Family Income and Higher Education Choices: The Importance of Accounting for College Quality

Author(s): Josh Kinsler and Ronni Pavan
Source: *Journal of Human Capital*, Vol. 5, No. 4 (Winter 2011), pp. 453-477
Published by: University of Chicago Press
Stable URL: http://www.jstor.org/stable/10.1086/663649
Accessed: 26-01-2016 13:52 UTC

# Family Income and Higher Education Choices: The Importance of Accounting for College Quality

Josh Kinsler
*University of Rochester*

Ronni Pavan
*University of Rochester*

In the examination of the determinants of educational choices, little attention has been devoted to the relationship between family income and the quality of higher education. Using the 1979 and 1997 waves of the NLSY, we show that family income significantly affects the quality of higher education, especially for high-ability individuals.While the impact of family income on college quality is significant in both samples, it has declined considerably over time for high-ability students. Overall, the trends we observe are highly consistent with increases in tuition across the quality spectrum, coupled with more generous merit-based aid at high-quality institutions.

## I. Introduction

It is a well-known result that the probability of attending a 2- or 4-year college varies significantly with both family income and individual ability.[1] Additionally, recent research illustrates that the attendance gaps across family income and ability groups have changed considerably from the 1980s to the 2000s. Belley and Lochner (2007) find a dramatic increase over time in the effect of family income on college attendance, and Lovenheim and Reynolds (2011) find that college nonattendance decreased substantially over time, particularly for high-ability students. The literature has focused extensively on the decision to attend college since the returns to obtaining a college degree are believed to be quite

[1] As an example, Ellwood and Kane (2000) use data from the High School and Beyond Survey and National Education Longitudinal Study of 1992 to show that 80 percent of the students from the top income quartile attended some type of postsecondary institution within 20 months of their high school graduation, as compared with 57 percent of those from the lowest income quartiles. The attendance gap is cut almost in half if math test scores from high school are taken into account. For additional examples, see Fuller, Manski, and Wise (1982), Carneiro and Heckman (2002), and Kane (2006).

[*Journal of Human Capital*, 2011, vol. 5, no. 4]

large. However, the 4-year college market is characterized by significant heterogeneity in quality and price, and this heterogeneity has been shown to have important effects on future wages (see, e.g., Dale and Krueger 2002; Black and Smith 2004, 2006; Hoekstra 2009). Yet, there is very little research regarding the effect family income and individual ability have on the quality of 4-year college a student initially attends or graduates from.[2]

Expanding our knowledge about how individuals choose college quality and how the factors affecting this choice have changed over time is important for evaluating and designing educational policies aimed at affecting the distribution of wealth in an economy. The cost of attending a 4-year college has increased dramatically over the past few decades, particularly at the top of the college quality distribution.[3] One might suspect that as a result the effect of family income on college quality has increased over time, indicating that quality is less accessible for the current generation of college-goers. However, the amount of financial aid and grants offered by top institutions has also increased significantly.[4] Thus, it is unclear whether family income has become more or less important in determining a student's college choice. Given the recent results indicating that higher-quality schools lead to higher wages, omitting college quality when analyzing the impact of family income on schooling outcomes will likely understate the role of education in generating persistent income inequality across generations.

In this paper, we use the National Longitudinal Surveys of Youth (NLSY) to examine how family income and individual ability affect college quality at the time of both initial attendance and graduation in the early 1980s and early 2000s. Using a series of straightforward multinomial logit models, we find positive and significant effects of family income on college quality in both the 1979 and 1997 cohorts of the NLSY. The importance of family income varies considerably across the ability spectrum, where controls for ability are constructed using the Armed Services Vocational Aptitude Battery (ASVAB) tests present in the NLSY. For example, in the 1979 cohort, a move from the first to the fourth quartile of family income increases the probability of attending a top-quartile college from 0.06 to 0.40 for high-ability individuals and from

[2] Exceptions include Fuller et al. (1982), Long (2004), and Lovenheim and Reynolds (2010).

[3] According to the College Board (http://trends.collegeboard.org/), average tuition at public 2-year colleges, public 4-year colleges, and private 4-year colleges has increased from 1980 to 2010 by approximately $2,000, $6,000, and $23,000, respectively (inflation adjusted).

[4] Although average published tuition and fees increased by about 15 percent in inflation-adjusted dollars at private not-for-profit 4-year colleges and universities from 2004–5 to 2009–10 and by about 20 percent at public 4-year institutions, the estimated average 2009–10 net price for full-time students, after considering grant aid and federal tax benefits, is about $1,100 lower (in 2009 dollars) in the private sector and about $400 lower in the public sector than it was 5 years ago. Information is obtained from http://trends.collegeboard.org/.

0.01 to 0.05 for average-ability individuals. Interestingly, the same movement across the family income distribution increases the probability of attending any 4-year college from 0.62 to 0.82 for high-ability individuals and from 0.23 to 0.30 for average-ability individuals. These numbers indicate that the impact of family income is highly concentrated at the top of the college quality distribution, particularly for high-ability students.

By utilizing both waves of the NLSY, we are also able to examine how the relationship between family income and college quality has changed over time. For average-ability individuals, the relationship between family income and college quality has strengthened slightly, though the change is not statistically significant. Conversely, the impact of family income on the quality of the initial college attended has decreased significantly for high-ability students. The importance of family income on the quality of the degree-granting school has also declined over time for high-ability individuals, though the change is not statistically significant. In general, the changes we observe over time in the importance of family income for the quantity and quality of schooling are highly consistent with an overall increase in tuition coupled with more generous merit aid at the highest-quality institutions. Family income has become more important for average-ability individuals at the extensive schooling margin, whereas family income has become less important for high-ability individuals along the college quality margin.

In Section III of the paper, we perform some simple back-of-the-envelope calculations to compare the effect family income has on the quantity and quality of education. To do this, we first estimate a set of Mincer wage regressions that allow the returns to college to differ with quality. We then combine our estimated returns to schooling with the predicted probabilities from our multinomial logit models to predict how wages change as individuals move from the first to the fourth quartile of family income. We find that for high-ability individuals, accounting for heterogeneity in college quality increases the dollar impact of a move from the first to the fourth quartile of family income by 31 percent in 1979 and 54 percent in 1997 relative to a model that treats 4-year colleges as homogeneous. For average-ability individuals, accounting for college quality does not appreciably alter the effect of a move from the first to the fourth quartile of family income.

In our analysis we consistently find that the importance of family income is relatively more pronounced in the quantity dimension for average-ability students, whereas it is much more evident in the quality dimension for high-ability students. This difference has important implications for policy since quality and tuition are highly correlated. Consider, for example, the federal Pell Grant, which has a maximum benefit of approximately $5,500 per year. This program can erase much of the impact of family income on the quantity of schooling for low-income, average-ability students since there are many colleges that cost less than

the maximum benefit. However, for high-ability, low-income students, the impact of family income on schooling outcomes will not be eliminated by such a program since the grant generally accounts for only a small fraction of tuition at high-quality institutions.

As previously mentioned, direct empirical evidence relating family income and college quality is rather sparse, though a number of studies have examined how sensitive students are to price when choosing between colleges. Fuller et al. (1982) find that students from high-income families are less price sensitive, and Long (2004) finds that low-income students in 1992 are as negatively affected by price as students in 1972. The paper that is most similar to ours is Lovenheim and Reynolds (2010). Using variation in housing wealth generated by the timing of the housing bubble, the authors find that less financially constrained families tend to purchase higher-quality schools. Our paper adds to this literature by examining directly how family income and college quality are related, in both the early 1980s and the 2000s.

The remainder of the paper is organized as follows. Section II discusses the data and sample restrictions in detail. The results of the analysis are contained in Section III. Section IV concludes with a brief discussion of how our work relates to the broader credit constraint literature.

