarized in Section 6. Other specifications will be discussed later in a sub-section devoted to robustness checks.

As for most non-linear models, the parameter estimates are not directly informative. For this reason, we focus our presentation of the results on two different statistics illustrating the mechanics of the model and present the parameter estimates in Table S5 found in a supplementary file. As the model is sequential and allows for the effect of income and AFQT scores to change when individuals reach 19, and to change again when they reach 22, we first compute the income and AFQT marginal effects on the probability of attaining an additional grade level at 18, 20 and 22.[25]

Our second set of marginal effects captures the impact of family income and AFQT scores on different measures of schooling attainment at a given age (highest grade completed and a college attendance indicator). They are obtained by compounding grade progression probabilities until age 26 and

[25] As is normally expected from a reduced-form model incorporating flexible unobserved heterogeneity, it is found to fit the data quite well. The fit is summarized in a table found in a supplementary material file (Table S3).

regressing those outcomes on measured family income. With these marginal effects, it is possible to relate our results to those reported in the literature in which observed schooling outcomes are regressed on income and a set of individual and family characteristics.

In order to illustrate the potentially changing impact of income with age, we measure income effects on educational outcomes at different ages (18, 22, 25) so to allow the age-income effect differential to set in. Schooling attainments measured at age 25-26 are likely to provide a more reliable indicator of total life-cycle schooling than age 22 as a significant proportion of individuals experience grade transitions beyond 22.

To obtain the marginal effects, we simulate educational choices between age 16 and 25 for a large number of individuals (10,000) reflecting the unobserved type distribution as well as random shocks. We then compute marginal effects on grade progression probabilities and highest grade completed and use the asymptotic distribution of parameter estimates to evaluate their standard errors.

5.1 The Effects of Income and AFQT Scores on Grade Progression

The marginal effects of income and net AFQT scores on the probability of obtaining an additional grade are reported in Table 3. When estimating the model, family income is measured in units of $10,000 (the base year is 2000) and AFQT scores are introduced as standardized residuals of a regression of AFQT scores on grade completed when ASVAB tests were administered. Note that our AFQT scores are already adjusted to account for differences across cohorts.

To obtain these effects, we set all observed regressors to their sample average. We evaluate marginal effects 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 of $10,000. For the 1979 cohort, the four income levels are $17,434, $39,487, $60,269 and $99,184, respectively. For the 1997 cohort, they are equal to $13,937, $46,726, $69,631 and $141,855. Expressed in terms of the standard deviation of the income distribution, a $10,000 increase is equivalent to about 33% of the standard deviation for the 1979 cohort and less (about 17%) for the 1997 cohort. To provide an overall picture, we also compute them at mean family income.

For the effects of AFQT scores, we proceed similarly and compute the marginal effect of AFQT's at each income quartile and at mean income. To facilitate comparability with income effects, we measure the impact of a change equal to 1/3 of the standard deviation of AFQT's to obtain a change comparable to the one we use for income. The marginal effects of income and AFQT scores on grade transition are averaged over types.

The results have been obtained for the model specification with correlated effects but also from a specification in which type probabilities depend only on initial grade level. Because the main results are practically identical and all the main conclusions persist when the distribution of types depends only on initial grade, we comment only on those obtained with correlated effects. Estimates obtained with alternative specifications will be discussed below.

5.1.1 Income Effects

Before summarizing the main features, we present the main estimates. In the 79 cohort, a $10,000 increase in income raises the probability of a grade progression at age 18 by 0.022 (at quartile 1), by 0.014 (at quartile 2), by 0.008 (at quartile 3), and by practically 0 (at quartile 4) in the 79 cohort. At age 20, the same amount increases the probability of a grade progression by 0.015 (at quartile 1), by 0.013 (at quartile 2), by 0.011 (at quartile 3), and 0.008 (at quartile 4). However, the importance of family income vanishes after age 21 as indicated by marginal effects ranging from 0.006 (quartile 1) to 0.004 (quartile 4) beyond 22.

We now examine the equivalent marginal effects for the 1997 cohort. A $10,000 increase in income raises the probability of a grade progression at age 18 by 0.009 (at quartile 1), by 0.008 (at quartile 2), by 0.007 (at quartile 3), and by 0.004 (at quartile 4). At age 20, the marginal income effects are also low on average as they lie between 0.006 (quartile 1) and 0.002 (quartile 4). As for those measured at age 18, they do not seem to vary with family income level. Finally, at age 22, income affects have practically vanished for all income levels.

There are 4 main features about grade progression income effects to retain from Table 3.

First, and aside from transitions taking place beyond age 21, family income effects on grade transitions have a concave shape as they tend to be generally higher at low income levels than at higher ones. This is true in both cohorts although concavity is much less pronounced in the 97 cohort.

A second finding is that income effects on pre-college transitions (from age 16 to 18) are as strong as they are on transitions taking place from age 19 onwards. For instance, when evaluated at mean income, the 79 cohort income effect on age 20 transitions (equal to 0.012) is practically identical to that measured at age 18 (which is equal to 0.010). For the more recent cohort, the average income effect at age 18 (averaging 0.007) actually exceeds the age 20 income effect (which is equal to 0.004) but these are so small that their differences may not raise substantial interest.

