2018

Degrees of Poverty: The Relationship between Family Income Background and the Returns to Education

Timothy J. Bartik
W.E. Upjohn Institute, bartik@upjohn.org

Brad J. Hershbein
W.E. Upjohn Institute, hershbein@upjohn.org

Upjohn Institute working paper ; 18-284

W.E. UPJOHN INSTITUTE
FOR EMPLOYMENT RESEARCH
Researching the causes and consequences of unemployment

Degrees of Poverty: The Relationship between Family Income Background and the Returns to Education

Upjohn Institute Working Paper 18-284

Timothy J. Bartik
email: bartik@upjohn.org
W.E. Upjohn Institute for Employment Research

Brad Hershbein
email: Hershbein@upjohn.org
W.E. Upjohn Institute for Employment Research

March 2018

ABSTRACT

Drawing on the Panel Study of Income Dynamics, we document a startling empirical pattern: the career earnings premium from a four-year college degree (relative to a high school diploma) for persons from low-income backgrounds is considerably less than it is for those from higher-income backgrounds. For individuals whose family income in high school was above 1.85 times the poverty level, we estimate that career earnings for bachelor's graduates are 136 percent higher than earnings for those whose education stopped at high school. However, for individuals whose family income during high school was below 1.85 times the poverty level, the career earnings of bachelor's graduates are only 71 percent higher than those of high school graduates. This lower premium amounts to $300,000 less in career earnings in present discounted value. We establish the prevalence and robustness of these differential returns to education across race and gender, finding that they are driven by whites and men and by differential access to the right tail of the earnings distribution.

JEL Codes: I24, I26, J24, J31

Key Words: inequality, return to education, career earnings profile, PSID, low-income

Acknowledgments:

We are grateful to Sarah Hamersma, Michael Hout, Melissa Kearney, Yulya Truskinovsky, and seminar participants at the 2017 PAA meetings, 2017 APPAM meetings, 2017 SEA meetings, Notre Dame, and the University of Michigan. We thank Wei-Jang Huang and Nathan Sotherland for excellent research assistance with the PSID. All errors are our own.

Growing earnings inequality in the United States has been the subject of much research, across a number of dimensions. This growth is thought to have many causes, ranging from international trade to technological change to institutional change (e.g., the declining real value of minimum wage, declines in unionization, and changes in corporate governance), and part of these structural changes have been linked with rising returns to education. Indeed, a large research literature has sought to describe or explain the growing gap between the more-educated and less-educated, as well as an increase in earnings dispersion within the more-educated group (Katz and Murphy 1992; Autor, Katz, Kearney 2006, 2008; Lemieux 2006). The overall growth in earnings inequality, and its relationship with the high (and perhaps still growing) returns to education, have led to heightened concerns about economic opportunity and mobility.

The relationship between earnings inequality and returns to education has also led to speculation among many policymakers and economists that increasing educational attainment, particularly among the poor, could help equalize economic opportunity, if not outcomes. Although an increase in the share of the population that is more highly educated would presumably decrease the returns to education through supply-side factors, possibly stopping or reversing its growth over the past few decades, it is certainly conceivable that providing more education to the poor might improve their relative earnings outcomes if not appreciably diminish overall inequality (Hershbein, Kearney, and Summers 2015).

Yet, another strain of research recognizes that returns to education are heterogeneous for different groups (Card 1999; Carneiro, Heckman, and Vytlacil 2001, 2011; Brand and Xie 2010), including for individuals whose parents have different levels of education (Altonji and Dunn 1996) or who vary in cognitive ability (Ashenfelter and Rouse 1998). This research literature has emphasized the role of selection in driving unobserved heterogeneity, noting that college attendance and completion vary by family background (Belley and Lochner 2007; Bailey and Dynarski 2011). Related to this literature are studies that investigate how child and adolescent experiences—especially poverty—shape career trajectories (Alexander, Entwhistle, and Olsen 2014). In these studies, a prominent mechanism for the role

of family socioeconomic background in affecting career earnings is through college attendance and attainment, with poverty found to be especially harmful for these indicators.

However, family income background may have career earnings effects well beyond its effect through college attendance and attainment. Lower-income family background may be associated with various events and experiences (e.g., family structure and environment, neighborhood influences and peer influences, school quality) that may lead to lower hard or soft skills or fewer connections, which might in turn lead to less ability to obtain better jobs. These effects may operate independently from educational attainment or interact with them heterogeneously. There has been very little work, to our knowledge, on how childhood poverty (or low family income more generally) affects the return to education, *conditional on achieving it*, over the whole career.[1] Family income background and educational attainment could interact because the skills and knowledge obtained through childhood experiences may have complementarity or substitutability with skills causally imparted by education attainment. For example, if a higher-family-income background provides better knowledge of social connections and better soft skills, this may complement better hard skills imparted by the educational system, so that a higher-income background and higher educational attainment together yield greater earnings returns than either would yield separately (Deming 2015). Alternatively, perhaps the educational system teaches students from a low-income background certain soft skills that students from a higher-income background pick up through other channels, so that school can to some extent substitute for family income background. Relations of complementarity between family background and educational attainment would lead to educational returns increasing with higher-income family backgrounds, while relations of substitutability might lead to educational returns declining with higher-income family backgrounds (Brand and Xie 2010).

[1] Torche (2011) investigates a related question of how intergenerational mobility varies by realized education of the children, finding a U-shaped pattern. However, her measures of mobility do not cover the whole career, and her focus is on absolute mobility rather than heterogeneity in the education earnings premium, per se.

Furthermore, family income background could correlate with empirically measured returns to education not due to true causal effects of education, but rather to how family income background affects who is selected to be educated. For example, low-income students who achieve higher educational attainment could be positively selected, with only the most motivated and talented succeeding in overcoming adversity. On the other hand, as policy pushes more low-income students to attain higher educational credentials, perhaps such marginal students will be negatively selected, with fewer skills.

In discussing education returns, and how they vary with family income background, we must be careful in distinguishing between different types of inequality. In the most extreme case, one might hope that although lower family-income background may handicap earnings prospects for individuals with less education, higher educational attainment has causal effects that would entirely eliminate these disadvantages. In this case, both the dollar and percentage returns to education would be much higher for persons from lower-income backgrounds, and in fact high enough to completely eliminate the handicap of family background, so that earnings of persons from different family-income backgrounds would be equalized at a sufficiently high enough level of education.

Less extremely, one might hope that the dollar earnings premium from education might be similar for persons from different income backgrounds, which, given lower baseline earnings at low education levels for persons from lower-income backgrounds, would imply higher *percentage* returns for persons from lower income backgrounds. Although such a situation would lead to identical benefit-cost analysis (assuming cost did not vary by income background) for individuals from different income groups, it would still have a progressive impact on most measures of the income distribution by having greater proportional effects on individuals from lower-income families.

More modestly still, advocates of reducing income inequality would hope that, at a minimum, the *percentage* returns to education for individuals from low-income backgrounds are at least not *less* than those for individuals from upper-income backgrounds. If the percentage returns are similar, then seeking to generally boost education attainment for all at least would not make the earnings distribution more

unequal. If the percentage returns to education are lower for persons from a low-income background, additional education if widespread—as has occurred in recent decades—can actually increase earnings inequality.

In this paper, we extensively document career earnings profiles by education levels (particularly bachelor's graduates and high school graduates), stratified by family income status during adolescence, using the Panel Study of Income Dynamics (PSID). (Future versions of this paper will also include profiles estimated from the National Longitudinal Study of Youth, 1979.) Unsurprisingly, and consistent with earlier research, we find that career earnings of bachelor's graduates who grew up in low-income households are substantially lower than similarly educated individuals who grew up in higher-income households.

However, and very surprisingly, we also find that the career *percentage* earnings premium from earning a bachelor's degree, relative to only a high school diploma, is much lower for individuals who grew up in low-income families. For individuals who grew up in families below 185 percent of the federal poverty line (the threshold for participation in the assisted school lunch program), the career earnings premium from a bachelor's degree is 71 percent, but for those who grew up in families above that income threshold, it is 136 percent. Thus, we find that education not only has much lower absolute returns for persons from low-income backgrounds, it also has much lower proportional returns.

In this paper, we present these basic results and explore their robustness to different assumptions and their variation across different groups. This is an exploratory and descriptive analysis that seeks to document stylized facts that can serve as a springboard for later causal analysis.

Summarizing our findings, the pattern described above is quite robust to our choice of sample and weighting within the PSID, what ages of earnings are included, and whether observations with zero earnings are included. However, our findings are sensitive to other variations. The higher education returns for high-family-income background individuals diminish when we exclude individuals who ever get graduate degrees, indicating a great deal of the higher college returns for the high-income-background

group is associated with postcollege credentials. Our findings are also sensitive to truncating the earnings distribution analyzed. Notably, the higher observed college earnings premium for individuals from higher-income backgrounds is all but eliminated when we exclude the top 1 percent of earnings, or if we look at the conditional median of earnings rather than the conditional mean.

However, differentials in education returns with family income background are particularly marked when we compare the low education returns for individuals whose family income background was in near poverty (100–200 percent of the poverty line) to the high returns for individuals whose family income background was above 400 percent of the poverty line. Percentage returns to higher education are actually quite high for (the very few) individuals from families in poverty who achieve it.

Finally, the differential college earnings premium by family-income background is more evident among men and whites. Among women and blacks, family income background does not seem to affect the percentage return to education.

Overall, we infer that the differential percentage return to education across family income background is largely driven by differential access to the right tail of the earnings distribution. College graduates who come from high-income families (400 percent or more of the poverty line), particularly men and whites and those who get graduate degrees, have a much higher chance of accessing the top of the earnings distribution than otherwise similar individuals who come from poorer families. In addition, among persons with a family income background that is "near-poor" (100–200 percent of the poverty line), the returns to college are not particularly high, especially for men and whites—these groups do almost as well with just a high school diploma.

The next section of the paper presents our data and methodology, with full details in Appendix A. The third section summarizes our initial results and discusses them in the context of earlier literature. The final section concludes.

DATA AND METHODOLOGY

Few existing data sets contain information on both economic circumstances during childhood or adolescence and earnings over the adult career. Administrative data in other countries, particularly Scandinavian ones, have been used to examine intergenerational outcomes dependent on earnings (Hirvonen 2008; Nilsen et al. 2012; Lundborg, Nilsson, and Rooth 2014). More recently, U.S. administrative data from tax records have been used by Raj Chetty and several coauthors to document intergenerational economic mobility and investigate the long-term effects of neighborhood characteristics on adult outcomes (Chetty et al. 2014; Chetty and Hendren 2015). However, U.S. administrative data seldom provide background information on even basic demographics, such as race or educational attainment, let alone detailed characteristics such as hours worked, colleges attended, or neighborhoods of residence.

Because these latter characteristics are integral in identifying and explaining the observed career earnings premium to a bachelor's degree by family income background, we employ rich and detailed longitudinal survey data, the Panel Study of Income Dynamics (PSID). The PSID, one of the longest running longitudinal household surveys, has followed several thousand families and their descendants since 1968. Families were interviewed annually through 1997 and have been interviewed biennially since, with the most recent data wave in 2013. Importantly for our purposes, these data allow linking children and adolescents earlier in their survey, when their parental income can be observed, to the educational and earnings profiles of those children as they progress through adulthood. Several studies investigating intergenerational income mobility have used the PSID (Solon 1992; Shin and Solon 2011), in large part because of the high quality of its earnings data across generations, although none to our knowledge has examined the observed education earnings premium over the career by family income background.