## II. Data

For our analysis, we turn to the 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. These individuals were interviewed annually through 1994 and are currently interviewed on a biennial basis. Similarly, the NLSY97 consists of a nationally representative sample of approximately 9,000 youths who were 12–16 years old as of December 31, 1996. Youths continue to be interviewed on an annual basis.

When we are examining schooling and labor market outcomes, our primary variables of interest are family income and student ability. Because high-quality schools tend to accept only high-ability students, it is imperative to control for student ability when estimating the effect of family income on college quality. Also, by controlling directly for student ability, we are able to isolate the role of family income in the short-term decisions of whether to attend college and which type of school to attend.[5] It is this short-run relationship that is the target of both federal and institutional need-based aid.

[5] In order to ensure that variation in family income reflects primarily differences in financial situations across families, we also control for a host of cultural and geographical factors that are likely to affect schooling decisions. In particular, we include controls for race, gender, education of the mother, education of the father, age of the mother at birth,

Both NLSY samples contain measures of family income and individual ability. Each year respondents are asked about total household income in the previous year.[6] Using this variable, we construct average household income for each respondent when he or she was 16 and 17 years of age.[7] Ability is captured by the ASVAB, which is a sequence of tests that cover basic math, verbal, and manual skills. Rather than use a single ability measure, such as the Armed Forces Qualifications Test, we construct separate measures of math and verbal ability for each respondent. Math ability is measured by averaging the scores on the Arithmetic Reasoning and Mathematics Knowledge portions of the ASVAB. Verbal ability is measured by averaging the scores on the Word Knowledge and Paragraph Comprehension portions of the ASVAB. Although the ASVAB tests are noisy measures of ability, they have been proven to be excellent controls with high explanatory power and, as a consequence, have been utilized often in this literature (see, e.g., Carneiro and Heckman 2002; Belley and Lochner 2007; Lovenheim and Reynolds 2011).

Because family income and student ability are utilized in all dimensions of the subsequent analysis, our primary sample selection criteria are driven by these two variables. In the NLSY79, many respondents are more than 18 years of age at the time of the first survey. This creates two issues. First, contemporaneous family income may not reflect constraints at the time college decisions were initially made, and for individuals already living outside the home, no measure of parental income is available. Second, ASVAB test scores for students older than 18 are endogenous to schooling choices. This will tend to overstate the importance of ability in predicting college graduation and college quality. To avoid these problems, we exclude anyone who is 19 or older in the first survey year, leaving four cohorts of respondents between the ages of 15 and 18 in 1979. Eliminating cohorts in the NLSY79, however, has implications for comparisons with the NLSY97 since it has five cohorts of respondents. Any comparison of the restricted NLSY79 sample with the full NLSY97 sample will be contaminated by differences in the age

number of siblings, a rural indicator, and an indicator for whether the family is intact. Geographical controls include the number of 2-year colleges, 4-year colleges, and 4-year colleges with average freshman Scholastic Aptitude Test (SAT) scores above 1,200 in the respondent's home state.

[6] Note that for young respondents, parents were also asked to report household income. For NLSY79 respondents born after 1961, 97 percent (90 percent) of the 1979 (1980) family income data were collected from parent interviews. In the NLSY97, parents of dependent children directly reported last year's income only in 1997. However, our constructed variable does not necessarily rely on this information since we utilize the income information when the respondent was 16 or 17. In order to test the robustness of our findings, we replicate our results using the 1997 information when available. The results were hardly affected.

[7] If family income is missing for one of these years, we simply use the available measure. If family income is not available at either 16 or 17 years of age, we sequentially look at family income from earlier ages to avoid dropping respondents. Respondents who report not living with their parents at these ages are excluded.

distribution of respondents. As a result, we eliminate the youngest cohort from the NLSY97 sample. In summary, only respondents born between 1961 and 1964 in the NLSY79 and respondents born between 1980 and 1983 in the NLSY97 remain in the sample. The oversample of minority groups is included in both the NLSY79 and NLSY97 cohorts.[8]

We construct variables to describe both college attendance and college graduation. We first determine whether the respondent ever attended college using responses regarding enrollment and then determine whether the respondent received a bachelor's degree, restricting ourselves to educational outcomes prior to 1989 for the NLSY79 cohort. The restriction for the earlier cohort is to ensure that across the samples, individuals have had the same amount of time for their schooling choices. When we utilize education as an input into wages, we use the highest degree received by the year in which the wage was earned.[9]

In addition to enrollment and degree receipt, we need information about which school the respondent attended and graduated from if we are going to analyze the impact of family income on college quality. Using the NLSY79 and NLSY97 geocode data, we are able to identify the specific colleges students attend. School identifiers are available for all years in the NLSY97 and are available beginning in 1984 in the NLSY79. In 1984, respondents are asked to list the most recent colleges attended, and we map these schools back to enrollment and degree receipts prior to 1984.[10] We then match the school identifiers available in the NLSY79 and NLSY97 to the Integrated Postsecondary Education Data System (IPEDS), which has a plethora of institutional information available, such as the type of school, location, faculty characteristics, and student characteristics.[11]

[8] Our results are robust to the exclusion of these samples. Results are available on request.

[9] For both the NLSY79 and NLSY97, we rely primarily on the highest degree received variable to determine whether a respondent eventually graduated from college. This variable becomes available in the NLSY79 only in 1988. At this time respondents are asked to report the year in which they received the highest degree. Because of attrition and missing variables, we use additional information obtained between 1979 and 1987 about reported college degrees and enrollment status to pin down the degree year. In the NLSY97, degree assignment is based on changes in the highest degree received variable. Note that when estimating the returns to schooling, we exclude individuals who have obtained advanced degrees.

[10] The school identifier associated with the first report of college enrollment is identified as the initial college attended. We assign school identifiers for the degree-granting schools across both samples in the following manner. If a student reports receiving a degree and is no longer enrolled, the school identifier of the most recent school is assumed to be the degree-granting school. If the respondent receives a degree and remains in school, we use the school identifier of the second most recent school as the degree-granting institution.

[11] Freshman SAT/ACT test score distributions and retention rates, two of the most widely used measures of school quality, are not available in IPEDS until 2001. To provide a more accurate measure of school quality for the NLSY79, we also collect data from the 1980 Barron's and Peterson's college guides. In 1980, typically the average freshman SAT/ACT scores are reported. IPEDS, however, reports the 25th and 75th percentiles of the freshman test score distributions in 2001 and beyond. To make these comparable, we

Our primary measure of school quality is the principal factor generated when using total SAT score, freshman retention rate, and faculty salary as noisy measures of unobserved quality.[12] Many schools, particularly in 1980, required students to complete just the ACT exam. Because we want to include as many schools as possible in our factor analysis, we use the SAT/ACT conversion tables provided by the Princeton Review to convert our ACT scores into SAT scores. For schools lacking an SAT/ACT measure, we impute quality by first regressing the estimated principal factors on a flexible function of freshman retention rate and faculty salary. We then predict what the principal factor is for those schools missing SAT/ACT scores. A similar procedure is utilized to impute quality for schools missing either freshman retention rate or faculty salary. Quality is imputed for 11 percent and 14 percent of schools as a result of missing either SAT/ACT scores, freshman retention rates, or faculty salary in the NLSY79 and NLSY97 samples, respectively. There are a handful of schools in both samples that are missing multiple measures of underlying school quality. We use the college guides in 1980 and 2001 to assign a quality quartile, which is described in greater detail below. Six percent and 2 percent of schools are assigned a quality quartile utilizing this method in the NLSY79 and NLSY97 samples, respectively.[13]

The final component of the data is labor market outcomes. We focus on hourly wages as the primary variable of interest. The NLSY79 and NLSY97 collect information on total labor income during the previous year. Using information on total hours, we are able to construct an hourly wage measure. We consider wages and incomes only for those with positive hours. Wages are deflated in year 2000 dollars, and we trim below \$1 and above \$1,000.[14]

In all of the analysis to follow, the oversamples of minority groups are included. As a result, in all education-related regressions, each observation is weighted according to the sample weights provided by the NLSY. For labor market outcomes, we weight individual wage observations according to the sample weights provided by the NLSY and the number of hours worked in a particular year. Weighting by the number of hours worked emphasizes the returns to schooling for those workers who are employed full-time.