Keeping in mind that the vast majority of grade progression transitions takes place before age 22, the third and most important finding remains the overwhelming evidence that the effects of family income on grade progression have decreased between the early 80's and the early 2000's. When comparing marginal effects measured at 18 and 20 at each specific income quartile obtained for the 79 cohort with those obtained for the 97 cohort, we observe that grade progression probabilities decrease by numbers ranging between 0.01 (at low income) to 0.005 (at high income). When measured at average income, these estimates translate into a 30% decrease in pre-college grade progression income effects and practically a 80% decrease for early college transitions. The effects of income at age 22 have also decreased by about 0.005 to approach 0.

As all these estimates measure the impact of income on a single grade transition probability, compounding those effects over early adulthood may be sufficient to generate even more significant differences in income effects when considering total schooling by age 22 or beyond. We shall return to this point below.

Finally, a fourth finding is that although we already noted the growing frequencies of belated grade progression between the 80's and the early 2000's, those transitions appear to be entirely disconnected from family income. While we also note a decrease in income effects between the early 80's and the early 2000's, their low magnitude obviate the need for further comments on their evolution.

5.1.2 The Effects of AFQT Scores

With respect to grade progression effects of AFQT scores, there are 3 main points to be made. First, AFQT scores have been the most important determinant of grade progression. Notwithstanding that the effects of age and education have been removed from AFQT measures, a 1/3 standard deviation difference in the AFQT residual (evaluated at mean income) changes grade progression probabilities by 0.022 at age 18, by 0.050 at age 20, and by 0.035 at age 22, in the early 80's. As a comparison, increments of $22,000, $41,000 and $70,000 would have been needed to generate a similar change in grade progression. In the early 2000's, the AFQT marginal effect on early grade transitions, equal to 0.021, was equivalent to a $30,000 increase while the effect on transitions between 19 and 21, equal to 0.016, was worth about $40,000. The AFQT effect on belated progression (at age 22), equal to 0.008, cannot be evaluated in dollars as income does not matter for those transitions.

Within cohort, a second finding is that AFQT scores had stronger effects on early college transitions (at age 20) than on both high school and belated college transitions in the 79 cohort. However, in the early 2000's, the effects of AFQT's appear to decrease with age. For instance, AFQT's have practically no impact on age 22 transitions. In the older cohort, AFQT had still strong effects on belated college transitions.

Third and foremost, and across cohorts, the effects of AFQT scores on grade transitions have decreased between the 1980's and the early 2000's. The only exception has been pre-college transitions (from 16 to 18) for which AFQT scores effects have remained constant. For the rest, they have been divided by more than 3. To illustrate this, and when evaluated at mean income level, a 1/3 standard deviation increase in AFQT raised early grade progression by 0.022 in the 79 cohort and by 0.021 in the 97 cohort. However, for transitions taking place at age 20, the same change in net AFQT raised grade progression by 0.050 in the 79 cohort but by only 0.016 in the 97 cohort. For belated college transitions, the AFQT effects dropped from 0.035 to 0.008. This indicates that differences in cognitive skills, to the extent that those are measured by AFQT scores, have become less important for college transitions but remained an important determinant of grade progression until age 18.

5.2 The Effects of Income and AFQT Scores on Educational Outcomes

After documenting the effects of income on grade progression probabilities, we now turn to marginal effects of income on educational outcomes as recorded by a given age. As noted earlier, much of the existing literature has focussed on outcomes measured around age 21-22 and ignored age-dependent income effects. To compute them, we proceed similarly and use simulated outcomes at age 18, 20, 22 and 25. The results, found in Table 4, summarize income and AFQT effects on highest grade completed and also on a college attendance indicator equal to 1 if an individual has ever attended college. To avoid repetitive comparisons, we report the effects evaluated at mean income.

5.2.1 Income Effects

In the 79 cohort, a \$10,000 change in income raised highest grade completed at age 18 by 0.021 , by 0.049 at age 20, by 0.065 at age 22 and by 0.073 by age 25.

In the 97 cohort, a \$10,000 change in income raised highest grade completed at age 18 by 0.017 , by 0.022 at age 20, by 0.027 at age 22 and by 0.027 by age 25). Aside from the relatively less important drop in income effects on grade attainment measured at age 18, income effects on schooling measured at 20, 22 and 25, have lost more than 50% of their magnitude.

When education outcomes are summarized by the college attendance indicator, the decrease is even more spectacular. In the 79 cohort, a \$10,000 differential raised the population proportion attending college by 1.4% but by the early 2000's, it did by only 0.3%.

As a result, it would have taken a \$140,000 differential in family income to generate one extra year of education in the early 1980's but a \$300,000 differential in the early 2000's. Put differently, a \$70,000 family income differential raised college participation by 10 percentage points in the early 1980's. In the early 2000's, a \$330,000 income differential had the same impact on college participation.

As already noted when we analyzed grade progression probabilities, this general finding goes against conventional wisdom as economists are often tempted to assimilate the documented increase in educational differences between high and low income quartiles (which often ignores belated college completion) to an increase in the effect of real income. Our findings suggest that it is far from being the case.