Our approach involves calculating the average family income of individuals when they are between the ages of 13 and 17, and renormalizing this income into a poverty ratio by dividing by the

federal poverty threshold for the observed year and household composition.[2] We use family income averaged over multiple years to deal with the mismeasurement and transitory income issues raised by Solon (1992), thus more accurately capturing economic circumstances during adolescence.[3] Using these family income ratios, we divide individuals into two groups: those with a low-income background, defined to be average family income below 185 percent of the federal poverty threshold; and higher-income backgrounds, those with family income above that threshold. Our choice of the 185 percent threshold (and binary groupings) is driven both by practical considerations—notably, allowing for sufficient sample sizes and simplicity in interpretation—and policy considerations. In particular, school-age children with family incomes less than 185 percent of the federal poverty guideline are eligible for the National School Lunch Program, which provides free or heavily subsidized lunches to students in school. Eligibility for the program is one of the most common indicators for low socioeconomic status available for students, and its prevalence in schools influences the receipt of Title I money to help pay for educational services. In addition, programs to enhance educational achievement or attainment of students in the K–12 system are frequently targeted by eligibility for a free or reduced-price lunch. Studies of how different educational reforms affect achievement or attainment often use this indicator to suggest the distributional implications of these reforms. Therefore, we do not consider this threshold arbitrary in demarcating childhood economic status. Nonetheless, we do explore how educational returns vary for somewhat more detailed groupings of family income background.

After this step, we use information on the timing of high school graduation (and, when pertinent, GED receipt), as well as postsecondary degree completion, to assign each individual the highest level of

[2] See https://www.census.gov/hhes/www/poverty/data/threshld/.

[3] While there are reliability gains to using just a few years of averaged income over a single year of income, Mazumder (2005) has noted that longer periods may be necessary to capture "permanent" income. However, such attenuation bias is a larger issue when trying to calculate an intergenerational earnings elasticity—a cardinal measure—than assigning families into a few ordinal earnings categories as we do here. Additionally, calculating family earnings over a longer span of childhood comes at the cost of sharply reduced sample sizes and less feasibility of capturing earnings over nearly the whole career. Thus, the chosen age range seemed a reasonable compromise.

education achieved by age 25. We are particularly interested in individuals whose highest education is either a high school diploma (excluding GEDs[4]) or at least a bachelor's degree. Although we recognize that some individuals earn a bachelor's degree after age 25 (as well as graduate degrees), we choose this age cutoff to maximize the observed earnings history under specific levels of education, although we again explore robustness to more flexible assignments.[5]

We follow over time the individuals for whom we can calculate a family income status and calculate simple (nonparametric) earnings profiles over the career, by family income and education groups.[6] More specifically, we estimate group-specific profiles using regressions of the following form:

$$y_{iacft} = \sum_a (\beta_a \times \boldsymbol{I}[age = a] \times \boldsymbol{I}[college = 1]) + \mu_{ct} + \varepsilon_{iacft}, \quad (1)$$

where y_{iacft} is annual reported earnings, including zeros (in year 2014 dollars, adjusted using the personal consumption expenditures [PCE] deflator), for individual *i*, at age *a*, in education group *c* and family income group *f*, measured in year *t*. The first term on the right side is a set of age dummies and separate coefficients for each education group (high school or at least a bachelor's degree), the μ_{ct} represent survey year dummies for each education group, and ε is the error term. By estimating Equation (1) separately for each family income group, we allow survey year effects and, of course, age profiles, to vary flexibly across education and family income groups.

We use the estimated coefficients β_a to construct fitted profiles for each education and family income group, applying the estimated survey year effect from 2012, the most recent earnings year, to

[4] Heckman and LaFontaine (2006) show that the causal returns to a GED are considerably less than an actual high school diploma.

[5] Since we also measure career earnings starting at age 25, the educational classification is correct at career start; later ages would miss larger portions of career earnings. Some of the robustness checks attempt to address issues relating to differential later-in-life degree acquisition by family income background.

[6] Imposing a quadratic age profile yields results very similar to those from a nonparametric profile, although we can reject the quadratic specification.

control for business cycle effects and secular (but group-specific) real wage growth. In our primary specifications, we begin our earnings profiles at age 25. Because the PSID data cover earnings from 1967 through 2012, the oldest birth cohort for whom we can measure family income during adolescence and earnings during adulthood are individuals born in 1950; the youngest cohort consists of individuals born in 1987. This cohort structure means that earnings observed at younger ages will consist of many birth cohorts, while those at older ages—up to 62, the oldest age we observe—will come disproportionately from earlier cohorts.[7]

While we relegate remaining details of data construction to Appendix A, it is important to describe a few salient features of the PSID that could affect interpretation. First, in the interest of maximizing our sample, we combine both the nationally representative cross-sectional sample and the low-income oversample and apply sample weights to maintain representativeness, although we also use only the cross-sectional sample (with or without weights) as a sensitivity check. Consequently, while the results should be representative of families that were living in the United States in 1968, they are generally not representative of all families living in the United States today, mostly due to a lack of representation of immigrants and Latinos.[8]

Second, the PSID consistently collects earnings only for household heads and their spouses.[9] Thus the analytic sample is necessarily restricted to individuals not living with their (sample member) parents, although a respondent is automatically counted as a head if he or she is living in a household with

[7] To address the possibility of cohort effects, we attempt to investigate profiles broken down by cohort groups, although data limitations limit us to looking at earnings through age 40. Although estimates are imprecise, we do not find strong differences between earlier (born in the 1950s) and later (born in the 1960s) cohorts. Consistent with earlier research (e.g., Autor, Katz, and Kearney 2008), the observed college premium rose for the later cohorts, perhaps weakly more so for lower-income individuals. Unlike for the full career range in our main analysis, we do not find strong evidence for differences in the college earnings premium by family income background through age 40, for either cohort group or overall, suggesting that earnings later in life are an important aspect of the pattern we find.

[8] We do not use the more recent immigrant or Latino samples of the PSID, as they do not allow for both determination of childhood economic circumstances and a (nearly) complete earnings profile.

[9] Earnings for all individuals are collected only from 1999 through 2007.

no other original sample respondents. This issue is mitigated to some extent by looking at earnings at ages 25 and older, when most children have become heads or spouses, but there is the potential that selection into these statuses varies by education within family income background, which could bias our estimated earnings profiles. Using a variant of the equation described above, we show in Appendix Table A.4 that this is not the case: while selection exists for college graduates (relative to high school graduates), it does not differentially affect low-income college graduates, and so comparing the observed career college earnings premium across income groups should be problematic due to selection concerns. (We confirm the same story holds more broadly for differential attrition from the survey in Table A.3).[10]

We present summary statistics for our analytic sample in Table 1. Statistics are shown for both the overall sample and each of the four education–family-income-background groups. For the entire sample, there are just over 47,000 person-year observations from about 4,400 unique individuals. Approximately 18 percent of the weighted observations are in the high school, low-income group, and 41 percent are in the high school, non-low-income group. The remaining 42 percent of the sample has a bachelor's degree, and unsurprisingly most of these come from the not-low-income group: only 3.7 percent of the whole sample are college graduates who grew up low-income. Indeed, it is the relative rarity of this group that makes the study of their earnings paths so interesting.[11] While average earnings—including zeros—are $46,000, the distribution is very diffuse, with a standard deviation of over $70,000.[12] Slightly over half the sample is female, about one-eighth is black, and the average family income during adolescence was about 3.7 times the poverty level.

Looking across the groups, it is clear that average earnings are higher not just for individuals from the more-educated groups, but for individuals who grew up in more fortunate economic

[10] Fitzgerald, Gottschalk, and Moffitt (1998) find that attrition in the PSID has not had serious impact on the representativeness of the data, especially when appropriate sample weights are used.

[11] For all those who grow up in low-income families, about 8 percent have earned a bachelor's degree by age 25, relative to 31 percent for individuals who grew up in higher-income families.

[12] Earnings are not top-coded in the PSID, although we examine sensitivity to outliers later on. We exclude the relatively few observations with imputed earnings, as described in more detail in Appendix A.

environments, conditional on education. This socioeconomic premium is particularly large for college graduates, at roughly 60 percent, although it is still rather considerable at 30 percent for the high school graduates. It also worth noting that both low-income groups are disproportionately female, at 58–60 percent. This likely reflects the overall greater high school dropout rates for men (Stark and Noel 2015), and may also result from low-income backgrounds having a more pernicious effect on education for men than for women (Autor et al. 2015; Chetty et al. 2016). Because of the association between income and race, the low-income groups are more heavily black than the non-low-income groups (although less so than in the unweighted data). Interestingly, although college graduates from non-low-income backgrounds grew up in families with substantially greater incomes than high school graduates (526 percent of the poverty threshold compared to 348 percent), this does not hold for the low-income group, where average family income is almost the same (131 percent compared to 123 percent).

RESULTS

We first present results for our baseline estimates. We then explore these findings more thoroughly using a variety of different samples and estimation strategies. These findings are meant as explorations to suggest causal hypotheses about the meaning and interpretation of our baseline estimates.

Baseline

We begin by presenting graphical career earnings profiles in Figures 1A and 1B, which respectively show trajectories of mean earnings by education for individuals who grew up in low-income families and for those who did not. Because the PSID switches to biennial surveying in 1997, we report earnings by two-year age bins for consistency.

In both graphs, college graduates earn considerably more than high school graduates throughout the career, but this is not surprising. More interesting is the difference in slopes and levels of the profiles. For high school graduates, the earnings slopes are quite similar across income backgrounds, with roughly

$700 increases every two years of age, although those from higher-income backgrounds earn about $10,000 more at each age up to about age 50.[13] In contrast, both slopes and levels diverge considerably for college graduates. From the mid-20s through the mid-40s, low-income college graduates on average increase their earnings by about $2,300 every two years, while higher-income college graduates have average increases more than twice as large, at roughly $5,200 every two years. Furthermore, while earnings appear to peak in the mid-40s for the former group, they continue to rise until age 50 for the latter group. Together, these factors imply that the average college graduate who grew up in a low-income family earns about as much at the peak of the career as the average college graduate from a higher-income family whose career is just beginning.

While these graphs are useful in illustrating when over the life cycle earnings differences occur, we construct a summary measure of career earnings by taking the present discounted value of the profiles, between ages 25 and 62, assuming a 3 percent discount rate and a base period of age 18. These values are presented in Table 2. For individuals from low-income families who obtain only a high school diploma, present discounted earnings are approximately $475,000, while for those who receive at least a bachelor's degree, this figure is about $810,000, a difference of $335,000 and a 70.6 percent increase. For individuals from higher-income families, high school graduates earn about $661,000 over the career, or about 39 percent more than individuals with the same level of education from poorer families. However, average career earnings for bachelor's graduates from the more well-to-do families reach $1.56 million. Not only is this amount nearly two times what low-income bachelor's graduates earn, it is 136 percent more than what higher-income-background high school graduates earn.

Put differently, the observed career earnings premium to (at least) a bachelor's degree, relative to a high school diploma, is nearly twice as large proportionally for individuals from non-low-income families as it is for individuals from low-income families, 136 percent to 71 percent. This proportional

[13] In proportional terms, the slope for low-income high school graduates is slightly larger, 2.7 percent compared to 2.0 percent, but not statistically significantly so.

difference is statistically significant at the 5 percent level and is quite large in practical terms. If low-income-background college graduates received the same proportional boost to career earnings as their peers from more fortunate backgrounds, their present discounted career earnings would be $1.12 million, or $312,000 (38.5 percent) more than what they are observed to earn. If low-income-background college graduates received the same *dollar* return to college graduation as their peers from higher-income backgrounds, their present discounted career earnings would be $1.38 million, or $565,000 (69.9 percent) more than their observed earnings. At an extreme, if a bachelor's degree completely eliminated any disadvantages associated with a low-income family background, and low-income-background college graduates earned the same $1.56 million as their peers from higher-income backgrounds, this $1.56 million would represent a boost of $752,000 (92.9 percent) from their $810,000 in actual career earnings.