In our schooling regressions we control for family income and ability

simply take the average of the 25th and 75th percentile scores when constructing college quality for the NLSY97 cohort.

[12] This approach is similar to the one employed by Black and Smith (2004). The authors note that additional or alternative measures lead to little or no change in the quality index. Note that we construct quality measures only for 4-year institutions and abstract from any quality differentials in 2-year schools.

[13] We are unable to assign any quality measure to 2 percent and 6 percent of schools in the NLSY79 and NLSY97. The primary reason for this is the reporting of invalid school codes in both NLSY samples.

[14] The estimated returns to quality are not sensitive to this trimming.

quartiles rather than include simple linear terms. Grouping the explanatory variable into quartiles has two positive effects. First, it helps identify nonlinearities in the impact of the variables on schooling choice. Second, it may help reduce the bias induced by the presence of measurement error.[15] When assigning individuals to quartiles of family income, we use the unconditional population distribution of earnings according to the 1980 and 2000 U.S. censuses to determine the appropriate cutoffs. We use this external source to ensure that any random sampling variation across the two samples in the distribution of family income does not bias the results. Ability quartiles are birth year specific to account for age differences at the time of the test. Finally, we also assign colleges to quartiles of the quality distribution.[16] The quartiles of college quality are generated with respect to the set of colleges respondents graduated from in each sample of the NLSY. When we construct the quartiles, colleges are weighted by the number of NLSY respondents who graduated from each school. For schools that respondents only attended but never graduated from, we use the cutoffs generated from the aforementioned procedure to assign a quality quartile.[17] For a sense of the difference in college quality across the quartiles, consider that the median first-quartile quality schools in 1979 and 1997 are Sam Houston State University and South Carolina State University, and the median fourth-quartile quality schools in 1979 and 1997 are New York University and the University of Texas at Austin.

Basic summary statistics for both the NLSY79 and NLSY97 respondents are shown in table 1. In terms of family background variables, respondents are less likely to grow up in an intact family, have fewer siblings, and have more educated parents in the NLSY97 cohort relative to the NLSY79 cohort. In real terms, families are wealthier in 1997, though the dispersion in family income has increased significantly. Note that the increased dispersion in family income evident in the NLSY is also present in the census. This implies that the income gap between the first and fourth quartiles is larger in the NLSY97 than it was in the NLSY79. Respondents in the NLSY97 cohort are more likely to drop

[15] As noted earlier, family income data are collected from both parents and respondents and, as such, are likely contaminated by measurement error. However, as long as measurement error does not cause a lot of misclassification, utilizing quartiles will reduce the bias. This strategy is very similar to the one proposed by Wald (1940). Note that when we use income reports using parental responses whenever possible, the results are essentially unchanged.

[16] Utilizing quartiles of college quality helps mitigate any remaining measurement error in our quality index. Note that our first step to reducing measurement error in college quality is to use multiple underlying measures of quality rather than a single measure, such as SAT scores.

[17] Note that we do not use these schools when generating the quartile cutoffs since they tend to be heavily skewed toward the bottom of the quality distribution. This would lead to an understatement of the effect of family income when examining quality choices for college graduates since their inclusion would substantially lower the average quality of the schools in the highest quartile.

TABLE 1
Summary Statistics, Respondents

| | NLSY79 | NLSY97 |
|---|---|---|
| Family background: | | |
| Urban | .768 | .780 |
| Intact family | .716 | .545 |
| Siblings | 3.219 | 2.303 |
| Hispanic | .067 | .113 |
| Black | .147 | .151 |
| Female | .487 | .501 |
| Mother's age at birth | 25.591 | 26.388 |
| Mother high school dropout | .326 | .153 |
| Mother high school graduate | .470 | .364 |
| Mother some college | .204 | .482 |
| Father high school dropout | .335 | .167 |
| Father high school graduate | .369 | .368 |
| Father some college | .296 | .465 |
| Family income (2000 $) | 48,574 | 64,548 |
| Standard deviation of family income | 30,815 | 57,743 |
| State-level geographic controls: | | |
| Unemployment rate | .078 | .044 |
| Number of 2-year colleges in state | 47.147 | 85.390 |
| Number of 4-year colleges in state | 74.578 | 104.689 |
| Number of 4-year above 1,200 SAT | 6.859 | 6.681 |
| Attainment: | | |
| High school dropout | .124 | .090 |
| GED | .058 | .046 |
| High school graduate | .265 | .164 |
| College dropout | .289 | .355 |
| 2-year college | .063 | .066 |
| 4-year college or more | .201 | .279 |
| Labor market outcomes: | | |
| Wage | 10.397 | 10.722 |
| Standard deviation of wage | 6.483 | 8.123 |
| Log wage | 2.179 | 2.186 |
| Standard deviation of log wage | .591 | .624 |
| Sample size | 5,161 | 4,707 |

Note.—Samples are restricted to respondents born between 1961 and 1964 in the NLSY79 and respondents born between 1980 and 1983 in the NLSY97 who have valid income and ability measures. Observations are weighted using the sampling weights provided by the NLSY. Statistics for the 1979 cohort use information only through 1989, allowing for simple comparisons between the two samples.

out and graduate from college relative to the NLSY79 cohort.[18] The patterns of wages across the two samples are similar to the pattern of family income, increases in both the levels, and dispersion in 1997.[19]

Tables 2 and 3 provide detailed information about college quality and tuition across the NLSY79 and NLSY97 samples. Table 2 illustrates that average SAT and ACT scores among incoming freshmen have improved, and tuition, as previously noted, increased dramatically in both level

[18] College dropouts are defined as individuals who were ever enrolled in college, either 2- or 4-year, but never report earning a degree.

[19] Family income and wages, as well as any other monetary measure utilized throughout the paper, are measured in 2000 dollars.

TABLE 2
Summary Statistics, Schools

| | 1979 | 1997 |
|---|---|---|
| ACT | 20.73 | 23.06 |
| | (3.00) | (3.01) |
| SAT math | 509.43 | 553.72 |
| | (68.75) | (62.13) |
| SAT verbal | 470.20 | 545.03 |
| | (62.62) | (56.77) |
| Log faculty salary | 10.79 | 10.98 |
| | (.18) | (.22) |
| Freshman retention rate | 76.57 | 78.95 |
| | (12.99) | (10.56) |
| Tuition | 5,665 | 11,485 |
| | (2,559) | (5,181) |
| Schools | 796 | 1,497 |

Note.—The sample of colleges in each cohort reflects 4-year colleges that NLSY respondents ever graduated from. College data for the 1979 cohort are obtained from IPEDS and Barron's and Peterson's college guides. College data for the 1997 cohort are obtained entirely from IPEDS. Schools are population weighted.