5.2.2 AFQT Scores

As noted earlier, and when examining the evolution of AFQT effects, a distinction must be made between pre-college transitions and transitions taking place beyond age 18. Not surprisingly and in line with the evolution of grade progression probability income effects, the effects of AFQT scores on grade attainments measured at age 20, at age 22 and at age 25 have all decreased substantially between the 80's and the early 2000's. This is the case at all income quartiles and also at mean income. More precisely, 1/3 standard deviation change in AFQT residuals in the 80's increased grade attainment at age 20 by 0.162 year, at age 22 by 0.254 and by 0.324 by age 25. In the early 2000's, the corresponding AFQT marginal effects had dropped to 0.084 (at age 20), 0.116 (at age 22) and 0.143 (at age 25). All in all, the effects of AFQT on schooling attainments measured at any point between age 19 and 25 have been divided by 2 between the early 80's and the early 2000's.

As for family income, the drop in AFQT score effects is more spectacular when stated in terms of college attendance. In the early 80's, a 1/3 standard deviation increase raised college attendance by 6.4 percentage points. In the early 2000's, it did so by only 1.5 percentage points.

To summarize, our findings show a strong decrease in the effect of real family income, at least until age 22. Moreover, what is true for family income is also true about AFQT scores. The only difference between family income and AFQT scores is that the effects of the latter have not vanished. Indeed, AFQT scores remained a key determinant of college completion in the early 2000's.

5.3 Robustness

To verify the robustness of our results, we implemented two different versions of the model.

First, and in order to check to what extent our main results are driven by the correlation between unobserved abilities and the projection of non-cognitive skill measures onto family characteristics, we estimate a version that does not use non-cognitive measures and thereby assumes that the distribution of types depends on initial grade only. In such a case, each type m is therefore endowed with a vector $\{\beta_{0m}^{16}, \beta_{0m}^{19}, \beta_{0m}^{22}\}$ with proportion $p_m(.)$ where

$$p_m(G_{16}) = \frac{\exp(\tilde{p}_m + \tilde{p}_{mG} \cdot G(16))}{1 + \sum_{j=2}^{M} \exp(\tilde{p}_j + \tilde{p}_{jG} \cdot G(16)}$$

This sort of specification, common in the econometrics of duration data and dynamic discrete choices, would only allow for an indirect correlation between unobserved heterogeneity and family income, AFQT and other regressors to the extent that grade attainment at grade 16 is also correlated with those variables.

The results are summarized in Table A1 found in appendix and indicate that most marginal effects remain comparable (up to the 3rd decimal). All features regarding the decreasing importance of income effects and AFQT scores, and which were discussed earlier, carry through.

Second, and in order to evaluate the potential sensitivity of our results to the presence of very high and very low income, we also re-estimate the models after removing the top 1% and the bottom 1% income levels. This may be justified by the potentially higher likelihood of mis-measurement of income at extreme levels.

Overall, removing those extreme income levels has no impact on our main results. The results, found in Table A2 in appendix, still indicate lower marginal effects on both grade progression and highest grade completed for the 97 cohort compared with the 79 cohort.

To illustrate this, let's focus on grade attainments and college participation. For highest grade completed, the age 25 income effect dropped from 0.073 (in the 79 cohort) to 0.040 (in the 97 cohort). For the incidence of income on college participation, we note a drop from 1.8% to 0.4%.

This indicates relatively clearly that our main results appear to be in no way explained by the behavior of individuals endowed with either very low or very high family income levels.

5.4 Comparison with a Relative Income Approach

As noted in Section 2, the existing literature has focussed mostly on documenting the evolution of education inequality in conjunction with income inequality and for this reason studied the evolution of average education outcomes at different income quantiles.

Although individual decisions are most likely not guided by relative income position, we estimated a version of our model specification with non-cognitive measures after replacing real income by a set of quartile indicators in order to obtain estimates of the evolution of educational differences which would be more comparable with those reported in the literature.

When estimating the model, we use the first quartile as the reference group and report marginal effects as the difference in schooling attainments between a given quartile and the first one. These educational differentials are partly affected by the evolution of real income effects and by changes in family income distribution. The estimates are presented in Table 5. We comment on grade attainments and proportion of college attendance at age 25.

Although reported educational differences may not be interpreted as emerging solely from changes in real income effects, the results are in line with the decreasing trend documented earlier. They show that there was a reasonably steep relative income gradient in the 1979 cohort for college attendance but that it has virtually disappeared for the 1997 cohort. Similarly, the attendance probability differential across quartiles measured in the 79 cohort turned out to be substantially larger than for the 97 cohort.

To be precise, the grade attainment differentials were equal to 0.223 for quartile 2, 0.254 for quartile 3, and 0.713 for quartile 4 in the 79 cohort. When expressed in terms of college enrollment proportions, the differentials were equal to 0.058 (quartile 2), 0.070 (quartile 3), and 0.190 (quartile 4).

Turning to the 97 cohort, the corresponding grade attainment differential between quartile 2 and quartile 1 was practically equal to 0 while the others were equal to 0.220 (quartile 3), and 0.201(quartile 4). The pattern is similar for college attainment differentials which are equal to -0.002 (insignificant) for quartile 2, 0.018 for quartile 3, and 0.027 for quartile 4.

Either in terms of grade attainments or college completion, educational differences between those in quartiles 2, 3 and 4 on one hand and those in quartile 1 on the other hand, have all decreased. This is particularly noticeable when considering the differentials between the second and first quartiles which have practically vanished.

To some extent, these results further illustrate the robustness of our general finding of weaker income effects on college attendance and grade completion for the 97 cohort. They indicate that the decrease in the effect of real income on higher education has most likely been sufficiently strong to counteract the effect of the increase in income dispersion and thereby prevent an increasing education gap between high and low family income families, after conditioning on both observed and unobserved characteristics.