To be clear, we are *not* claiming that coming from a low-income background *causally* reduces the return to a bachelor's degree. Rather, we find the magnitude of the observed *correlational* difference in the premium striking, in part because the selection issues and literature described earlier suggest returns should be (weakly) higher for those with a low propensity to complete college, and in part because such large proportional differences are not observed across race or sex.[14]

Of course, a myriad of factors could explain the relatively low earnings premium to college for individuals from low-income backgrounds, including measurement issues inherent in the data as well as more substantive issues. In the following subsections, we explore a number of features of these estimates in an attempt to understand these estimates more fully.

[14] Using the PSID data and the same approach, the observed proportional increase in present discounted career earnings from a bachelor's degree, relative to a high school diploma, is 137 percent for (all) whites, 162 percent for blacks, 173 percent for men, and 119 percent for women. Similar findings are obtained when one examines the cross-sectional (synthetic cohort) returns to education and how they vary with age using data sets such as the Current Population Survey or the American Community Survey (Bartik, Hershbein, and Lachowska 2016).

Decomposing Baseline Differences in Returns: College Versus High School Earnings

How much of this large difference in relative percentage returns across family income backgrounds is due to differences in college earnings versus differences in high school earnings? To do such decomposition, it is convenient to reexpress the differences in the observed premium to a college education in logarithmic differences, as shown in the top part of Table 3. For individuals who grew up in a higher-income family background (greater than 185 percent of the poverty line), the 136.3 percent career college earnings premium from Table 2 corresponds to a logarithmic value in Table 3 of 86 log points (i.e., $\ln(2.363) \times 100 = 86.0$). For those who grew up in a lower-income family background (less than 185 percent of the poverty line), the 70.6 percent college premium corresponds to 53.4 log points. The log difference is thus 32.6 points in Table 3, which corresponds to the 65.7 percentage point difference in Table 2.

The advantage of this logarithmic formulation is that the difference across income groups in the college earnings premium exactly equals the differences in the income background premium across education groups. That is, Table 3 shows that, among those with a high school diploma only, the higher-income background group actually earns about 33 log points more over the career than the lower-income background group. But this difference across groups is roughly doubled, at 66 log points, for those who earn a bachelor's degree.

As this discussion suggests, the differential return to a bachelor's degree across income groups is driven both by income background differences in the earnings of high school graduates and those in college graduates (as well as selection into each education group). While both are empirically important, the greater dispersion in the earnings of college graduates—which has been growing over time (Autor, Katz, and Kearney 2006)—plays the larger role, and is a topic to which we return.

Further Exploring Differentials by Age

Before considering variants to the baseline specification, we further explore earnings differentials by age, beyond the picture provided by Figure 1. To do so, Table 4 considers the difference in predicted

earnings (logged) across the two income background groups, for each educational attainment category, by two-year age bins. This formulation allows us to highlight how the differential education return varies over the career.

As shown in Table 4, log differences across income groups in earnings are quite similar at younger ages for the two education groups, roughly around 30–40 points. But as individuals age, the log differences between income background rapidly increase for the college-graduate group, through about age 50, while remaining static for the high school graduate group. More specifically, log differences for the former group reach 50 points during an individual's 30s, and then rise to around 100 points as one approaches age 50. In contrast, log differences in earnings for high school graduates generally stay between 30 and 40 points into one's 50s, although the estimates bounce around a bit at older ages.

Thus, individuals from higher-income backgrounds earn greater proportional career returns from completing college due to relative earnings increases that begin in the early 30s, accelerate rapidly until about age 50, and then are maintained until the effective end of our sample, at ages in the early 60s. However, it is worth noting that at the oldest of these ages the standard errors of our estimates increase due to fewer earnings observations.[15]

Sensitivity to the Sample and to Weighting

Our baseline estimates use the combination of the Survey Research Center (cross-sectional or SRC) sample with the Survey of Economic Opportunity (low-income or SEO) sample of the PSID, with the estimates weighted using PSID-provided weights. As explained above, we include the SEO sample because we want more individuals with a low-income background in our estimation sample, precisely the purpose of the SEO sample component.

[15] The drop in predicted earnings for college graduates from low-income backgrounds (relative to both high school graduates from the same background and to college graduates from more affluent backgrounds; see Figure 1) also plays a role, albeit a smaller one due to the high discounting at these ages.

We use weighting because the SEO oversamples lower-income persons, particularly in the South, and weighting is necessary for sample statistics to approximate population statistics. However, while the weights correct for representation by income background, they may not do so by education *within* income background, which could threaten the representativeness of our estimates of differential educational returns (Solon, Haider, and Wooldridge 2015).[16]

To test the sensitivity of our results to such sample and weighting assumptions, Table 5 considers alternative sampling choices using the PSID. We first consider two options: using only the nationally representative SRC sample with sample weights, and using the SRC sample but without weights. It is reassuring that the nationally representative SRC sample results are very similar to the weighted SRC-SEO results. If anything, the differences in the returns to education across income background groups are somewhat greater in the SRC sample, weighted or unweighted. For example, in the SRC weighted sample, the log return to a college degree is 41 points for individuals from a low-income family background, compared to 85 points for those individuals from a higher-income family background, a difference of 44 log points. The baseline SRC-SEO weighted estimates show a difference of 33 log points. Even without weighting, the SRC-only sample shows a difference of 37 log points. However, because the SRC sample has fewer individuals in it with a low-income background, restricting estimation to the SRC sample tends to increase standard errors somewhat, although not enough to change statistical significance appreciably.

On the other hand, when we use the unweighted SRC-SEO sample, the rate of return to education for the low-income background group tends to increase, and the difference between the two groups in the percentage return to a college education (using either absolute percentages or log differences) decreases and is no longer statistically significant. A possible explanation of this phenomenon is provided by later results in this paper, which suggests that returns to college education tend to be higher for individuals with extremely low-income backgrounds. We suspect that the SEO oversample of such individuals tends

[16] Brown (1996) discusses some of the sampling issues of the SEO oversample in the PSID.

to increase the rate of return to education. Weighting corrects for this phenomenon by making the low-income-background group (less than 185 percent of poverty) more representative of the population from this income background and less disproportionately from below the poverty line.

The last row of Table 5 emphasizes that in all cases, even when the percentage returns to education are no longer significantly higher for the higher-income group, the absolute dollar value of the returns to college is still significantly higher for the latter group. For example, even in the SRC-SEO unweighted sample, the career dollar return to a college education is $414,000 greater for individuals from a higher-income background than it is for individuals from a low-income background. Both the proportional and absolute return to college for the two income background groups may be of interest for policy purposes.

Other Sample Restrictions

Table 6 considers other restrictions in the estimation sample. First, rather than starting earnings at age 25, we estimate conditional earnings starting at age 20. (The education groups are still classified by education attainment as of age 25). This change adds career earnings for all education and income background groups, but proportionally more for the high school graduate group, who spend more of this age period working. While this reduces the career return to a college education, as expected, it does not drop much differentially by family income background. Relative to the baseline case of a differential across income groups of 32.6 log points, including earnings from age 20 reduces the differential only to 30.6 log points.[17]

In the next column of Table 6, we exclude all observations with zero earnings, essentially comparing earnings profiles of workers instead of all individuals. If education differentially affects the likelihood of employment based on family income background, restricting the sample to observations

[17] Note that the tendency of higher-income individuals to complete their degrees faster (and thus have additional earnings between age 20 and 25 as college graduates) would tend to *increase* the log differential. That it falls slightly suggests that relative earnings of high school graduates across income groups play a role at these ages.

with positive earnings should shed light on the magnitude of this mechanism in driving the differential. Perhaps unsurprisingly, this restriction slightly reduces the observed college premium for the low-income-background group, as earnings increase proportionately more for high school graduates. For individuals from a higher-income background, the proportional college premium also falls, but by somewhat less. Therefore, this restriction actually increases the relative college premium for the higher-income group, from about 33 log points at baseline to 42 points in the positive-earnings-only sample.[18]

Finally, we consider dropping from the sample anyone who ever obtains a postcollege degree. This restriction does make a major change to the earnings differential across income groups. In particular, it significantly reduces expected lifetime earnings for the higher-income background group with a college degree, from $1.56 million in the baseline sample to $1.23 million when those who obtain graduate degrees are excluded. As a result, the estimated percentage rate of return to college falls substantially for the higher-income background group, from 86 log points at baseline to 63 log points. This 63-point return to college is not substantively or significantly different from the estimated 62 log point return to college for the low-income background group, whose college premium actually rises slightly.[19] The dollar return to a college education for the higher-income background group still exceeds the dollar return for the low-income background group, but by much less—the differential is only $205,000, relative to $566,000 at baseline.

The results from restricting the sample to those who never earn a graduate degree certainly suggests that the college return differential by income background is a story more about advanced degrees than bachelor's degrees, per se. However, in the human capital investment framework, the additional return that a graduate degree may bring is part of the option value to a bachelor's degree. Under such a

[18] This is consistent with larger differences in the likelihood of employment with greater education among the low-income group, but the magnitude is relatively modest.

[19] At age 25, 17.4 percent of bachelor's graduates from the higher-income background group have a graduate degree, while a similar 16.3 percent of graduates from the low-income group do. However, these shares diverge with age. By age 30, they are 25.6 and 18.6 percent, respectively, and at the maximum observed ages they are 36.5 and 27.8 percent.

conceptual framework, if individuals who grew up in a more economically advantaged family are more likely to earn a graduate degree, conditional on earning a bachelor's degree, then the advanced degree is a possible mechanism for their proportionally greater return to earning (at least) a bachelor's degree, and one that has important implications for policies that promote undergraduate (but not necessarily postgraduate) access and success. Moreover, compositional differences in who obtains a graduate degree (and in what field) may also play a role, something to which we return below.

More Flexible Income Background Groups

Although there is some reason to particularly focus on the free and reduced-price lunch cutoff, as mentioned above, we want to explore more flexible specifications for how family income background might affect the return to education. In Table 7, we consider as an alternative dividing individuals into four groups by family income background: greater than 400 percent of the poverty line, 200–400 percent of the poverty line, 100–200 percent of the poverty line, and below the poverty line. In examining the return to a college degree, we compare each of the three lower-income groups to the highest-income group.

We find that the variation in the percentage return to a college degree follows a somewhat U-shaped pattern with family income background. The biggest contrast in college premium is between the highest-income-background group and the one with family income between 1 and 2 times the poverty line. The highest-income group has a log return to a college degree of about 83 points, which is 44 points higher than the 39-point return for the "near-poor" income group. The second-highest-income group (200–400 percent of poverty) has a premium that is in between, at around 63 log points. But for individuals who grow up below the poverty line, the observed college premium is remarkably high: roughly 103 log points higher (but not statistically significantly so) than the 83 points of the highest-income-background group. Additionally, the absolute dollar return for college graduates who grew up in

poverty is larger for all but the highest-income group, a pattern largely driven by the very low earnings of individuals with only a high school diploma from the poorest group.[20]

A comparison of the poverty-income group with the near-poor group is instructive. In log terms, college graduates from the first group have earnings about 31 points greater than those from the second group; among high school graduates, on the other hand, earnings are 32 log points *less*. Therefore, about half of the 63.5 log point difference between these two groups in the return to college is due to the difference in college earnings, and half is due to the difference in high school earnings.