TABLE 3
Correlates with Principal College Quality Measure

| | ACT | SAT Math | SAT Verbal | ln(Faculty Salary) | Retention | Tuition | Quality |
|---|---|---|---|---|---|---|---|
| | A. NLSY79 ($N$ = 313) | | | | | | |
| ACT | 1 | | | | | | |
| SAT math | .820 | 1 | | | | | |
| SAT verbal | .790 | .888 | 1 | | | | |
| ln(faculty salary) | .347 | .415 | .348 | 1 | | | |
| Retention | .557 | .485 | .440 | .101 | 1 | | |
| Tuition | .581 | .553 | .632 | .041 | .44 | 1 | |
| Quality factor | .812 | .895 | .848 | .624 | .686 | .530 | 1 |
| | B. NLSY97 ($N$ = 1,181) | | | | | | |
| ACT | 1 | | | | | | |
| SAT math | .910 | 1 | | | | | |
| SAT verbal | .914 | .957 | 1 | | | | |
| ln(faculty salary) | .543 | .657 | .583 | 1 | | | |
| Retention | .718 | .749 | .711 | .685 | 1 | | |
| Tuition | .568 | .550 | .592 | .328 | .458 | 1 | |
| Quality factor | .839 | .910 | .873 | .825 | .929 | .526 | 1 |

Note.—Quality factor is the principal factor generated when using total SAT score, freshman retention rate, and faculty salary as noisy measures of unobserved quality. For further details, see the text. The sample of colleges in each cohort reflects 4-year colleges that NLSY respondents ever graduated from. College data for the 1979 cohort are obtained from IPEDS and Barron's and Peterson's college guides. College data for the 1997 cohort are obtained entirely from IPEDS. Schools are population weighted.

and variability. Table 3 shows how various noisy measures of college quality correlate with our principal factor measure of quality and with tuition. In the NLSY79, average test scores have the highest correlation with our quality measure, whereas in the NLSY97, the freshman retention rate has the highest correlation with our quality measure. Note that in both samples, quality and tuition are highly correlated, suggesting that family income may play a significant role in the choice of college quality. One interesting result in table 3 is that there is a significant increase in the correlation between SAT scores and faculty salary over time. This change is consistent with increased ability sorting across colleges, as suggested by Hoxby (2009). It could also indicate that our measures of college quality are more precise in the NLSY97. Both explanations would likely lead us to understate the change in the impact of family income on college quality. In the case of stronger sorting, the true difference in quality between a first- and fourth-quartile school has expanded, meaning that the smaller effect of family income in the NLSY97 is even more surprising. Measurement error in college quality will result in an understatement of the impact of family income on college quality as a result of misclassification.[20] To the extent that there is more measurement error in the NLSY79, the change in the impact of family income will be attenuated.

Finally, table 4 provides additional evidence that family income and educational attainment, including college quality, are closely related. The table shows the distribution of educational attainment conditional on graduating from high school for each quartile of family income in the NLSY79 and NLSY97 samples. In both samples, individuals from the lowest quartile of family income have the highest probability of having a high school diploma and the lowest probability of graduating from a top-quartile college. Individuals from the fourth quartile of family income display the exact opposite pattern. Across the samples, there is a clear indication that the probability of graduating from a 4-year college has increased for all quartiles of family income. It is also true that the correlation between family income and graduating from a high-quality school has decreased slightly. The usefulness of table 4 in identifying the effect of family income on attainment is limited, however, since family income is likely correlated with many other factors that affect schooling decisions. The next section addresses this issue by examining the relationship between schooling choices and family income conditional on ability and a host of other background characteristics.

[20] Note that when we construct quality quartiles, the measurement error in our outcome is no longer classical. For example, a fourth-quartile college can be misclassified only as a lower-quality school.

TABLE 4
Relationship between Family Income and Education

| | Quartile of Family Income | | | |
|---|---|---|---|---|
| | First | Second | Third | Fourth |
| | A. NLSY79 | | | |
| High school | .849 | .749 | .676 | .488 |
| 2-year college | .054 | .073 | .093 | .058 |
| 4-year college: | | | | |
| First quartile | .033 | .069 | .058 | .072 |
| Second quartile | .025 | .039 | .060 | .097 |
| Third quartile | .021 | .038 | .048 | .114 |
| Fourth quartile | .015 | .027 | .059 | .149 |
| Not classified | .003 | .005 | .006 | .021 |
| *N* | 1,611 | 1,196 | 903 | 549 |
| | B. NLSY97 | | | |
| High school | .805 | .717 | .582 | .450 |
| 2-year college | .061 | .067 | .085 | .064 |
| 4-year college: | | | | |
| First quartile | .050 | .070 | .075 | .090 |
| Second quartile | .026 | .052 | .102 | .100 |
| Third quartile | .033 | .047 | .080 | .132 |
| Fourth quartile | .022 | .042 | .070 | .151 |
| Not classified | .004 | .006 | .007 | .013 |
| *N* | 1,159 | 1,109 | 1,126 | 830 |

Note.—Samples are restricted to respondents born between 1961 and 1964 in the NLSY79 and respondents born between 1980 and 1983 in the NLSY97 who obtained at least a high school diploma and have valid income and ability measures. Observations are weighted using the sampling weights provided by the NLSY. Statistics for the 1979 cohort use information only through 1989, allowing for simple comparisons between the two samples. See the text for a discussion of how the quartiles of family income and college quality are constructed.

## III. Results

In this section we analyze the impact of family income on the quality and quantity of education obtained, highlighting how the introduction of quality enhances our understanding of the college attendance decision. At the end of the section we also provide a simple back-of-the-envelope calculation that allows us to compare the impact family income has on the quantity and quality margins of higher education. The metric that we utilize to compare the effect of family income across the two schooling margins is the differential in predicted wages across workers with different educational backgrounds.

### A. *Effect of Family Income on Quality Conditional on Attendance and Graduation*

Our analysis of the relationship between schooling and family income begins with the effect family income has on college quality. We first examine how family income affects the quality of the initial college

TABLE 5
Family Income and Initial College Quality, Multinomial Logit

| | A. NLSY79: Quartile of College Quality | | | B. NLSY97: Quartile of College Quality | | |
|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 2 | 3 | 4 |
| Income quartile 2 | −.520 | −.184 | 1.068** | .431 | .052 | .230 |
| | (.307) | (.308) | (.542) | (.271) | (.300) | (.298) |
| Income quartile 3 | −.328 | −.701* | 1.578*** | .552** | .385 | .318 |
| | (.386) | (.386) | (.508) | (.245) | (.282) | (.312) |
| Income quartile 4 | −.264 | .630 | 2.261*** | .783*** | .750* | 1.107*** |
| | (.453) | (.397) | (.624) | (.277) | (.401) | (.317) |
| Income quartile 4 × high ability | .511 | −.198 | .590 | .040 | −.058 | .267 |
| | (.462) | (.567) | (.578) | (.419) | (.422) | (.390) |
| Math quartile 2 | −.089 | 1.021** | .307 | .630* | .663 | .431 |
| | (.380) | (.494) | (.694) | (.346) | (.455) | (.671) |
| Math quartile 3 | .130 | .630 | .247 | .838** | 1.399*** | .955 |
| | (.375) | (.598) | (.798) | (.339) | (.475) | (.610) |
| Math quartile 4 | .347 | 1.571** | 1.477* | 1.270*** | 2.451*** | 2.320*** |
| | (.423) | (.639) | (.781) | (.328) | (.479) | (.595) |
| Verbal quartile 2 | .606 | .068 | .418 | −.289 | .028 | −.286 |
| | (.477) | (.422) | (.995) | (.347) | (.466) | (.609) |
| Verbal quartile 3 | .871* | .508 | .719 | .237 | .388 | .659 |
| | (.529) | (.451) | (1.028) | (.343) | (.481) | (.607) |
| Verbal quartile 4 | 1.069** | 1.041** | 1.459 | .162 | .661 | 1.083* |
| | (.528) | (.451) | (1.003) | (.374) | (.495) | (.624) |
| Additional controls | Y | Y | Y | Y | Y | Y |
| Observations | | 1,162 | | | 1,574 | |