6 Changes in Educational Selection

In order to comprehend the sources of changes in marginal effects of family income and AFQT scores, it is necessary to evaluate the relative importance of unobserved heterogeneity within each cohort and to quantify its evolution. This is a crucial step toward understanding our main results.

6.1 The Importance of Unobserved Heterogeneity

There are 3 specific questions that we need to answer. Which observed characteristics are more highly correlated with unobserved heterogeneity? Has the dependence of the distribution of unobserved heterogeneity on observed characteristics changed ? What happened to the distribution of unobserved heterogeneity across cohort and to what extent has the relative importance of unobserved heterogeneity changed between the early 1980's and the early 2000's?

The distribution of unobserved heterogeneity is summarized in Table 6. In both cohorts, we find evidence of 4 distinct types. In the 79 cohort, more than half of the population is endowed with an intercept of -2.0 while about a quarter has -0.4. Dispersion is also high in the 97 cohort as we find an intercept of -2.5 for 10%, -1.1 for 26%, and a total of 65% at -0.7 or -0.5.

To obtain a clearer picture and to answer these questions formally, we simulate the distribution of unobserved heterogeneity using the parameters measuring the impact of initial grade and non-cognitive measures on type probabilities. We then regress realized unobserved heterogeneity on each individual characteristic one-by-one and on the totality of characteristics. We report the R-squares of each regression in column 1 (79 cohort) and in column 2 (97 cohort) of Table 7. We also ran separate regressions of simulated schooling (as measured by age 25) on individual characteristics as well as on unobserved heterogeneity. This allows us 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. Those results are in column 3 (79 cohort) and column 4 (97 cohort).

There are three elements to retain from the first two columns of Table 6 on the decomposition of the distribution of unobserved heterogeneity. First, the R-squares of the regressions incorporating income and AFQT scores are both equal to 0.075 in the 79 cohort. While these indicate relatively low correlations, AFQT's and income are the characteristics with the highest level of correlation with unobserved types. The intact family indicator is also relatively correlated with unobserved heterogeneity.

Second, in the 97 cohort, AFQT score is the variable with the highest R-square, which is equal to 0.058, but the movement in family income R-squares (from 0.075 to 0.027) indicates that income has become relatively less correlated with the unobserved factor.

Third, and when taken globally, our estimates indicate that the overall correlation between unobserved heterogeneity and family characteristics has decreased. This is exemplified by the lower R-square obtained for the 97 cohort (equal to 0.089) compared to the one for the 79 cohort (equal to 0.15).[26]

Turning to the decomposition of educational outcomes (in the last columns of Table 7), we note that among observed individual and family characteristics, only the "Intact Family" indicator, has become notably more important over the period considered. Its R-square going from 0.027 to 0.088 shows that being raised within a nuclear family has become a better predictor of education in the early 2000's than in the 80's.[27] However, mother's education remains the best predictor.

In line with our main results, a second notable finding is the drop in both AFQT R-squares, which went from 0.343 in the 79 cohort to 0.262 in the 97 cohort, and family income, which dropped from 0.162 to 0.076. De-

[26]Overall, both the Rotter scale and the Rosenberg scale are weakly correlated with observed regressors as indicated by the R-squares of the regressions used to generate their predicted values, and equal to 0.06 and 0.08 respectively. AFQT score is the characteristic that has the highest level of correlation with both measures. The effect of income on the Rotter index is very weak and its effect on the Rosenberg self-esteem measure is insignificant (see table S4 in the supplementary file).

[27]It is however impossible to say if this relationship is causal or just explained by a change in composition of the sub-population experiencing family separations. It is also interesting to note that Blacks also tend to be more educated, after conditioning on all characteristics. This finding is not incoherent with the generally lower proportion of Blacks attending college found in many US data sets.

spite that, AFQT score remains the most important predictor of educational attainments in the early 2000's before mother's education.

The third and most striking finding, is the spectacular increase in the relative importance of unobserved heterogeneity. In the early 80's, unobserved heterogeneity accounted for about 50% of schooling attainments. By the early 2000's, unobserved heterogeneity has become by far the most important determinant of education and accounts for almost 75% of the explained variations in schooling. It has therefore become more important than the entire set of individual and family characteristics.

To summarize, our results suggest that classical educational selection based partly on cognitive abilities and/or family income (present in the 79 cohort) is being gradually dominated by a different form of selection which reserves a more important role to non-cognitive abilities and preferences and a lesser role to cognitive ability and family income. Put differently, the evolution of income and AFQT score effects indicates that college expansion has benefitted mostly lower income and lower ability individuals. Indeed, the decrease in the effect of real income on higher education has most likely been sufficiently strong to counteract the increase in income dispersion and thereby prevent an increasing education gap between high and low family income families.

6.2 Impact of Ignoring Unobserved Heterogeneity

Evaluating the implication of ignoring unobserved taste for schooling is also a crucial issue to comprehend our main results. The consequences of ignoring dynamic selection on estimates of the effect of income on educational outcomes are however difficult to anticipate as our model allows for age-differentiated effects of unobserved heterogeneity and observed characteristics on schooling. If there are going to be consequences of ignoring dynamic selection on income effects, we expect them to be more important on the recent cohort since we found unobserved heterogeneity to be relatively more important for it.[28]

[28]In some specific econometric settings, such as Proportional Hazards models in which unobserved heterogeneity hits the hazard rate multiplicatively, the impact of ignoring dynamic selection may be derived analytically. See Lancaster, 1990, for a formal analysis in a continuous duration setting, and Cameron and Heckman (1998) for a discussion within a discrete duration setting.