One interpretation of the relatively high earnings for college graduates from the poverty group is selection: members of this group who earn a bachelor's degree may have particularly high (unobserved) skills, which are likely needed to overcome the greater barriers this group faces in getting a college degree. On the flip side, the lower high school earnings of this group suggest its members face particularly dire job prospects with limited education, perhaps because they lack the social capital (job contacts or reputation) needed to get good jobs without a higher degree (Putnam 2015).

Looking Beyond Conditional Means to Other Moments of the Earnings Distribution

Thus far we have looked only at the conditional means of earnings for different groups. In Table 8, we investigate other moments of the earnings distribution, which turns out to be important in understanding the differential college premium by family income background.

[20] Of the individuals who grew up in poverty in the sample, 15.8 percent earned a bachelor's degree by age 25. In ascending order for the remaining income groups, these shares are 18.5 percent, 36.7 percent, and 66.0 percent. Among those with bachelor's degrees by age 25, the shares that ever obtain a graduate degree are 22.2, 30.8, 31.0, and 40.8 percent. Thus, it is interesting that those who grow up in poverty have the highest college premium despite the smallest share earning a graduate degree. That the college premium is proportionally lower for those below the 185 percent of poverty threshold, but not at the 100 percent threshold, is a result of the relatively greater number of college graduates (and relatively fewer high school graduates) who grew up with family incomes between 100 and 185 percent of the poverty line.

First, we examine the impact of outliers by simply eliminating all observations that are in the top 1 percent of the sample earnings distribution in each survey wave.[21] Doing so greatly reduces the differential in the observed college earnings premium between the two income background groups, and the remaining difference is no longer statistically significant. The predominant effect of dropping the top 1 percent of earnings observations mainly is to reduce mean expected earnings for the higher-income background group with a college degree, from $1.6 million in lifetime earnings to $1.2 million. This lowers the log return to college education for the higher-income background group from 86 points to 60 points, which is no longer significantly greater than the (almost unchanged) 53-point return for the low-income group. (Absolute dollar returns are still significantly greater, at $204,000, but this is less than half the difference when the full earnings distribution is included.) These results suggest, in line with the graduate degree results in Table 6, that the baseline numbers are driven in no small part by some college graduates from the higher-income background group who obtain very high earnings in some years. We do not believe that this apparent differential access to the far-right tail of the earnings distribution—even for college graduates—for those from different income backgrounds obviates our baseline results. On the contrary, a host of work by Piketty and Saez and many others documents the fastest relative income growth at the top of the earnings distribution.[22] If individuals from low-income backgrounds are less likely to have access to this part of the earnings distribution—even with high levels of education—there may be considerable consequences for both earnings inequality and intergenerational mobility.

To further illustrate the importance of the tails of the earnings distribution, we consider earnings at different percentiles of that distribution, specifically earnings at the 25th percentile, the median, the 75th percentile, and the 90th percentile. These are shown in the remaining columns of Table 8. At the 25th percentile, the log return to college is actually much higher (175 points) for individuals from low-

[21] Earnings are *not* top-coded in the PSID. The 99th percentile of sample earnings (in $2014) varies from approximately $75,000 in the mid-1970s, when the oldest sample respondent is in her mid-20s, to about $400,000 before the Great Recession, when the oldest sample respondent is in her late 50s.

[22] Autor, Katz, and Kearney (2008) provide a review.

income backgrounds than it is for those from high-family-income backgrounds (87 points), and even the absolute dollar return is slightly higher. This apparent reversal from baseline occurs for two reasons. First, at the 25th percentile of the earnings distribution, earnings for high school graduates from a low-income family essentially collapse, decreasing from $475,000 at the mean to just $83,000 at the 25th percentile. Earnings for high school graduates from higher-income families fall as well, but not by as much, from $661,000 to $226,000. Apparently, the lower tail of earnings for low-income-background high school graduates is quite low indeed. Second, at the 25th percentile, earnings for college graduates are not that different for individuals from the different income backgrounds; both are approximately $500,000. In conjunction with the results from excluding the extreme right tail, these patterns suggest not only a much larger variance in earnings for college graduates coming from a higher-income background than for college graduates from poorer families, but that the lower tail of earnings is fairly similar for both groups. Put differently, a higher-income background—conditional on having a bachelor's degree—stretches the right tail of the earnings distribution.

Switching to the median, we find that the relative (log) return to college is slightly (not statistically significantly) higher for the low-income group, and that the absolute dollar returns are similar. Thus, the median or typical college graduate enjoys an earnings return that does not vary much with her family income-background, in contrast with the "average" college graduate. This again emphasizes the importance of the tails of the earnings distribution.

The higher percentage return to college for the higher-income-background group begins to re-emerge as we move to the 75th and 90th percentiles. However, these percentage returns are only weakly significantly higher for the higher-income-background group, in part because the standard errors for the low-income-background group tend to increase, due to diminished density, higher up in the earnings distribution. The dollar returns to a college degree, however, are both substantively and statistically much greater for the higher income group at the 75th and 90th percentiles. For example, at the 90th percentile, a college degree increases lifetime earnings for the higher-income-background group by almost $1.4

million, compared with just under $500,000 for the low-income-background group. This difference of almost $900,000 is clearly statistically (and economically) significant.

Overall, a striking pattern is that at the various percentiles, the dollar return to a college degree remains fairly flat, between roughly $400,000 and $500,000, for the low-income background group. In contrast, the dollar return to a college degree for the higher-income background group dramatically escalates as when moving from lower to higher percentiles, increasing from a little over $300,000 to almost $1.4 million between the 25th percentile and 90th percentiles. Once again, the right tail of the earnings distribution for college graduates who come from low-income families is considerably shorter than it is for their peers from more-affluent families.

This quantile analysis is illustrative, but the importance of location in the earnings distribution in contributing toward the family background college earnings gap can perhaps most clearly be seen by looking at the entire earnings distribution. We present the summary measure of the ratio difference in Figure 2. In the figure, the x-axis represents the percentile of the cumulative career earnings distribution, and the y-axis represents the difference in the ratio at a given cumulative earnings percentile.[23] Since the bottom quarter of the earnings distributions consists of zeroes or very small values, which make ratios unstable or infeasible to calculate, we focus on the top three quartiles of the distribution. From the 25th percentile to about the 45th percentile, the calculated ratio difference is negative, implying that the observed proportional college earnings premium is actually higher for individuals from low-income families than it is for those from higher-income families. However, as the bootstrapped confidence intervals show, a null of no difference in the ratios cannot be rejected. The ratio difference is actually weakly positive at the median, but again a zero difference lies within the confidence interval, consistent

[23] To calculate these distributions for each family-income–education group, we create 5,000 bootstrap replication draws of individuals, with replacement, from the analytic sample. For each group in each replicate draw, we estimate the empirical cumulative distribution function at each age, and then sum across ages for each centile of the cumulative distribution function. This yields 5,000 replicates of cumulative career earnings at each centile for each group. We calculate the ratio difference for each centile of each replicate and then take the median of the ratios across replicates.

with the estimate for the median in Table 8. The difference continues to grow monotonically through the earnings distribution, becoming statistically significant at the 75th percentile, and quite large above the 95th percentile.

Variation in Results by Gender and Race

We now consider how our baseline results vary when we stratify by gender or race, in an attempt to more fully understand the causes or at least correlates of these varying returns to college by family income background. These results are presented in Table 9.

By gender, the lower percentage returns to college for lower-income-background individuals is a phenomenon that is entirely driven by men and does not occur at all for women. For men, the log return to college is 34 points for the lower-income-background group and 99 points for the higher-income-background group. This difference of 65 log points is far greater than the 33 points in the baseline sample that pooled both genders. The relative increase of 32 points in the income-background differential is about two-thirds due to a higher college earnings premium (from an income-background premium among college graduates of 66 log points at baseline to 86 points for men, an increase of 20 points), and about one-third due to a reduction in the high school earnings differential (from an income-background premium among high school graduates of 33 points at baseline to 21 points for just men, a reduction of 12 points).[24] Low-income background men with only a high school education have much higher earnings than the pooled sample of individuals with the same education and family-income background, with lifetime earnings of $645,000 compared to $475,000; given the well-known gender patterns in employment and occupation among these groups, this is perhaps unsurprising. Moreover, high-income-background men with a college degree do extraordinarily well, with average lifetime earnings of $2.14 million, compared to $1.56 million for the equivalent combined gender sample, and far greater than the

[24] That is, earnings for college graduate men diverge more across income backgrounds than do earnings for the pooled sample, while earnings for high school graduate men diverge less across income backgrounds.

$905,000 average earnings of college graduate men from poorer backgrounds. Consequently, the differential college premium by income background is exceedingly high for men.

For women, in contrast, the percentage earnings return to college is virtually the same regardless of family background. The log return to college is about 69 points for women from both family-income-background groups. College graduate women from higher-income backgrounds do not average the extraordinarily high earnings that their male counterparts do, making $1.04 million to their $2.14 million for their male counterparts. Furthermore, women from a low-income background with only a high school diploma do not fare as well earnings-wise as their male counterparts. These low-income-background high school graduate women have expected lifetime earnings of $354,000, compared with $645,000 for their male counterparts. These two differences depress the college return for women from a high-income background and increase the college return for women from a low-income background, which on net yield similar college premia.

By race, the differential higher college returns for individuals from a higher-income background are quite strong for whites but do not occur for blacks. Looking first at whites, the log return to college for higher-income background individuals is 84 points, whereas it is only 11 points for those from a low-income background, and this latter return is not significantly different from zero. That is, white college graduates from a low-income background on average earn only $66,000 more over their careers than their high school graduate peers from the same income background. Because college graduates from higher-income families have much greater earnings, the relative college premium is much larger for them.

Conversely, blacks have relatively high log returns to college that vary little by family-income background: 103 points for the higher-income group and 101 points for the low-income group. The striking difference between blacks and whites occurs among the low-income group, where college graduate earnings are much greater for blacks, at $1.03 million, than for whites, at $622,000. On the other hand, for the low-income background group, black high school graduate earnings are just $378,000, substantially less than the $555,000 of their white counterparts. Thus, the relatively high college premium

for blacks is driven both by low earnings for high school graduates and relatively high earnings for college graduates, a phenomenon that exists for both income background groups.

The finding that the differential college earnings premium by family-income background persists and is even larger for men and whites—groups that are highly advantaged in the labor market—but completely disappears for women and blacks again points to the importance of the right tail of the earnings distribution.

Some Further Results for Men and Whites

For men and whites, the groups for which we see the largest differences across income background in the relative returns to a bachelor's degree, we explore how our results vary across some of the dimensions previously explored in this paper. Specifically, we see how results change when we restrict the sample to positive earnings, omit holders of graduate degrees, look at other income background thresholds, drop the top 1 percent of earnings observations, and focus on conditional median earnings rather than conditional mean earnings.