Note.—Respondents are clustered at the primary sampling unit, and we utilize robust standard errors. The dependent variable indicates the quality of the initial college the respondent attended measured in quartiles. The first quartile of college quality is the excluded category. High ability indicates that a respondent is in the fourth quartile of both math and verbal ability. Observations are weighted using the sampling weights provided by the NLSY. In order for the samples to be comparable, only schooling choices through 1989 are considered for the NLSY79 sample. Additional controls include indicator variables for urban, intact family, race, gender, mother (father) is a high school graduate, mother (father) has some college, and continuous controls for mother's age, number of siblings, unemployment rate when exiting high school, and the number of 2-year, 4-year, and 4-year colleges with average freshman SAT scores above 1,200 in the state.
* $p < .1$.
** $p < .05$.
*** $p < .01$.

attended conditional on ever attending a 4-year school. Then, we examine the impact family income has on the quality of the degree-granting institution for those respondents who receive a 4-year college degree. We start by examining attendance since much of the previous literature has focused on the impact family income has on the likelihood of attending a 4-year college. We then shift to graduation since the labor market returns to higher education are typically associated with degree receipt. It is also simpler to "monetize" the impact of family income on schooling when education is defined according to degree receipt.

Table 5 lists the coefficients from a multinomial logit model in which respondents choose among the four quartiles of college quality, conditional on ever attending a 4-year college. Panels A and B of the table show results for the NLSY79 and NLSY97 samples, respectively. When

estimating the impact of family income on the quality of the first college ever attended, we control for quartiles of math and verbal ability, a host of family background characteristics, and a set of supply-side variables, such as the number of 4-year colleges with average freshman SAT scores above 1,200 in the respondent's home state. A full list of controls can be found both in note 5 and in the note to table 5. Additionally, we include an interaction term that indicates that the respondent is from the top quartile of family income, math, and verbal ability. While the nonlinearity of the multinomial logit specification will allow the effect of changes in family income to vary across the ability distribution, we include a direct interaction for the highest-income and highest-ability respondents for additional flexibility.[21] Table 6 has a structure identical to table 5, the only difference being that in table 6 the outcome is college quality conditional on graduating from a 4-year college. All regressions are clustered at the primary sampling unit level, and we use a heteroskedastic-robust estimator for the variance.

The results in tables 5 and 6 indicate that family income and ability have a significant effect on the quality of the first school a student ever attends in the early 1980s and 2000s and also on the quality of the school a student graduates from in the early 1980s. However, it is difficult to interpret the findings from tables 5 and 6 as a result of the nonlinearity of the multinomial logit model. In addition, we would expect the effect of family income to vary considerably across the ability spectrum since low-ability students are unlikely to either attend or graduate from a top-quartile school regardless of family income. To provide some greater context, table 7 shows some predicted probabilities for the average high-ability and the average respondent in each sample as we vary the level of family income. Panel A of the table shows predicted probabilities for school quality when initially attending (based on the results from table 5), and panel B shows predicted probabilities for school quality when graduating (based on the results from table 6).

The results in panel A of table 7 are quite striking. For the average high-ability college attendee in the 1979 cohort, the probability of attending a top-quartile school changes from 0.074 to 0.511 as family income changes from the first to the fourth quartile.[22] This equates to a marginal effect of 0.437, which is both economically and statistically significant. Even when we examine the average college attendee in the 1979 sample, the probability of attending a top-quartile college changes by 0.253 as family income changes from the first to the fourth quartile. The impact of family income on the likelihood of attending a top-quartile college in the 1997 cohort is also positive and statistically sig-

[21] We have also completed all of the subsequent analysis eliminating any student whose math and verbal abilities are below the median. The results are not sensitive to this restriction. Results are available on request.

[22] High-ability individuals are those in the top quartile of both math and verbal ability. In 1979 (1997) this group accounts for 17 percent (14 percent) of our sample.

TABLE 6
Family Income and Quality of Degree-Granting School, Multinomial Logit

| | A. NLSY79: Quartile of College Quality | | | B. NLSY97: Quartile of College Quality | | |
|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 2 | 3 | 4 |
| Income quartile 2 | −.267 | −.096 | −.005 | .298 | −.245 | −.285 |
| | (.464) | (.448) | (.540) | (.355) | (.318) | (.322) |
| Income quartile 3 | .486 | .244 | 1.117** | .690** | .271 | .245 |
| | (.470) | (.461) | (.559) | (.339) | (.348) | (.307) |
| Income quartile 4 | .637 | 1.112** | 1.262** | .445 | .475 | .459 |
| | (.542) | (.485) | (.541) | (.370) | (.404) | (.376) |
| Income quartile 4 × high ability | −.414 | −.814 | .179 | −.046 | .209 | .512 |
| | (.515) | (.528) | (.435) | (.467) | (.429) | (.418) |
| Math quartile 2 | .901 | .268 | 2.572* | 1.384*** | .614 | −.037 |
| | (.841) | (.998) | (1.481) | (.484) | (.547) | (.827) |
| Math quartile 3 | .876 | .443 | 3.211** | 1.072** | 1.299*** | .678 |
| | (.730) | (.970) | (1.503) | (.489) | (.502) | (.665) |
| Math quartile 4 | .997 | 1.218 | 3.649** | 1.113** | 1.865*** | 1.522** |
| | (.759) | (.879) | (1.481) | (.516) | (.494) | (.639) |
| Verbal quartile 2 | .215 | −1.038 | −.995 | −.592 | .091 | .620 |
| | (.751) | (.929) | (.972) | (.442) | (.675) | (.947) |
| Verbal quartile 3 | .173 | −.822 | −1.339 | −.019 | .497 | 1.601* |
| | (.782) | (.952) | (1.018) | (.472) | (.647) | (.949) |
| Verbal quartile 4 | .233 | −.407 | −.763 | −.246 | .568 | 1.890** |
| | (.789) | (.943) | (.982) | (.503) | (.645) | (.949) |
| Additional controls | Y | Y | Y | Y | Y | Y |
| Observations | | 714 | | | 1,118 | |

Note.—Respondents are clustered at the primary sampling unit, and we utilize robust standard errors. The dependent variable indicates the quality of the college the respondent graduated from measured in quartiles. The first quartile of college quality is the excluded category. High ability indicates that a respondent is in the fourth quartile of both math and verbal ability. Observations are weighted using the sampling weights provided by the NLSY. In order for the samples to be comparable, only schooling choices through 1989 are considered for the NLSY79 sample. Additional controls include indicator variables for urban, intact family, race, gender, mother (father) is a high school graduate, mother (father) has some college, and continuous controls for mother's age, number of siblings, unemployment rate when exiting high school, and the number of 2-year, 4-year, and 4-year colleges with average freshman SAT scores above 1,200 in the state.
* $p < .1$.
** $p < .05$.
*** $p < .01$.

nificant for both the average high-ability and the average attendees. However, the marginal effects are significantly smaller relative to the 1979 cohort. For high-ability individuals, the impact of a move from the first to the fourth quartile of family income decreases over time by 0.251 with a *t*-statistic equal to 2.73. For the average student, the decrease is 0.136 with a *t*-statistic equal to 2.43. The approximately fourfold increase in the probability of attending a fourth-quartile college when family income is low is the reason for the significant decline over time in the importance of family income. This change is largely consistent with the recent expansion of need-based aid at many top institutions.