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 income effects obtained at the average income level. Those are found in Table 8. To ease comparison, we also reproduce the estimates of Table 3 (our most general model) in the bottom panel of Table 8.

The impact of ignoring unobserved heterogeneity comes out relatively clearly. First, and as normally expected, income effects on early grade progression are much less affected than later ones as the effect of dynamic selection sets in more clearly as age progresses. In the 79 cohort, the average income effect on grade progression probabilities taking place at age 20 and 22 are multiplied by 2, going from 0.012 to 0.023 at 20, and from 0.005 to 0.009 at 22. In the 97 cohort, income effects are even more inflated. For instance, income effects at age 20, equal to 0.004 with unobserved heterogeneity, are multiplied by 4 to reach 0.016. At 22, the average income effect ignoring heterogeneity is positive (0.006) while that obtained when incorporating is equal to 0 at 4 decimals.

When translated into grade attainment income effects, ignoring heterogeneity inflates income effects in both cohorts but in different proportions. It multiplies them by 2 in the 79 cohort and by 3 to 4 in the 97 cohort. At age 25, ignoring heterogeneity moves income effects on grade attainment from 0.073 year to 0.147 year and raised college participation income effects from 1.4 to 3.8 percentage points in the 79 cohort. In the 97 cohort, it moves grade attainment income effects from 0.027 year to 0.115 year and inflates college participation income effects from 0.3 to 2.3 percentage points.

The consequences of ignoring unobserved heterogeneity are easy to summarize. Not accounting for dynamic selection inflates income effects on both grade progression probabilities and schooling attainments within each cohort. Because unobserved heterogeneity is found to be relatively more important in the early 2000's, the income effect inflation is even more serious for the 97 cohort. While ignoring unobserved heterogeneity would still reveal decreasing income effects, it would do so from inaccurate estimates. Moreover, ignoring dynamic selection would hide the fact that income effects have practically vanished.

Finally, and upon comparing AFQT effects with and without heterogeneity, the conclusion is similar. In both cohorts, ignoring unobserved heterogeneity drives AFQT marginal effects at much higher levels than those obtained when unobserved heterogeneity is properly accounted for.

7 Interpretations and Conclusion

To appreciate our main results, it is useful to put them in perspective with the existing literature. As noted earlier, the vast majority of papers have concentrated on measuring gaps in enrollments between children of different family income quartiles. By doing so, most authors ignored 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. Some conclude in favor of an increasing gap in education between top and bottom quartiles of the distribution while others report a more stable pattern.

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. In the paper, we present evidence that income effects on higher education attainments have lost most of their impact between the 1980's and the early 2000's. Our findings suggest that the decrease in the effects of real income on higher education has most likely been sufficiently strong to counteract the increase in income dispersion and thereby prevent an increasing education gap between high and low family income families. At the same time, classical educational selection based partly on cognitive abilities and/or family income appears to be gradually dominated by a different form of selection which reserves a more important role to non-cognitive abilities and preferences than in the past.

As noted earlier, our main finding is in accord with the evolution of college attendance in the US. The number of individuals graduating with a Bachelor's degree has fluctuated between 1.5 and 2.0 million per year in the past 20 years and is larger than what was observed in the early 80's.[29] When coupled with the increased capacity at lower quality and lower tuition institutions, this may explain the decreasing impact of real family income and AFQT score on educational attainment.

The effect of family income on education inequality is however far from being a closed research topic. Our results raise questions that may generate

[29]Source: 120 years of American Education: A Statistical Portrait, National Center for Education Statistics (1993).

avenues for future research. One would be to evaluate to what extent family income-based education inequality might have been replaced by family income-based "trajectory" inequality. That is it would be crucial to know if those coming from lower income families have been forced to experience trajectories leading to college graduation that use intensive labor supply (while in school) in order to finance their education. If so, it would also be interesting to know if those using more labor supply intensive trajectories have been penalized for doing so.

Our paper has focussed on the US, which is known to provide high incentives to education but relatively low investment levels for disadvantaged children. It would be interesting to see how income effects have evolved in countries where a higher share of public expenditures goes to skill investment among the disadvantaged and where university tuition are lower.[30]

Finally, it would be highly relevant to investigate the factors lying behind the increase in relative importance of unobserved heterogeneity. These are fundamental issues that we are currently examining in ongoing research.

8 References:

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

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

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

[4] Ashworth, Jared, V. Joseph Hotz, Arnaud Maurel, and Tyler Ransom (2018), Changes across Cohorts in Wage Returns to Schooling and Early Work Experiences, Working paper, Duke University.

[5] Bajari, Patrick, Fox, Jeremy and Steve Ryan (2007) "Linear Regression Estimation of Discrete Choice Models with Nonparametric Distributions

[30] Landerso and Heckman (2017) study differences in intergenerational mobility between Denmak and the US and conclude that despite differences in public policies, they disclose a similar correlation between family resources and educational attainment.

of Random Coefficients" *American Economic Review.* Apr 2007, Vol. 97, No. 2: Pages 459-463

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

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

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

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

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

[10] Bound, John, Michael Lovenheim, and Sara Turner (2010): "Why have college completion rates declined? An analysis of changing student preparation and collegiate resources," *American Economic Journal: Applied Economics*, 2(3), 129–57.