These results are presented for men in Table 10 and for whites in Table 11. In both cases, the income-background differential in the college premium is stronger and more robust than for the overall sample. Indeed, for both men and whites, the differential college earning premium persists even when dropping the (group-specific) top 1 percentile of earnings. In fact, the differential premium is about as high in both these cases as it is for the overall sample over the full earnings distribution. When removing individuals who ever obtain graduate degrees from the male and white samples, the relative differential is again reduced substantially and loses statistical significance. However, whereas in the overall sample the differential was reduced to an economically insignificant 0.6 log points, the differentials remain economically meaningful (if imprecisely estimated due to smaller sample sizes), at 22 log points for men and 27 log points for whites. In more practical terms, these translate into percentage college premia that are one-and-one-half times to twice as large for individuals from higher-income families than those from low-income families. Thus, while graduate degrees appear to drive much of the differential college

earnings premium, they do not drive all of it. Of course, because the earnings distributions of men and whites are shifted to the right relative to that of the overall sample, the right tail of the distribution is effectively larger, which likely explains why dropping the top 1 percent of earnings or graduate-degree recipients reduces, but does not eliminate, the differential college premium.

For the remaining sample and moment variants, the pattern of results is quite similar to what we previously found. The contrast in relative returns is little changed when we focus on only positive earnings observations, is increased when we compare a very-high-income-background group (greater than 400 percent of the poverty line) to a near-poor group (100–200 percent of the poverty line), and is sharply reduced when we focus on the conditional median of earnings rather than on the conditional mean.

CONCLUSION

These results are clearly descriptive and barely scratch the surface in investigating heterogeneity across family income background in the observed returns to higher education. Nonetheless, they suggest that the large differences in the observed percentage return to education across individuals from different family-income backgrounds are most likely associated with the right tail of the earnings distribution for college graduates from higher-income families. Individuals who earn a graduate degree, are white or male, or come from a family with income above 400 percent of the poverty line, are much more likely to achieve very high earnings than similar individuals from a low-income family, and particularly from a near-poor one (family income between 100–200 percent of the poverty line), whose earnings boost from college is particularly small.

We plan on conducting considerably more work on this topic, given the importance of educational returns to many important public policies related to income distribution and poverty. More specifically, we will continue investigating mechanisms that explain the tails of the earnings distribution,

as well as allow for more accurate inferential methods.[25] We also aim to further and more flexibly clarify where in the socioeconomic spectrum college return differentials are most concentrated.

In addition, we plan to explain more precisely how these education return differentials manifest over the life course, from market factors such as labor force participation and experience, occupation, and industry, as well as "premarket" factors including postsecondary education characteristics (type of degree, field of study, and selectivity of college attended) and neighborhood and family characteristics (racial and income segregation, single parenthood, and others) that previous research has identified in influencing labor market outcomes.[26] We are also interested in whether the differentials have evolved across cohorts as wage inequality and the return to college have increased.

[25] Currently our standard errors are calculated via the delta method, which is unlikely to perform well for quantile estimates and nonlinear combinations of parameter estimates that may not be sufficiently asymptotically normal. We are exploring various bootstrap approaches.

[26] Kearney and Levine (2016), using the NLSY79, find suggestive evidence that the observed rate of return to an additional year of education is lower in states with greater inequality as measured by the 50/10 household income ratio; and the inequality gradient is steeper among children with less-educated mothers.

APPENDIX A: CONSTRUCTION OF DATA

To investigate the relationship between economic circumstances during adolescence and the observed earnings premium to higher education, we use a longitudinal data set: the Panel Study of Income Dynamics (PSID).

PSID

The PSID is the world's longest, continuously run panel study of households. It began tracking approximately 18,000 individuals from 5,000 families in 1968 and has followed members of these families and their descendants as they have formed their own households. In addition to extensive information in income and demographics, the PSID collects data on employment, education, wealth, marriage and fertility, health, spending, personal behavior, and other topics. Surveying took place annually between 1968 and 1997 and biennially afterward; the most recent wave released is from 2013.

The PSID was designed to be nationally representative of U.S. households in 1968, but its sampling frame means that it is less representative of U.S. households today, underrepresenting immigrant and Latino households, in particular. (A short-lived Latino sample was implemented for a few years starting in 1990, and a longer-lasting immigrant sample was implemented in the late 1990s.) In addition to the core, nationally representative, sample of households that began in 1968 (SRC sample), the PSID oversampled lower-income households as part of the Survey of Economic Opportunity (SEO sample). Because of the short time frame of the Latino and immigrant samples, we restrict our focus to the original cross-sectional (SRC) and SEO samples. Sample weights are provided to make the combined samples nationally representative, and we discuss the construction of our analytic sample below.

Determination of Family Income Background

For our purposes of determining household income during adolescence, it is critical to link individuals to their parents. We perform this linkage using the variables for 1968 interview number and

person number, ER30001 and ER30002, respectively. Because we are interested in average family income when an individual is between the ages of 13 and 17, we use the reported age variable after some cleaning[27] to identify the family income years of interest. (Since the income year is the calendar year prior to the year of interview, we account for the adjustment in children's ages to the previous year.) For each income year in which the child is between 13 and 17, we take the family income (if neither head's nor spouse's earnings are imputed by PSID) and divide this measure by the Census poverty threshold for the family, provided directly as a PSID variable, to obtain an income-to-needs ratio for the family that year.

While in some cases it is possible to calculate this ratio for each year the child is between the ages of 13 and 17, it is more common that we observe family income at only a few points over this age interval; for the oldest cohorts for which we can identify family income (those born in 1950), we observe income only for the year in which the cohort was 17 years old. Following the advice of Solon (1992), we take the mean of the ratio measures over the points observed for each individual in order to calculate a final income-to-needs ratio.[28] Individuals with a ratio no greater than 185 percent are assigned to low-income family background status; these individuals would nominally be eligible for the federal assisted lunch program, one of the more common indicators of low-income background observed in student populations. Individuals with a ratio greater than 185 percent are assigned to non-low-income family background status.

As shown in Table A.1, there are 9,144 individuals in the PSID for whom we can identify parental income background.

[27] The reported age contains some errors of nonmonotonicity and invalid skips. We clean the variable with the following procedure: using the reported age in wave *t*, we impute the age in every other wave using the year of survey and the reported age. We then calculate the sum of squared deviations from the actual reported age in each wave and the imputed age. We repeat this process for all other waves and find the wave with the smallest sum of squared deviations; we assign the imputed age from this wave as the analytic age.

[28] We obtain nearly identical assignments if we instead calculate average family income and average poverty threshold and then divide the first average by the second. We prefer the average of ratios because family composition is not necessarily fixed.

Valid Own Earnings

These 9,144 individuals are collectively interviewed 256,614 times through the 2013 interview year (out of a possible maximum of 304,211 times in which they were eligible to be interviewed). Most of these interview periods, however, occur during childhood. Once we restrict the sample to person-years in which the individual is at least 25 during the earnings year (or at least 26 during the interview year), Table A.1 shows that the count falls to 90,108. Of these 90,108 observations, 82,540 had valid, non-imputed labor earnings, including zeros. The difference between these numbers is only partially due to the exclusion of imputed earnings (the PSID has a far smaller proportion of earnings imputations than the March CPS, for example). Subsection D below discusses this selection issue. Note that our earnings measure includes both wages and salaries as well as the labor portion of farm and own-business income.

Because we are interested in the earnings profiles of specific education groups, the last six lines of Table A.1 show unique person and person-year (observation) counts for those with high school and college-level education by age 25. More specifically, the first three of these rows show counts for those with exactly 12 years of completed schooling, those with exactly 16 years of schooling, and those with 16 or more years of schooling. The latter three rows are instead based on diploma and degree attainment, rather than grade completion. This constitutes our preferred measure because it more accurately captures the educational credential received; however, our results are robust to using the grade completed education measure.[29] Across the two education measures, the degree receipt classification shows substantially fewer high school diplomas than the grade completed classification. This is largely due to how GEDs get counted; the later approach does not distinguish between GEDs and high school diplomas in the data while the former approach does, and we exclude GEDs from the high school diploma counts. As Heckman and LaFontaine (2006) show that GEDs do not provide labor market returns comparable to a

[29] A drawback to the degree timing variables is that, unlike highest grade completed, they are not asked of every sample member in each year. Rather, degree type and timing are generally asked only of heads and spouses, and even then only beginning in 1985.

high school diploma, using degree receipt instead of grade completed is again preferred. This classification system yields 4,394 individuals with 46,789 earnings-year observations.

Cohort Distribution

Table A.2 shows the distribution of individuals and observations across the two PSID core samples (SRC/cross-sectional and SEO/low-income), education classification, family income background, and birth cohort. Unsurprisingly given the large amount of research on socioeconomic status and educational attainment (see Bailey and Dynarski [2011] for a review), individuals who grow up low-income are less likely to obtain a bachelor's degree. Our preference for combining the SRC and SEO core samples (and applying sample weights) is motivated to maintain large enough sample sizes for college graduates who grow up low-income, although we note that our results are unchanged (albeit less precise) using just the cross-sectional sample, with or without weights.

The nature of the data—a panel that individuals enter as they age—implies that more sample observations are for earlier birth cohorts, as these cohorts are in the analytical sample for longer and at older ages. Whereas the median birth year for individuals in the sample is 1964, the median birth year for observations is 1958. This structure also implies that estimated earnings at earlier ages will comprise more cohorts than at older ages.

Attrition and Selection

As with any panel survey—and especially one as long as the PSID—careful attention must be paid to how attrition could influence the estimates. In particular, career earnings estimates by education and family income background could be biased if one of these groups was differentially likely to attrite. In Table A.3, we present estimates from a linear probability model of the likelihood of attriting (not being interviewed for whatever reason) across survey waves by education group, low-income background, and the interaction of the two. Each regression also includes dummies for the survey year, and the four columns present results with and without survey weights and age dummies for individuals. Under the

preferred specification with both of these controls (column 4), individuals who grow up low-income are about 5 percentage points more likely to attrite and college graduates about 6 percentage points less likely to do so, relative to a mean of 20 percent. Importantly, however, the estimate for the interaction between college graduate and low-income background is small (about 2 percentage points) and not statistically significant in any specification with age controls or weights. Thus, college graduates who grow up low-income are not differentially likely to attrite relative to college graduates who grew up in more favorable economic circumstances. While this finding does not completely eliminate the threat of attrition to cause bias in the estimates (e.g., through within-group unobserved heterogeneity), it suggests that attrition is not a first-order concern.

Another but related issue is differential selection into head or spouse status. The PSID regularly collects earnings information only for someone who is the head or spouse of a family unit, so adult children living with their parents (or adult siblings) will have missing earnings information. This issue is mitigated to some extent by looking at earnings at ages 25 and older, when most children have become heads or spouses, but there is the potential that selection into these statuses varies by education and family income background. Table A.4 investigates this possibility by estimating the effect of education, income background, and their interaction on the likelihood of being a head or spouse, with the format similar to Table A.3. Although family income background does not seem to substantively affect being a head or spouse, college graduates are about 5 percentage points less likely to be a head or spouse in any given wave, off of a mean of 87 percent. Once more, however, there is no differential impact of college graduates who grow up low-income relative to those who do not. Thus, selection into earnings eligibility should not present undue bias when comparing observed education earnings premia by family-income background.

REFERENCES

Alexander, Karl, Doris Entwhistle, and Linda Olsen. 2014. *The Long Shadow: Family Background, Disadvantaged Urban Youth, and the Transition to Adulthood*. New York: Russell Sage Foundation.

Altonji, Joseph G., and Thomas A. Dunn. 1996. "The Effects of Family Characteristics on the Return to Education." *Review of Economics and Statistics* 78(4): 692–704.

Ashenfelter, Orley, and Cecilia Rouse. "Income, Schooling, and Ability: Evidence from a New Sample of Identical Twins." *Quarterly Journal of Economics* 113: 253–284.