Panel B of table 7 illustrates the predicted probabilities of graduating from a top-quartile college for the average high-ability college graduate and the average college graduate when family income is either low or

TABLE 7
Predicted Probability of Attending a Top-Quartile School Conditional on College Attendance/Graduation

| | Family Income | | |
|---|---|---|---|
| | Quartile 1 | Quartile 4 | Change |
| | A. Initial Attendance | | |
| NLSY79 attendees: | | | |
| High ability | .074** | .511*** | .437*** |
| | (.030) | (.057) | (.056) |
| Average | .036** | .289*** | .253*** |
| | (.016) | (.043) | (.041) |
| NLSY97 attendees: | | | |
| High ability | .282*** | .468*** | .186*** |
| | (.056) | (.049) | (.073) |
| Average | .136*** | .253*** | .117*** |
| | (.027) | (.026) | (.038) |
| | B. Graduation | | |
| NLSY79 graduates: | | | |
| High ability | .182*** | .436*** | .254*** |
| | (.064) | (.053) | (.081) |
| Average | .133*** | .298*** | .164*** |
| | (.048) | (.042) | (.058) |
| NLSY97 graduates: | | | |
| High ability | .355*** | .479*** | .124* |
| | (.069) | (.046) | (.075) |
| Average | .217*** | .287*** | .069 |
| | (.046) | (.031) | (.052) |

Note.—Respondents are clustered at the primary sampling unit, and we utilize robust standard errors. Predicted probabilities in panels A and B are generated using the estimated coefficients from tables 5 and 6, respectively. The first two columns of probabilities indicate the probability of attending a top-quartile, 4-year college conditional on family income being in the first and fourth quartiles, respectively. The high-ability individual is constructed by taking the average of all the controls other than ability and family income for individuals in the estimation sample who are in the fourth quartile of math and verbal ability. The average individual is constructed by taking the average of all the controls other than family income for all individuals in the estimation sample.
* $p < .1$.
** $p < .05$.
*** $p < .01$.

high. The general pattern of results is similar to that for the initial attendance margin; however, the marginal effects of moving from the first to the fourth quartile of family income on the probability of graduating from a top-quartile college are smaller when compared to the marginal effects along the attendance margin. In addition, while the impact of family income on the quality of the degree-granting college has declined over time, the changes are no longer statistically significant. We view this largely as a result of the small samples available since the change over time for the average high-ability graduate is still considerable.

Table 7 makes clear that there are differences in the effect of family

income on school quality across the initial attendance and graduation margins. These differences are the consequence of heterogeneity in college dropout behavior. In results not reported, we find that dropping out is most prevalent among low-income students attending low-quality colleges. As a result, the dispersion in family income across the quartiles of college quality is smaller in the sample of graduating students as compared to the sample of college attendees.

### *B. Effect of Family Income on Quantity and Quality of Higher Education*

The fact that the set of students who graduate differs from the set of students who initially attend draws attention to the issue of selection more broadly. Looking only at students who choose to attend or graduate from college might hide important patterns that are linked to the extensive margin. In particular, low-income college attendees might be selected quite differently than high-income college attendees. In addition, the selection process for each income group may have changed differentially across the NLSY79 and NLSY97.

To avoid estimating the impact of family income on school quality using differentially selected samples, we estimate broader schooling models in which we assume that all respondents face a choice set that includes high school, 2-year college, first-quartile 4-year college, and so on.[23] Students who never attend any type of college are assumed to choose high school for both initial attendance and graduation. We estimate a multinomial logit model using the same controls as in the initial attendance (table 5) and graduation models (table 6).

Table 8 displays some key predicted probabilities from the estimated models.[24] The results for the initial attendance choice are in panel A and the results for 4-year college graduation are in panel B. Columns 1–3 examine the probability of attending or graduating from any 4-year college as family income changes from low to high, and columns 4–6 examine the probability of attending or graduating from a top-quartile college as family income changes from low to high.

Moving from the first to the fourth quartile of family income increases the probability of attending any 4-year college, and this increase is always economically and statistically different from zero. Similarly to the estimates provided by Belley and Lochner (2007), the increase in the probability of attendance for the average individual is larger in the 1997 sample (0.125) than in the 1979 sample (0.070), though the change is

[23] Note that high school dropouts are excluded from the analysis. For individuals who ultimately obtain graduate degrees, we consider choices regarding only their undergraduate schooling. We have also estimated additional models in which we allow individuals to choose to be college dropouts. The marginal effects reported in table 8 and the monetized effect of family income are not sensitive to these extensions. Results are available on request.

[24] We do not display the coefficients from the full attendance and graduation models though they are available on request.

TABLE 8
Higher Education Choice Probabilities Conditional on Graduating from High School

| | Any 4-Year College | | | Quartile 4 4-Year College | | |
|---|---|---|---|---|---|---|
| | Family Income Quartile 1 (1) | Family Income Quartile 4 (2) | Change (3) | Family Income Quartile 1 (4) | Family Income Quartile 4 (5) | Change (6) |
| | A. Initial Attendance | | | | | |
| NLSY79 respondents: | | | | | | |
| High ability | .620*** | .819*** | .199*** | .061*** | .401*** | .339*** |
| | (.040) | (.039) | (.055) | (.019) | (.045) | (.043) |
| Average | .225*** | .295*** | .070** | .008*** | .053*** | .045*** |
| | (.022) | (.027) | (.033) | (.003) | (.012) | (.011) |
| NLSY97 respondents: | | | | | | |
| High ability | .735*** | .846*** | .111*** | .217*** | .392*** | .175*** |
| | (.033) | (.029) | (.043) | (.045) | (.044) | (.062) |
| Average | .308*** | .434*** | .125*** | .026*** | .065*** | .039*** |
| | (.025) | (.028) | (.036) | (.006) | (.012) | (.012) |
| | B. Graduation | | | | | |
| NLSY79 respondents: | | | | | | |
| High ability | .494*** | .699*** | .205*** | .093*** | .299*** | .206*** |
| | (.052) | (.046) | (.073) | (.030) | (.037) | (.047) |
| Average | .087*** | .132*** | .045** | .007** | .022*** | .014*** |
| | (.013) | (.018) | (.022) | (.003) | (.006) | (.005) |
| NLSY97 respondents: | | | | | | |
| High ability | .594*** | .683*** | .089 | .199*** | .247*** | .122** |
| | (.041) | (.040) | (.060) | (.044) | (.036) | (.055) |
| Average | .164*** | .236*** | .072*** | .019*** | .036*** | .017** |
| | (.018) | (.020) | (.027) | (.005) | (.008) | (.007) |

Note.—Respondents are clustered at the primary sampling unit, and we utilize robust standard errors. Predicted probabilities stem from a multinomial logit school choice model described in the text. The choice set includes high school, 2-year college, and all four quality quartiles of 4-year college. Columns 1–3 consider the probability of attendance at any 4-year college, and cols. 4–6 consider the probability of attendance at a top-quartile 4-year college. The high-ability respondent is constructed by taking the average of all the controls other than ability and family income for individuals in the estimation sample who are in the fourth quartile of math and verbal ability. The average respondent is constructed by taking the average of all the controls other than family income for all individuals in the estimation sample. In addition to family income and ability, we control for family background and geographical variation in access to schools, as in tables 5 and 6.
* $p < .1$.
** $p < .05$.
*** $p < .01$.

not statistically significant. The pattern is reversed for the average high-ability individual, with the impact of family income on the probability of 4-year college attendance much larger in 1979 (0.199) than in 1997 (0.111), though again the difference is not statistically significant.