[11] Bulman, George, Robert Fairlie, Sarena Goodman and Adam Isen (2016), Parental Resources and College Attendance: Evidence from Lottery Wins, NBER Working Paper No. w22679.

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

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

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

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

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

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

[18] Deming, David J. (2017) The Growing Importance of Social Skills in the Labor Market. *Quarterly Journal of Economics*, 132(4): 1593–1640.

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

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

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

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

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

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

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

[26] Lancaster, Tony (1990), The Econometric Analysis of Transition Data, Cambridge University Press

[27] Landerso, Rasmus and James J. Heckman (2017) The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the US, *Scandinavian Journal. of Economics* 119(1), 178–230, DOI: 10.1111/sjoe.12219

[28] Lochner, Lance J. and A. Monge-Naranjo (2011), The nature of credit constraints and human capital, *American Economic Review* 101(6), 2487-2529.

[29] Loken, Katrine (2007), Family Income and Children's Education: Using the Norwegian Oil Boom as a Natural Experiment, *Labour Economics*, 17, 118-129.

[30] Lovenheim, Michael and C. Lockwood Reynolds (2007), Changes in Postsecondary Choices by Ability and Income: Evidence from the National Longitudinal Surveys of Youth, *Journal of Human Capital*, 5, 70-109.

[31] Lovenheim, Michael F. and C. Lockwood Reynolds (2011), The Effect of Housing Wealth on College Choice: Evidence from the Housing Boom, *Journal of Human Resources* 48, 1-35.

[32] Nielsen, Eric R. (2015), The Income-Achievement Gap and Adult Outcome Inequality, Finance and Economics Discussion Series 2015-041. Washington: Board of Governors of the Federal Reserve System.

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

[34] Train, Kenneth (2008) EM Algorithms for nonparametric estimation of mixing distributions, *Journal of Choice Modelling* Vol 1, Issue 1, Pages 40-69.

Table 1: Summary Statistics in 1979 and 1997

	1979 cohort	1997 cohort
	Educational Attainment	
Highest grade completed	13.1	13.6
Proportion attended college	0.455	0.579
Proportion graduated from college	0.218	0.306
	Parental Income	
Average	$54,155	$65,572
Std dev	$33,308	$58,136
Average among college attendants	$65,319	$77,974
Highest grade completed by income quartile		
Quartile 1	12.0	12.4
Quartile 2	12.8	13.1
Quartile 3	13.3	14.1
Quartile 4	14.2	14.7
	AFQT	
Average	169	170.8
Std dev	30.2	30
Average among college attendants	185.8	182.2
	Other characteristics	
Male	0.494	0.499
Mother's education	11.7	13.2
Intact Family	0.788	0.581
Rural	0.238	0.264
Number of Siblings	3.1	2.3
Black	0.115	0.135
Hispanics	0.078	0.107
Mother's age at birth	26.3	25.8
Sample size	2,151	2,651

Note:
1979 income: $28,886 (25%); $50,163 (50%); $71,210 (75%)
1997 income: $29,623 (25%); $53,010 (50%); $82,285 (75%)
Attended college = 1 if completed grade 13 or more
Graduated college = 1 if completed grade 16 or more

Table 2: Marginal Effects of Observed Characteristics on HGC - Obtained from OLS

	1979 cohort		1997 cohort	
Variable	**Estimate**	**T-statistic**	**Estimate**	**T-statistic**
Mother's education	0.167	9.76	0.190	10.07
Intact Family	0.195	2.00	0.899	9.64
Rural	0.061	0.70	-0.015	-0.16
Number of Siblings	-0.098	-5.00	0.000	0.01
Black	1.118	8.51	0.835	6.27
Hispanic	0.643	4.28	0.252	1.74
Mother's age at birth	0.034	5.44	0.033	3.54
Male	-0.146	-1.99	-0.545	-6.47
AFQT	1.014	23.12	1.057	22.17
Family income	0.119	9.34	0.044	5.45
Mother's education	0.173	10.09	0.186	9.89
Intact Family	0.219	2.22	0.798	8.27
Rural	0.061	0.69	-0.023	-0.23
Number of Siblings	-0.097	-4.94	0.005	0.12
Black	1.124	8.43	0.869	6.53
Hispanic	0.657	4.34	0.278	1.93
Mother's age at birth	0.034	5.44	0.033	3.53
Male	-0.146	-1.99	-0.565	-6.72
AFQT	1.020	23.18	1.046	21.94
Family income - Q2	0.257	2.36	0.222	1.82
Family income - Q3	0.495	4.35	0.652	5.04
Family income - Q4	0.994	8.25	0.821	6.04
Sample size	2,151		2,651	

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

Table 3: Marginal Effects of Family Income and AFQT scores on Grade Transition Probabilities