Autor, David, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. 2015. "Family Disadvantage and the Gender Gap in Behavioral and Educational Outcomes." Northwestern University Working Paper.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney. 2006. "The Polarization of the U.S. Labor Market." *American Economic Review* 96(2): 189–194.

———. 2008. "Trends in U.S. Wage Inequality." *Review of Economics and Statistics* 90(2): 300–323.

Bailey, Martha J., and Susan M. Dynarski. 2011. "Inequality in Postsecondary Education." In *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*, Greg J. Duncan and Richard J. Murnane, eds. New York: Russell Sage.

Bartik, Timothy J., Brad Hershbein, and Marta Lachowska, 2016. "The Merits of Universal Scholarships: Benefit-Cost Evidence from the Kalamazoo Promise." Upjohn Institute Working Paper No. 16-252.

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

Brand, Jennie E., and Yu Xie. 2010. "Who Benefits Most from College? Evidence for Negative Selection in Heterogeneous Economic Returns to Higher Education." *American Sociological Review* 75(2): 273–302.

Brown, Charles. 1996. "Notes on the 'SEO' or 'Census' Component of the PSID." Mimeo.

Card, David. 1999. "The Causal Effect of Education on Earnings." In *Handbook of Labor Economics 3A*, Orley Ashenfelter and David Card, eds. Amsterdam: Elsevier, pp. 1801–1863.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil. 2001 "Estimating the Return to Education When It Varies among Individuals." University of Chicago Working Paper.

———. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101(6): 2754–2781.

Chetty, Raj, and Nathaniel Hendren. 2015. "The Impacts of Neighborhoods on Intergenerational Mobility: Childhood Exposure Effects and County-Level Estimates." Harvard University Working Paper.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where Is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics* 129(4) 1553–1623.

Chetty, Raj, Nathaniel Hendren, Jeremy Majerovitz, and Benjamin Scuderi. 2016. "Childhood Environment and Gender Gaps in Adulthood." NBER Working Paper No. 21936.

Deming, David J. 2015. "The Growing Importance of Social Skills in the Labor Market." NBER Working Paper No. 21473.

Fitzgerald, John, Peter Gottschalk, and Robert Moffitt. 1998. "The Impact of Attrition in The Panel Study of Income Dynamics on Intergenerational Analysis." *Journal of Human Resources* 33(2): 300–344.

Heckman, James J., and Paul LaFontaine. 2006. "Bias Corrected Estimates of GED Returns." *Journal of Labor Economics* 24(3): 661–700.

Hershbein, Brad, Melissa S. Kearney, and Lawrence H. Summers. 2015. "Increasing Education: What It Will and Will Not Do for Earnings and Earnings Inequality." Brookings Institution mimeo, Washington, DC.

Hirvonen, Lalaina H. 2008. "Intergenerational Earnings Mobility among Daughters and Sons: Evidence from Sweden and a Comparison with the United States." *American Journal of Economic and Sociology* 67(5): 777–826.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *Quarterly Journal of Economics* 107(1): 35–78.

Kearney, Melissa S., and Philip B. Levine. 2016. "Income Inequality, Social Mobility, and the Decision to Drop Out of High School." *Brookings Papers on Economic Activity* Spring: 333–396.

Lemieux, Thomas. 2006. "Increasing Residual Wage Inequality: Composition Effects, Noisy Data, or Rising Demand for Skill?" *American Economic Review* 96(3): 461–498.

Lundborg, Petter, Anton Nilsson, and Dan-Olof Rooth. 2014. "Parental Education and Offspring Outcomes: Evidence from the Swedish Compulsory School Reform." *American Economic Journal: Applied Economics* 6(1): 253–278.

Mazumder, Bhashkar. 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *Review of Economics and Statistics* 87(2): 235–255.

Nilsen, Oivind Anti, Kjell Vaage, Arild Aakvik, and Karl Jacobsen. 2012. "Intergenerational Earnings Mobility in Norway: Levels and Trends." *Scandinavian Journal of Economics* 114(1): 1–23.

Putnam, Robert D. 2015. *Our Kids: The American Dream in Crisis*. New York: Simon and Schuster.

Shin, Donggyun, and Gary Solon. 2011. "Trends in Men's Earnings Volatility: What Does the Panel Study of Income Dynamics Show?" *Journal of Public Economics* 95(7-8): 973–982.

Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *American Economic Review* 82: 393–408.

Solon, Gary, Steven J. Haider, and Jeffrey Wooldridge. 2015. "What Are We Weighting For?" *Journal of Human Resources* 50(2): 301–316.

Stark, Patrick, and Amber M. Noel. 2015. *Trends in High School Dropout and Completion Rates in the United States: 1972–2012*. Washington, DC: U.S. Department of Education.

Torche, Florencia. 2011. "Is a College Degree Still the Great Equalizer? Intergenerational Mobility across Levels of Schooling in the United States." *American Journal of Sociology* 117(3): 763–807.

Figure 1A Career Earnings Profile by Education, ≤185% FPL

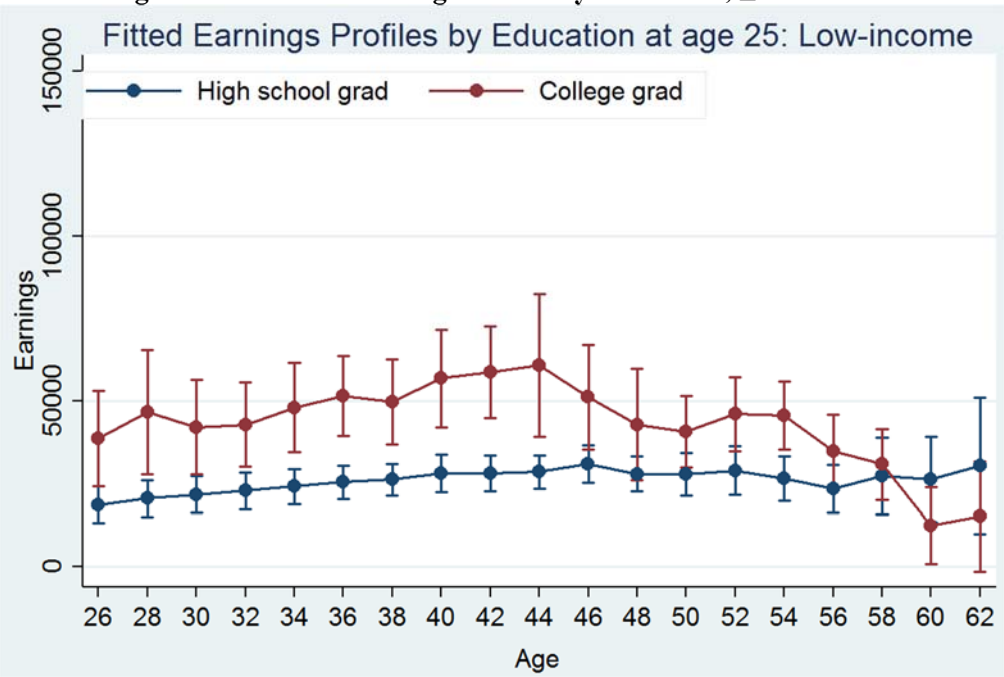
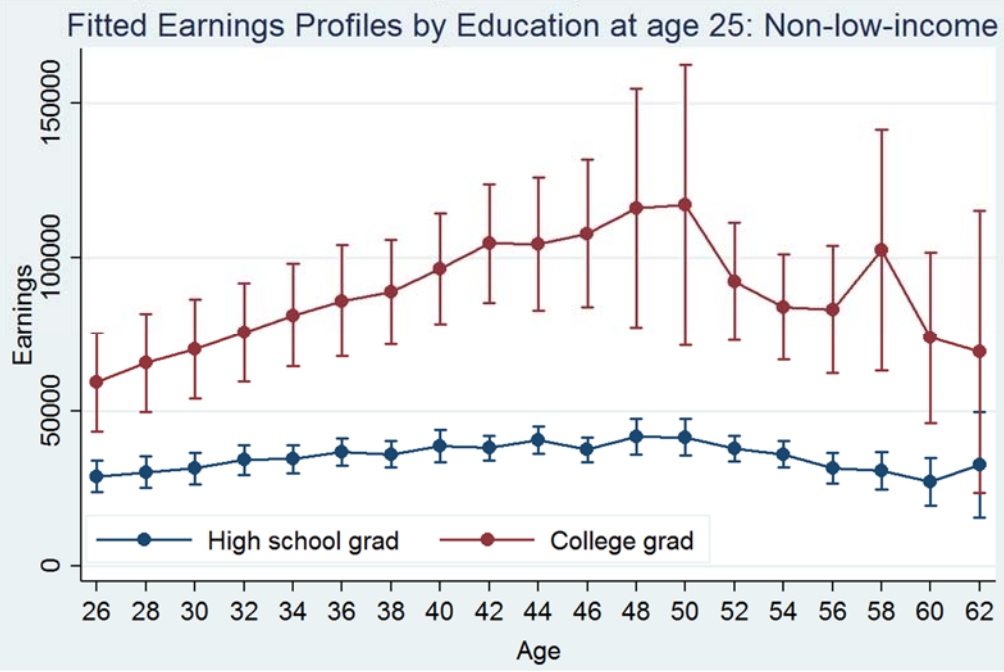
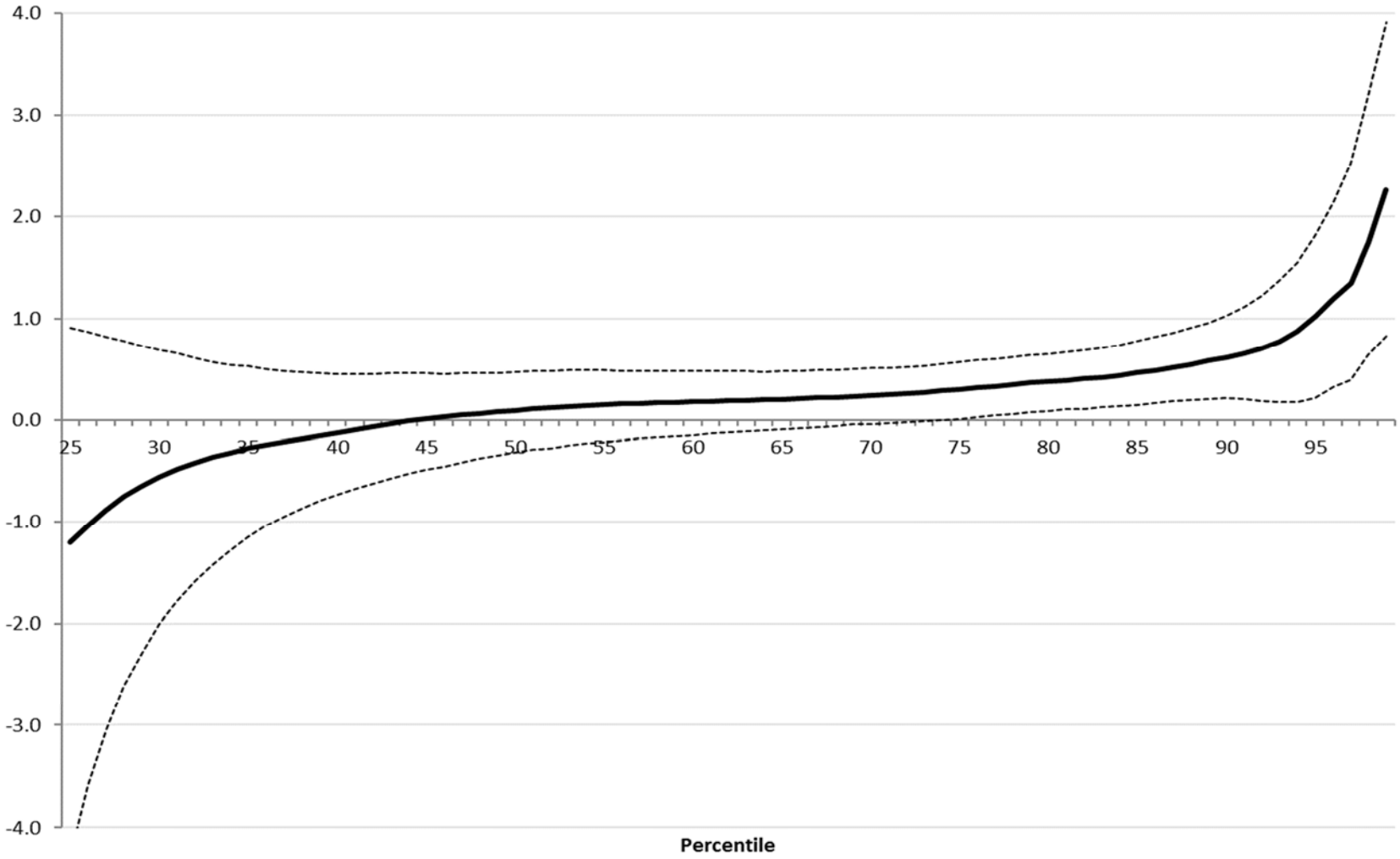