Columns 4–6 in panel A show the increase in the probability of attending a fourth-quartile quality college as family income changes. Comparing these results to the results pertaining to attendance at any 4-year college is informative about where the gradient between college attendance and family income originates. For the average high-ability student, we see that the entire family income, 4-year college attendance gradient

stems from individuals in the top-quartile schools. For example, in 1979 (1997), a move from the first to the fourth quartile of family income increases the probability of attendance at a top-quartile college by 0.339 (0.175) while increasing the probability of attendance at any 4-year college by only 0.199 (0.111). In fact, the gradient must be negative for all other quartiles to compensate for the large changes in attendance at top-quartile schools. The impact of family income for the average respondent is instead present not only at the top-quartile schools but also for the other quartiles (0.045 being smaller than 0.070 and 0.039 smaller than 0.125). Over time, we do not see a change in the impact of family income for the average student, indicating that the increase in the overall attendance gradient over time stems primarily from increased attendance at lower-quality schools. In contrast, for high-ability students the decrease over time in the effect of family income on the probability of attendance at a top-quartile school is statistically different from zero ($t$-statistic = 2.17) and steeper than the decrease in the importance of family income on attendance overall. Again, the attendance patterns evident in panel A are highly consistent with an increase over time in both average tuition and merit-based aid at top-quality schools.

The pattern of results for graduation in panel B of table 8 is quite similar to that from the initial attendance margin in panel A. In all but one case, a move from the first to the fourth quartile of family income has a positive and statistically significant effect on the probability of graduating from a 4-year college or top-quartile 4-year college for both respondent types in either cohort. The magnitude of the marginal effects of family income on the probability of graduating from a 4-year college is quite similar to that of the attendance marginal effects. However, the effect of family income on the probability of graduating from a top-quartile 4-year college is considerably smaller. Again, the root cause of this change is the fact that most 4-year college dropouts are from low-income and low-quality colleges. Comparing the marginal effects across the 4-year college graduation margin and fourth-quartile college graduation margin again indicates that for high-ability individuals in both cohorts, essentially all of the change in college attendance is occurring at the top-quartile schools. For the average respondent, the impact of family income occurs across the quality distribution. While the impact of family income on 4-year college graduation and graduation from a top-quartile college has decreased in magnitude over time for the average high-ability respondent, the changes are not statistically significant.

### *C. Monetizing the Effect of Family Income on Schooling Outcomes*

While table 8 provides evidence of an important effect of family income on both the quantity and quality of education, it is difficult to compare

the two effects. However, if we convert the impact of family income on educational choices into monetary terms, we can more easily contrast the impact of family income on the two higher education margins. There are well-documented wage returns to obtaining a college degree, and as a result, increasing the income of a student's family will increase the student's future wages through his or her educational attainment. Additionally, graduating from a better school will likely lead to greater wage returns. By allowing the return to education to depend on the quality of the institution, we can therefore quantify the importance of the quality margin with respect to the quantity margin.

In order to create the metric for our comparison, we need to estimate the impact of education on wages. Estimating returns to education is a massive task and has generated an incredibly vast literature (see, e.g., Card 2001). A smaller literature has also tried to estimate the returns to college quality, consistently finding a large impact of quality on income.[25] In this analysis we abstract from all the selection issues and simply provide some ordinary least squares results to be used in our back-of-the-envelope calculations. Although these estimates are likely to be contaminated by selection biases, it is interesting to note that they are in range with most of the instrumental variable estimates presented in the literature.

After estimating the "returns" to schooling, we link them to the probabilities that we have estimated in the previous subsection. Combining these two parts, we are able to provide a simple calculation that quantifies the increase in expected wages for a student of a certain ability group that moves from the first to the fourth quartile of the family income distribution.

In table 9 we report results for the Mincer regressions controlling for education, potential experience, total hours worked, race, gender, state-level unemployment rates, and observed ability measures.[26] The regressions for the 1997 cohort utilize a much younger sample and fewer observations than the 1979 cohort. To address this issue, we estimate the coefficients for the 1979 cohort both on the whole sample and on wage data prior to 1989. Interestingly, the estimated schooling returns hardly change.

Not surprisingly, we find large and significant returns to obtaining a college degree for all three samples. There appears to have been a slight increase over time in the return to obtaining a college degree, though

[25] See Brewer, Eide, and Ehrenberg (1999), Black and Smith (2004, 2006), and Hoekstra (2009). One prominent exception is Dale and Krueger (2002), though the authors' identification approach has received much criticism. See, e.g., Hoxby (2009).

[26] We assume that family income has no direct effect on an individual's earnings. However, if there is a direct effect, the returns to college quality will be biased since, as we have already seen, college quality is correlated with family income. In results not reported, we find that including family income directly into the wage equation has no appreciable impact on the returns to college quality.

TABLE 9
Wage Returns to Schooling

| | 1979 All | | 1979–89 | | 1997 | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| 2-year degree | .207*** | .209*** | .240*** | .241*** | .228*** | .229*** |
| | (.026) | (.026) | (.028) | (.028) | (.034) | (.035) |
| BA, quality Q1 | | .375*** | | .390*** | | .360*** |
| | | (.031) | | (.032) | | (.045) |
| BA, quality Q2 | | .442*** | | .408*** | | .422*** |
| | | (.036) | | (.038) | | (.035) |
| BA, quality Q3 | | .486*** | | .460*** | | .486*** |
| | | (.037) | | (.044) | | (.030) |
| BA, quality Q4 | | .546*** | | .581*** | | .586*** |
| | | (.043) | | (.048) | | (.035) |
| BA | .449*** | | .452*** | | .465*** | |
| | (.020) | | (.026) | | (.025) | |
| Math ability Q2 | .083*** | .082*** | .101*** | .101*** | .086*** | .087*** |
| | (.025) | (.025) | (.022) | (.022) | (.024) | (.024) |
| Math ability Q3 | .165*** | .165*** | .180*** | .180*** | .118*** | .122*** |
| | (.026) | (.026) | (.028) | (.028) | (.027) | (.027) |
| Math ability Q4 | .201*** | .192*** | .191*** | .185*** | .174*** | .163*** |
| | (.034) | (.033) | (.033) | (.033) | (.030) | (.030) |
| Verbal ability Q2 | .103*** | .105*** | .070*** | .071*** | .029 | .031 |
| | (.024) | (.024) | (.021) | (.021) | (.022) | (.022) |
| Verbal ability Q3 | .068** | .071** | .050** | .054** | .030 | .030 |
| | (.028) | (.028) | (.025) | (.025) | (.025) | (.025) |
| Verbal ability Q4 | .096*** | .094*** | .064* | .061* | −.003 | −.009 |
| | (.035) | (.035) | (.034) | (.034) | (.029) | (.029) |
| Additional controls | Y | Y | Y | Y | Y | Y |
| Observations | 49,771 | 49,771 | 21,810 | 21,810 | 17,092 | 17,092 |
| $R^2$ | .302 | .304 | .232 | .235 | .161 | .164 |

Note.—Respondents are clustered at the primary sampling unit, and we utilize robust standard errors. The dependent variable is the hourly wage constructed using previous years' total income and hours worked. Observations are weighted using a combination of the sampling weights provided by the NLSY and the number of hours reported. Each regression includes as additional control variables gender, race, potential experience, potential experience squared, total accumulated hours, and the contemporaneous state unemployment rate.
* $p < .1$.
** $p < .05$.
*** $p < .01$.

the change is not significant. We also find significant wage returns to college quality that are relatively stable both over time and over the life cycle. For example, in the NLSY97, respondents who graduate from a first-quartile 4-year college earn 0.36 log points more than high school graduates, and graduates from a fourth-quartile 4-year college earn 0.59 log points more than high school graduates. These coefficients are statistically significantly different from each other at a 1 percent level.[27]

[27] While the estimated return to quality is large, it is entirely in line with recent estimates that use plausibly exogenous variation in college quality. For example, Hoekstra (2009) uses a regression discontinuity approach to estimate the return to attending a flagship state university relative to a student's next-best option. The estimated 20 percent return

TABLE 10
Percentage Change in Predicted Wage Moving from the First to the Fourth Quartile of Family Income

| | Ignoring Quality (1) | Accounting for Quality (2) | Increase (%) (3) |
|---|---|---|---|
| NLSY79 respondents: | | | |
| High ability | .086 | .112 | 31.0 |
| Average | .021 | .022 | 4.7 |
| NLSY97 respondents: | | | |
| High ability | .038 | .058 | 53.5 |
| Average | .034 | .035 | 2.6 |

Note.—Numbers reflect the percentage change in predicted wage associated with a move from the first to the fourth quartile of family income. We combine predicted probabilities from school choice estimates with estimated returns to schooling to generate the above estimates.