	Income				AFQT			
	1979		1997		1979		1997	
	m.e.	std dev	m.e.	std dev	m.e.	std dev	m.e.	std dev
Age 18								
at Income Q1	0.022	0.003	0.009	0.004	0.026	0.003	0.021	0.004
at Income Q2	0.014	0.002	0.008	0.004	0.023	0.002	0.021	0.004
at Income Q3	0.008	0.002	0.007	0.004	0.021	0.002	0.021	0.003
at Income Q4	0.000	0.002	0.004	0.003	0.019	0.002	0.020	0.003
at Income mean	0.010	0.002	0.007	0.004	0.022	0.002	0.021	0.003
Age 20								
at Income Q1	0.015	0.005	0.006	0.005	0.050	0.005	0.019	0.005
at Income Q2	0.013	0.005	0.005	0.005	0.049	0.005	0.017	0.005
at Income Q3	0.011	0.005	0.004	0.005	0.049	0.005	0.016	0.006
at Income Q4	0.008	0.005	0.002	0.006	0.047	0.005	0.014	0.006
at Income mean	0.012	0.005	0.004	0.005	0.050	0.005	0.016	0.006
Age 22								
at Income Q1	0.006	0.003	0.001	0.004	0.035	0.003	0.010	0.004
at Income Q2	0.005	0.003	0.001	0.004	0.035	0.003	0.008	0.004
at Income Q3	0.004	0.003	0.000	0.004	0.034	0.003	0.008	0.004
at Income Q4	0.004	0.003	-0.001	0.004	0.034	0.003	0.007	0.004
at Income mean	0.005	0.003	0.000	0.004	0.035	0.003	0.008	0.004

Table 4: Marginal Effects of Family Income and AFQT scores on Educational Attainment

	Income				AFQT			
	1979		1997		1979		1997	
	m.e.	std dev	m.e.	std dev	m.e.	std dev	m.e.	std dev
Highest Grade Completed								
Age 18	0.021	0.004	0.017	0.010	0.042	0.004	0.053	0.010
Age 20	0.049	0.014	0.022	0.021	0.162	0.014	0.084	0.021
Age 22	0.065	0.022	0.027	0.030	0.254	0.022	0.116	0.030
Age 25	0.073	0.025	0.027	0.040	0.324	0.026	0.143	0.040
College Attendance								
Age 25	0.014	0.007	0.003	0.007	0.064	0.007	0.015	0.007

Table 5: Marginal Effects of Family Income and AFQT scores on Educational Attainment - Relative Income

	Income				AFQT			
	1979		1997		1979		1997	
	m.e.	std dev	m.e.	std dev	m.e.	std dev	m.e.	std dev
	Income Quartile 2 vs Income Quartile 1							
Grade Transition Probability								
Age 18	0.026	0.002	0.000	0.004	0.020	0.002	0.020	0.004
Age 20	0.032	0.004	-0.001	0.005	0.051	0.004	0.019	0.005
Age 22	0.016	0.003	-0.004	0.004	0.039	0.003	0.007	0.004
Highest Grade Completed								
Age 18	0.036	0.004	-0.001	0.011	0.018	0.003	0.056	0.010
Age 20	0.131	0.013	-0.003	0.021	0.171	0.013	0.095	0.021
Age 22	0.193	0.019	-0.012	0.029	0.305	0.019	0.130	0.029
Age 25	0.223	0.023	-0.035	0.036	0.399	0.023	0.163	0.036
College Attendance								
Age 25	0.058	0.007	-0.002	0.007	0.093	0.007	0.016	0.007
	Income Quartile 3 vs Income Quartile 1							
Grade Transition Probability								
Age 18	0.046	0.002	0.050	0.003	0.019	0.002	0.019	0.003
Age 20	0.046	0.004	0.022	0.005	0.049	0.004	0.019	0.005
Age 22	-0.002	0.003	0.002	0.004	0.037	0.003	0.006	0.004
Highest Grade Completed								
Age 18	0.054	0.003	0.141	0.010	0.013	0.002	0.055	0.010
Age 20	0.192	0.013	0.185	0.020	0.159	0.012	0.092	0.020
Age 22	0.260	0.019	0.214	0.029	0.289	0.019	0.126	0.029
Age 25	0.254	0.022	0.220	0.036	0.365	0.022	0.157	0.036
College Attendance								
Age 25	0.070	0.007	0.018	0.007	0.093	0.007	0.013	0.007
	Income Quartile 4 vs Income Quartile 1							
Grade Transition Probability								
Age 18	0.063	0.002	0.041	0.004	0.017	0.002	0.020	0.003
Age 20	0.111	0.004	0.040	0.005	0.047	0.004	0.019	0.005
Age 22	0.038	0.003	-0.004	0.004	0.038	0.003	0.006	0.004
Highest Grade Completed								
Age 18	0.065	0.003	0.114	0.010	0.010	0.002	0.056	0.010
Age 20	0.401	0.013	0.192	0.021	0.147	0.011	0.091	0.020
Age 22	0.636	0.019	0.230	0.029	0.287	0.018	0.117	0.029
Age 25	0.713	0.023	0.201	0.036	0.377	0.022	0.142	0.036
College Attendance								
Age 25	0.190	0.007	0.027	0.007	0.077	0.006	0.013	0.007