Figure 1B Career Earnings Profile by Education, >185% FPL



SOURCE: Authors' calculations from the PSID.
NOTE: Mean earnings by age are in year 2014 dollars, adjusted with the PCE deflator from the Bureau of Economic Analysis, and are calculated including zeros but dropping imputations. Whiskers represent 95 percent confidence intervals for each point estimate, where the underlying standard errors are robust to heteroskedasticity and intraperson correlation.

Figure 2 Difference in College–High School Career Earnings Ratios, between Higher and Low-Income Backgrounds, by Percentiles



SOURCE: Authors' calculations from the PSID.
NOTE: The solid line plots the difference between the career college/high school earnings ratio for individuals from families with earnings above 185 percent of the federal poverty threshold and the same ratio for individuals from families with earnings below 185 percent of the federal poverty threshold. See text for details of this construction. The dashed lines represent 95 percent confidence intervals calculated from the bootstrap replications (i.e., the 2.5th percentile of the replications at each percentile of the earnings distribution on the x-axis represents the lower confidence bound, and the 97.5th percentile of the replications at each earnings percentile represents the upper confidence bound.

Table 1 PSID Sample Summary Statistics

	All		High school, low-income		High school, not low-income		College, low-income		College, not low-income	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Annual earnings ($2014)	46,002	70,450	26,292	27,126	34,340	35,305	42,829	40,368	68,125	102,282
Birthyear	1960.3	8.4	1959.4	8.2	1959.5	7.5	1958.6	8.9	1961.7	9.0
Age	36.61	8.82	36.29	8.70	36.92	8.88	37.66	9.37	36.31	8.74
Earnings year	1996.9	9.7	1995.7	9.8	1996.4	9.5	1996.3	10.2	1998.0	9.6
Female	0.536	0.499	0.583	0.493	0.524	0.499	0.595	0.491	0.522	0.500
Black	0.124	0.329	0.358	0.479	0.084	0.278	0.427	0.495	0.027	0.163
Other race	0.019	0.137	0.035	0.184	0.015	0.122	0.011	0.104	0.017	0.129
FPL % when 13–17	367.5	303.3	123.2	41.3	347.7	153.2	131.0	41.3	526.2	397.6
Share FPL < 185%	0.214	0.410								
HS, low-income	0.177	0.381								
HS, not low-income	0.408	0.492								
College, low-income	0.037	0.188								
Coll., not low-income	0.378	0.485								
N (observations)	47,254		15,322		16,376		2,586		12,970	
Unique individuals	4,395		1,358		1,460		227		1,353	

SOURCE: Authors' calculations from the PSID.
NOTE: See Appendix A for details of sample construction. Annual earnings include zeros. The FPL % is family income expressed as a percentage of the federal poverty threshold for that family. Statistics calculated using PSID-supplied sample weights.

Table 2 Present Discounted Value of Career Earnings, by Education and Family Income Background

	Earnings ($2014)	College − High school	College / High school	Difference-in-differences	Difference-in-ratios
Low-income					
High school grad	474,522				
	(31,622)	335,062	1.706		
		(77,223)	(0.187)		
College grad	809,583				
	(70,452)				
				565,791***	0.657**
Non-low-income				(154,921)	(0.289)
High school grad					
	661,017				
	(25,673)	900,853	2.363		
		(134,302)	(0.220)		
College grad	1,561,870				
	(131,825)				

SOURCE: Authors' calculations from the PSID.
NOTE: * p<0.10, ** p<0.05, *** p<0.01. Cumulative earnings from ages 25–62, taken from the PSID sample as described in the text, are discounted at an annual rate of 3 percent from the perspective of an 18-year-old. Standard errors robust to heteroskedasticity and intrapersonal correlation and calculated via the delta method are in parentheses.

Table 3 Decomposition of Differences across Income Background Groups in the Percentage Return to College: College Earnings versus High School Earnings

Log point return to college, non-low-income	86.0
Standard error (SE)	(9.3)
Log point return to college, low-income	53.4
SE	(11.0)
Difference in log point return to college	32.6
SE	(14.4)
Difference in log college earnings: Non-low-income minus low-income	65.7
SE	(12.1)
Difference in log high school earnings: Non-low-income minus low-income	33.1
SE	(7.7)

SOURCE: Authors' calculations from the PSID.
NOTE: This table takes the data in Table 2 and calculates the (natural) logarithmic differences in relative earnings with family income background. These are stated as logarithmic percentage differences, which are 100 times the actual difference in logarithmic terms. Thus, 86.0 in the above table is equal to 100 times [ln(1,561,870) – ln(661,017)] from Table 2. The advantage is that the difference in the log percentage return to college across income background groups exactly equals the difference in log college earnings minus the difference in log high school earnings (i.e., additivity).

Table 4 Differences in Earnings across Income Background Groups, by Educational Attainment and Age

Age	26	28	30	32	34	36	38	40	42	44	46	48	50	52	54	56	58	60	62
High School: Non-low-income minus low-income	40	36	35	38	34	36	32	33	32	37	22	44	45	32	37	36	17	6	8
	(12)	(12)	(12)	(11)	(11)	(10)	(10)	(11)	(11)	(10)	(11)	(12)	(14)	(15)	(16)	(20)	(26)	(30)	(44)
College: Non-low-income minus low-income	36	28	46	53	50	48	56	52	57	54	78	109	114	76	67	94	125	193	153
	(17)	(20)	(16)	(14)	(14)	(14)	(14)	(15)	(15)	(21)	(21)	(32)	(26)	(19)	(18)	(23)	(29)	(61)	(65)

SOURCE: Authors' calculations from the PSID.
NOTE: Age gives results for two-year age bins. All differences across income groups are stated as 100 times [ln(earnings for higher-income background group at that age bin for that education category) minus ln(earnings for lower-income background group for that same age bin and education category)]. Standard errors (in parentheses) are adjusted to be compatible (that is, also multiplied by 100). These log differences are from the same baseline specification that underlies Figure 1 and Tables 2 and 3.

Table 5 Sensitivity of Results to Inclusion of SEO Sample and to Weighting

	Baseline: SRC-SEO weighted		SRC weighted		SRC unweighted		SRC-SEO unweighted	
	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income
PV of HS earnings	661.0	474.5	671.9	514.0	658.1	485.0	615.7	414.4
	(25.7)	(31.6)	(27.1)	(41.9)	(25.0)	(38.0)	(20.4)	(18.1)
PV of coll. earnings	1,561.9	809.6	1,569.0	773.2	1,480.0	752.1	1,449.0	832.8
	(131.8)	(70.5)	(134.1)	(84.5)	(97.6)	(71.4)	(89.6)	(49.7)
Difference	900.9	335.1	896.8	259.2	822.1	267.1	832.8	418.4
	(134.3)	(77.2)	(136.8)	(94.3)	(100.7)	(80.9)	(91.9)	(52.9)
Ratio	2.363	1.706	2.335	1.504	2.249	1.551	2.353	2.010
	(0.220)	(0.187)	(0.221)	(0.205)	(0.171)	(0.191)	(0.165)	(0.149)
Ratio difference		0.657**		0.831***		0.698***		0.343
		(0.289)		(0.301)		(0.256)		(0.222)
Ln college return	86.0	53.4	84.8	40.8	81.0	43.9	85.6	69.8
	(9.3)	(11.0)	(9.5)	(13.6)	(7.6)	(12.3)	(7.0)	(7.4)
Difference ln return		32.6**		44.0***		37.2***		15.8
		(14.4)		(16.6)		(14.5)		(10.2)
Absolute difference college return		565.8***		637.5***		555.0***		414.4***
		(154.9)		(166.1)		(129.2)		(106.1)

SOURCE: Authors' calculations from the PSID.
NOTE: * p<0.10, ** p<0.05, *** p<0.01, shown for cross-differences only. All earnings numbers are in thousands of dollars and are the present value of estimated lifetime earnings for a particular income background/educational attainment group. Logarithmic differences are 100 times the difference in natural logarithm of indicated variables. Standard errors (in parentheses) are robust to heteroskedasticity and intraperson correlation. SRC is Survey Research Center sample component of PSID; SEO is Survey of Economic Opportunity component of PSID (oversample of low-income families in poverty). Differences in the right column of each column pair always represent the higher-income group minus the lower-income group. See text for description of family income background.

Table 6 Sensitivity of Results to Sample Restrictions

	Baseline estimates		Includes age 20+ earnings		Drop zero earnings		Drop graduate degrees	
	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income
PV of HS earnings	661.0	474.5	790.9	561.0	785.6	635.2	658.7	429.5
	(25.7)	(31.6)	(31.8)	(36.8)	(27.1)	(37.0)	(31.1)	(31.4)
PV of coll. earnings	1,561.9	809.6	1,721.4	898.7	1,751.7	931.1	1,233.8	799.8
	(131.8)	(70.5)	(144.8)	(83.7)	(145.7)	(71.8)	(77.1)	(78.1)
Difference	900.9	335.1	930.5	337.7	966.1	295.9	575.1	370.3
	(134.3)	(77.2)	(148.2)	(91.5)	(148.2)	(80.7)	(83.2)	(84.2)
Ratio	2.363	1.706	2.176	1.602	2.230	1.466	1.873	1.862
	(0.220)	(0.187)	(0.203)	(0.183)	(0.201)	(0.142)	(0.147)	(0.227)
Ratio difference		0.657**		0.574**		0.764***		0.011
		(0.289)		(0.273)		(0.246)		(0.270)
Ln college return	86.0	53.4	77.7	47.1	80.2	38.2	62.8	62.2
	(9.3)	(11.0)	(9.3)	(11.4)	(9.0	(9.7)	(7.8)	(12.2)
Difference ln return		32.6**		30.6**		42.0***		0.6
		(14.4)		(14.7)		(13.2)		(14.5)
Absolute difference college return		565.8***		592.8***		670.2***		204.8*
		(154.9)		(174.2)		(168.7)		(118.3)

SOURCE: Authors' calculations from the PSID.
NOTE: * p<0.10, ** p<0.05, *** p<0.01, shown for cross-differences only. All earnings numbers are in thousands of dollars and are the present value of estimated lifetime earnings for a particular income background/educational attainment group. Logarithmic differences are 100 times the difference in natural logarithm of indicated variables. Standard errors (in parentheses) are robust to heteroskedasticity and intraperson correlation. Differences in the right column of each column pair always represent the higher-income group minus the lower-income group. See text for description of family income background.