Finally, we combine the estimated returns to schooling with the probabilities previously estimated to find the predicted log wage increase associated with a move from the first to the fourth quartile of family income. The results of this exercise are shown in table 10. Column 1 illustrates the predicted change in log wages that results strictly from changes in the quantity of schooling when family income changes from the first to the fourth quartile. For the average high-ability respondent, wages are predicted to increase by 8.6 percent and 3.8 percent for the 1979 and 1997 cohorts, respectively. The changes for the average respondent are smaller, at 2.1 percent and 3.4 percent, respectively. When we account for college quality, wages are predicted to increase by 11.2 percent and 5.8 percent when we increase family income from the first to the fourth quartile for the average high-ability respondent in the 1979 and 1997 cohorts. These results suggest that for high-ability individuals, the connection between family income and future wages through schooling would be understated by 31 percent and 54 percent if we were to ignore the quality aspect of higher education. Not surprisingly, for the average respondent, the effect of family income on future wages through the quality of schooling is quite small. The average respondent is unlikely to attend a top-tier school regardless of family income. The numbers for high-ability students are large in the 1979 cohort but shrink considerably in the 1997 sample, consistent with the results of tables 7 and 8.

## IV. Conclusion

In this paper we examine how family income affects both the quality and quantity of higher education using the NLSY79 and NLSY97. We

is very close to our estimate, despite the fact that the gap in quality between a first- and fourth-quartile school in our sample is probably significantly larger than the quality gap experienced by most individuals in Hoekstra's sample.

find that family income has a large effect on college quality for high-ability students and a large effect on college attendance and graduation for average-ability students. By mapping the effects of quality and quantity on wages, we are able to show that the impact of family income on the quality of higher education constitutes a significant portion of the overall link between family income and future wages through schooling choices. Our cross-sectional estimates of the effect of family income on college quality, attendance, and subsequent wages should be interpreted as descriptive in nature, as we cannot rule out the possibility that family income is correlated with other unobserved attributes (e.g., noncognitive ability) that affect schooling choices.

As previously noted, the changes we observe over time in the importance of family income for the quality and quantity of schooling are largely consistent with recent trends in the field of higher education. Overall, college tuition has been rising rapidly, and consistent with this trend, we find that the effect of family income on college attendance has increased slightly across the NLSY79 and NLSY97. However, financial aid has become significantly more generous at the nations's top universities. For example, families with income below $60,000 pay no tuition to attend Harvard University. In line with the trends in financial aid at top colleges, we find that for high-ability students the effect of family income on the probability of attending a top-quartile school decreased significantly across the two waves of the NLSY.

The fact that family income is a strong predictor of both college quality and college attendance and that the changes in the importance of family income over time coincide with changes in prices is consistent with the existence of credit constraints in the higher-education market. Of course, these patterns could also be explained by a consumption theory of schooling or simply by omitted variable bias related to unobserved student ability.[28] While we are unable to say precisely whether credit constraints are responsible for the patterns we observe in the data, we can say rather confidently that statistical tests for the presence of borrowing constraints that fail to account for quality are likely flawed. As an example, Cameron and Taber (2004) compare several instrumental variable estimates of the returns to schooling to test for the presence of borrowing constraints. The basic idea of the test is that the estimated returns to schooling using an instrument related to financial constraints should be larger in magnitude with respect to the other instrumental variable estimates of the returns to schooling if individuals are credit constrained. Because they do not find that such returns are larger, they conclude that borrowing constraints are not important. However, as pointed out by Carneiro and Heckman (2002), this argument

[28] While the cross-sectional estimates of the impact of family income are likely biased by unobserved student attributes, the estimated change in the effect of family income will be consistent as long as the relationship between family income and the unobserved student attributes has remained stable over time.

is faulty if there exists a quality margin to the schooling choice. They demonstrate that when individuals adjust on the quality margin as well as on the quantity margin, instrumental variables that induce constrained students to attend lower-quality schools can lower the estimated Mincer return to schooling. Our results indicate that indeed the concern of Carneiro and Heckman is valid.

**References**

Belley, Philippe, and Lance Lochner. 2007. "The Changing Role of Family Income and Ability in Determining Educational Achievement." *J. Human Capital* 1 (1): 37–89.

Black, Dan A., and Jeffrey A. Smith. 2004. "How Robust Is the Evidence on the Effects of College Quality? Evidence from Matching." *J. Econometrics* 121:99–124.

———. 2006. "Estimating the Returns to College Quality with Multiple Proxies for Quality." *J. Labor Econ.* 24 (3): 701–28.

Brewer, Dominic J., Eric R. Eide, and Ronald G. Ehrenberg. 1999. "Does It Pay to Attend an Elite Private College? Cross-Cohort Evidence on the Effects of College Type on Earnings." *J. Human Resources* 34 (1): 104–23.

Cameron, Stephen V., and Christopher Taber. 2004. "Estimation of Educational Borrowing Constraints Using Returns to Schooling." *J.P.E.* 112 (1): 132–82.

Card, David. 2001. "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems." *Econometrica* 69 (5): 1127–60.

Carneiro, Pedro, and James J. Heckman. 2002. "The Evidence on Credit Constraints in Post-secondary Schooling." *Econ. J.* 112 (482): 705–34.

Dale, Stacy B., and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Q.J.E.* 117 (4): 1491–1527.

Ellwood, David T., and Thomas J. Kane. 2000. "Who Is Getting a College Education? Family Background and the Growing Gaps in Enrollment." In *Securing the Future: Investing in Children from Birth to College*, edited by Sheldon Danziger and Jane Waldfogel. New York: Russell Sage Found.

Fuller, Winship C., Charles F. Manski, and David A. Wise. 1982. "New Evidence on the Economic Determinants of Postsecondary Schooling Choices." *J. Human Resources* 17 (4): 477–98.

Hoekstra, Mark. 2009. "The Effect of Attending the Flagship State University on Earnings: A Discontinuity-Based Approach." *Rev. Econ. and Statis.* 91 (4): 717–24.

Hoxby, Caroline M. 2009. "The Changing Selectivity of American Colleges." *J. Econ. Perspectives* 23 (4): 95–118.

Kane, Thomas J. 2006. "Public Intervention in Post-secondary Education." In *Handbook of the Economics of Education*, vol. 2, edited by Eric A. Hanushek and Finis Welch. New York: Elsevier.

Long, Bridget Terry. 2004. "How Have College Decisions Changed over Time? An Application of the Conditional Logistic Choice Model." *J. Econometrics* 121: 271–96.

Lovenheim, Michael F., and C. L. Reynolds. 2010. "The Effect of Housing Wealth on College Choice: Evidence from the Housing Boom." Working paper, Cornell Univ.

———. 2011. "Changes in Postsecondary Choices by Ability and Income: Evi-

dence from the National Longitudinal Surveys of Youth." *J. Human Capital* 5 (1): 70–109.
Wald, Abraham. 1940. "Fitting Straight Lines If Both Variables Are Subject to Error." *Ann. Math. Statis.* 11:284–300.