Table 6: Type Proportions and Location Parameters

	1979		1997	
Type	**Location**	**Proportion**	**Location**	**Proportion**
1	-2.0	0.515	-2.5	0.097
2	-1.8	0.000	-2.3	0.000
3	-1.6	0.000	-2.1	0.000
4	-1.4	0.001	-1.9	0.000
5	-1.2	0.054	-1.7	0.000
6	-0.8	0.186	-1.3	0.000
7	-0.6	0.000	-1.1	0.257
8	-0.4	0.242	-0.9	0.000
9	-0.2	0.000	-0.7	0.601
10	0.0	0.001	-0.5	0.045

Table 7: R-squares from Regressions of Type-specific Location Parameter and Highest Grade Completed

	Location		Highest Grade Completed	
	1979	**1997**	**1979**	**1997**
Mother's education	0.016	0.028	0.185	0.125
Intact Family	0.041	0.033	0.027	0.088
Rural	0.003	0.000	0.002	0.000
Number of siblings	0.005	0.004	0.060	0.011
Black	0.052	0.003	0.014	0.006
Hispanic	0.005	0.001	0.022	0.010
Mother's age at birth	0.006	0.009	0.017	0.050
Male	0.012	0.000	0.000	0.008
AFQT	0.075	0.058	0.343	0.262
Family income	0.075	0.027	0.162	0.076
Unobserved Heterogeneity			0.339	0.487
All of the above	0.150	0.089	0.643	0.659

R-squares from regressions where variables are individually considered.
Dependent variables in column heading. All regressions include initial grade.

Table 8: Marginal Effects of Family Income and AFQT scores on Educational Attainment - Comparison of Models

	Income				AFQT			
	1979		1997		1979		1997	
	m.e.	std dev	m.e.	std dev	m.e.	std dev	m.e.	std dev
				Single Type Model				
Grade Transition Probability								
Age 18	0.012	0.002	0.008	0.002	0.024	0.002	0.023	0.002
Age 20	0.023	0.004	0.016	0.003	0.075	0.004	0.058	0.003
Age 22	0.009	0.003	0.007	0.004	0.047	0.003	0.040	0.004
Highest Grade Completed								
Age 18	0.014	0.003	0.015	0.005	0.025	0.003	0.044	0.004
Age 20	0.080	0.013	0.057	0.013	0.244	0.012	0.196	0.012
Age 22	0.128	0.019	0.095	0.021	0.433	0.019	0.360	0.020
Age 25	0.147	0.023	0.115	0.027	0.542	0.023	0.506	0.027
College Attendance								
Age 25	0.038	0.007	0.023	0.006	0.127	0.007	0.081	0.005
				10-type Model				
Grade Transition Probability								
Age 18	0.010	0.002	0.007	0.004	0.022	0.002	0.021	0.003
Age 20	0.012	0.005	0.004	0.005	0.050	0.005	0.016	0.006
Age 22	0.005	0.003	0.000	0.004	0.035	0.003	0.008	0.004
Highest Grade Completed								
Age 18	0.021	0.004	0.017	0.010	0.042	0.004	0.053	0.010
Age 20	0.049	0.014	0.022	0.021	0.162	0.014	0.084	0.021
Age 22	0.065	0.022	0.027	0.030	0.254	0.022	0.116	0.030
Age 25	0.073	0.025	0.027	0.040	0.324	0.026	0.143	0.040
College Attendance								
Age 25	0.014	0.007	0.003	0.007	0.064	0.007	0.015	0.007

Table A1: Marginal Effects of Family Income and AFQT scores on Educational Attainment - Model without Non-cognitive Measures

	Income				AFQT			
	1979		1997		1979		1997	
	m.e.	std dev	m.e.	std dev	m.e.	std dev	m.e.	std dev
Grade Transition Probability								
Age 18	0.011	0.002	0.007	0.003	0.024	0.002	0.023	0.003
Age 20	0.014	0.005	0.006	0.005	0.046	0.005	0.022	0.005
Age 22	0.005	0.004	0.001	0.004	0.033	0.004	0.012	0.004
Highest Grade Completed								
Age 18	0.016	0.006	0.021	0.010	0.035	0.006	0.067	0.010
Age 20	0.041	0.016	0.035	0.021	0.104	0.016	0.117	0.021
Age 22	0.069	0.024	0.045	0.029	0.201	0.024	0.163	0.029
Age 25	0.078	0.028	0.051	0.037	0.295	0.028	0.213	0.037
College Attendance								
Age 25	0.012	0.006	0.007	0.007	0.031	0.006	0.027	0.007

Table A2: Marginal Effects of Family Income and AFQT scores on Educational Attainment - Trimmed Sample

	Income				AFQT			
	1979		1997		1979		1997	
	m.e.	std dev	m.e.	std dev	m.e.	std dev	m.e.	std dev
Grade Transition Probability								
Age 18	0.009	0.002	0.007	0.004	0.024	0.002	0.020	0.004
Age 20	0.014	0.004	0.004	0.005	0.062	0.004	0.012	0.005
Age 22	0.004	0.003	0.001	0.004	0.037	0.003	0.003	0.004
Highest Grade Completed								
Age 18	0.011	0.003	0.017	0.011	0.028	0.003	0.052	0.011
Age 20	0.046	0.013	0.027	0.021	0.188	0.013	0.079	0.021
Age 22	0.066	0.019	0.035	0.030	0.300	0.020	0.100	0.030
Age 25	0.073	0.023	0.040	0.040	0.371	0.023	0.111	0.040
College Attendance								
Age 25	0.018	0.007	0.004	0.007	0.093	0.007	0.017	0.007