Table 7 Results with Alternative Breakdown of Family Income Background

	Baseline estimates		Alternative income background groupings			
	>185% pov.	≤185% pov.	>400% pov.	200–400% pov.	100-200% pov.	≤100% pov.
PV of HS earnings	661.0	474.5	785.8	616.3	519.0	376.7
	(25.7)	(31.6)	(62.7)	(25.8)	(38.2)	(39.4)
PV of coll. Earnings	1,561.9	809.6	1801.0	1151.0	767.1	1051.1
	(131.8)	(70.5)	(198.5)	(90.6)	(81.3)	(140.8)
Difference	900.9	335.1	1015.0	534.7	248.2	674.3
	(134.3)	(77.2)	(208.2)	(94.2)	(89.8)	(146.2)
Ratio	2.363	1.706	2.292	1.868	1.478	2.790
	(0.220)	(0.187)	(0.312)	(0.166)	(0.191)	(0.474)
Ratio difference		0.657**		0.424	0.814**	-0.498
		(0.289)		(0.353)	(0.366)	(0.567)
Ln college return	86.0	53.4	82.9	62.5	39.1	102.6
	(9.3)	(11.0)	(13.6)	(8.9)	(12.9)	(17.0)
Difference ln return		32.6**		20.5	43.9**	-19.7
		(14.4)		(16.3)	(18.8)	(21.8)
Absolute difference college return		565.8***		480.3**	766.8***	340.7
		(154.9)		(228.5)	(226.8)	(254.4)

SOURCE: Authors' calculations from the PSID.
NOTE: * p<0.10, ** p<0.05, *** p<0.01, shown for cross-differences only. All earnings numbers are in thousands of dollars and are the present value of estimated lifetime earnings for a particular income background/educational attainment group. Logarithmic differences are 100 times the difference in natural logarithm of indicated variables. Standard errors (in parentheses) are robust to heteroskedasticity and intraperson correlation. Differences in the right column of each column pair always represent the higher-income group minus the lower-income group; for alternative income groupings, differences are relative to the highest income group. See text for description of family income background.

Table 8 Other Conditional Moments and Quantiles of the Earnings Distribution

	Baseline		No 99th pctile		25th percentile		Median		75th percentile		90th percentile	
	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income
PV of HS earnings	661.0	474.5	647.7	474.1	225.8	83.2	548.0	360.0	937.0	707.0	1,342.0	1,050.0
	(25.7)	(31.6)	(22.5)	(31.3)	(44.5)	(43.2)	(24.4)	(31.9)	(40.7)	(42.7)	(59.5)	(164.3)
PV of coll. earnings	1,561.9	809.6	1,182.3	805.0	537.0	478.4	1,062.0	803.0	1,731.3	1,097.0	2,719.0	1,546.0
	(131.8)	(70.5)	(32.3)	(71.1)	(29.7)	(103.1)	(41.6)	(85.9)	(55.0)	(72.9)	(149.8)	(293.6)
Difference	900.9	335.1	534.6	330.9	311.2	395.2	514.0	443.1	794.3	389.9	1,377.0	496.0
	(134.3)	(77.2)	(39.4)	(77.7)	(53.5)	(113.3)	(47.6)	(90.1)	(69.9)	(75.8)	(165.0)	(266.7)
Ratio	2.363	1.706	1.825	1.698	2.378	5.749	1.938	2.231	1.848	1.551	2.026	1.472
	(0.220)	(0.187)	(0.081)	(0.187)	(0.487)	(3.275)	(0.113)	(0.302)	(0.102)	(0.123)	(0.148)	(0.274)
Ratio difference		0.657**		0.127		-3.371		-0.293		0.297*		0.554*
		(0.289)		(0.204)		(3.311)		(0.322)		(0.160)		(0.311)
Ln college return	86.0	53.4	60.2	52.9	86.6	174.9	66.2	80.2	61.4	43.9	70.6	38.7
	(9.3)	(11.0)	(4.4)	(11.0)	(20.5)	(56.2)	(5.9)	(13.9)	(5.4)	(9.0)	(7.1)	(24.6)
Difference ln return		32.6**		7.2		-88.3		-14.1		17.4*		31.9
		(14.4)		(11.9)		(59.8)		(15.1)		(10.5)		(25.6)
Absolute difference college return		565.8***		203.7**		-84.0		70.9		404.4**		881.0***
		(154.9)		(87.1)		(125.3)		(101.9)		(103.0)		(313.6)

SOURCE: Authors' calculations from the PSID.
NOTE: * p<0.10, ** p<0.05, *** p<0.01, shown for cross-differences only. All earnings numbers are in thousands of dollars and are the present value of estimated lifetime earnings for a particular income background/educational attainment group. The column pair "No 99th percentile" drops earnings observations in the top percentile in each survey wave. Quantile estimates (including median) are percentiles from the conditional earnings distribution (see Equation [1]). Logarithmic differences are 100 times the difference in natural logarithm of indicated variables. Standard errors (in parentheses) are robust to heteroskedasticity and intraperson correlation. Differences in the right column of each column pair always represent the higher-income group minus the lower-income group. See text for description of family income background.

Table 9 Present Discounted Value of Career Earnings, by Education and Family Income Background, Gender and Race

	Baseline		Men		Women		Whites		Blacks	
	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income
PV of HS earnings	661.0	474.5	794.1	644.6	520.2	354.1	673.5	555.4	508.0	377.9
	(25.7)	(31.6)	(40.9)	(63.0)	(28.0)	(26.0)	(28.9)	(52.8)	(42.5)	(28.8)
PV of coll. earnings	1,561.9	809.6	2,143.1	905.2	1,040.1	703.6	1,556.5	621.8	1,416.5	1,032.0
	(131.8)	(70.5)	(260.1)	(126.9)	(74.1)	(64.3)	(137.0)	(85.7)	(106.4)	(112.4)
Difference	900.9	335.1	1,349.0	260.6	519.9	349.5	882.9	66.4	908.4	654.0
	(134.3)	(77.2)	(263.3)	(141.6)	(79.2)	(69.4)	(140.0)	(100.6)	(114.6)	(116.0)
Ratio	2.363	1.706	2.699	1.404	1.999	1.987	2.311	1.120	2.788	2.731
	(0.220)	(0.187)	(0.356)	(0.240)	(0.179)	(0.233)	(0.226)	(0.187)	(0.314)	(0.363)
Ratio difference		0.657**		1.295***		0.012		1.191***		0.057
		(0.289)		(0.429)		(0.294)		(0.293)		(0.480)
Ln college return	86.0	53.4	99.3	34.0	69.3	68.7	83.7	11.3	102.5	100.5
	(9.3)	(11.0)	(13.2	(17.1)	(8.9)	(11.7)	(9.8	(16.7)	(11.2)	(13.3)
Difference ln return		32.6**		65.3***		0.6		72.4***		2.0
		(14.4)		(21.6)		(14.7)		(19.4)		(17.4)
Absolute difference college return		565.8***		1088.4***		170.4*		816.5***		254.4
		(154.9)		(299.0)		(105.3)		(172.4)		(163.1)

SOURCE: Authors' calculations from the PSID.
NOTE: * p<0.10, ** p<0.05, *** p<0.01, shown for cross-differences only. All earnings numbers are in thousands of dollars, and are the present value of estimated lifetime earnings for a particular income background/educational attainment group. Logarithmic differences are 100 times the difference in natural logarithm of indicated variables. Standard errors (in parentheses) are robust to heteroskedasticity and intraperson correlation. Differences in the right column of each column pair always represent the higher-income group minus the lower-income group. See text for description of family income background.

Table 10 Alternate Moments and Restrictions for Men

	Baseline		Drop zero earnings		Drop graduate degrees		More extreme income grouping		No 99th pctile		Median	
	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	>400% pov.	100–200% pov.	Non-low-inc.	Low-income	Non-low-inc.	Low-income
PV of HS earnings	794.1	644.6	946.7	822.4	774.1	563.3	932.4	687.7	770.9	638.2	698.6	489.3
	(40.9)	(63.0)	(43.1)	(69.3)	(44.9)	(55.3)	(100.5)	(72.4)	(34.2)	(62.7)	(44.3)	(64.2)
PV of coll. earnings	2,143.1	905.2	2,357.3	1,068.9	1,563.0	910.6	2,573.4	793.3	1,516.9	864.7	1,398.4	887.1
	(260.1)	(126.9)	(283.7)	(131.3)	(98.7)	(150.6)	(395.0)	(165.3)	(53.6)	(118.0)	(52.1)	(239.2)
Difference	1,349.0	260.6	1,410.5	246.5	788.8	347.3	1,641.0	105.6	746.1	226.5	699.9	397.8
	(263.3)	(141.6)	(286.9)	(148.5)	(107.0)	(160.4)	(407.6)	(180.5)	(63.6)	(133.6)	(69.0)	(247.8)
Ratio	2.699	1.404	2.490	1.300	2.019	1.617	2.760	1.153	1.968	1.355	2.002	1.813
	(0.356)	(0.240)	(0.320)	(0.194)	(0.173)	(0.311)	(0.518)	(0.269)	(0.112)	(0.228)	(0.148)	(0.544)
Ratio difference		1.295***		1.190***		0.402		1.607***		0.613**		0.189
		(0.429)		(0.374		(0.356)		(0.584)		(0.254)		(0.564)
Ln college return	99.3	34.0	91.2	26.2	70.3	48.0	101.5	14.3	67.7	30.4	69.4	59.5
	(13.2)	(17.1)	(12.9)	(14.9)	(8.6)	(19.2)	(18.8)	(23.3)	(5.7)	(16.8)	(7.4)	(30.0)
Difference ln return		65.3***		65.0***		22.2		87.2***		37.3**		9.9
		(21.6)		(19.7)		(21.1)		(30.0)		(17.7)		(30.9)
Absolute difference college return		1088.4***		1164.0***		441.5**		1535.4***		519.6***		302.1
		(299.0)		(323.1)		(192.9)		(445.8)		(148.0)		(257.3)

SOURCE: Authors' calculations from the PSID.
NOTE: See previous tables.

Table 11 Alternate Moments and Restrictions for Whites

	Baseline		Drop zero earnings		Drop graduate degrees		More extreme income grouping		No 99th pctile		Median	
	Non-low-inc.	Low-income	Non-low-inc.	Low-income	Non-low-inc.	Low-income	>400% pov.	100–200% pov.	Non-low-inc.	Low-income	Non-low-inc.	Low-income
PV of HS earnings	673.5	555.4	802.5	702.0	672.4	462.4	798.6	564.6	658.3	553.9	553.0	441.1
	(28.9)	(52.8)	(30.5)	(60.7)	(35.5)	(48.3)	(66.9)	(54.7)	(25.1)	(52.3)	(29.7)	(57.1)
PV of college earn.	1,556.5	621.8	1,808.5	818.8	1,285.2	679.0	1,796.3	700.8	1,197.1	613.1	1,047.8	639.5
	(137.0)	(85.7)	(152.1)	(94.9)	(79.0)	(99.1)	(204.8)	(104.9)	(34.9)	(83.3)	(37.7)	(131.9)
Difference	882.9	66.4	1,006.0	116.8	612.7	216.6	997.7	136.2	538.9	59.1	494.8	198.4
	(140.0)	(100.6)	(154.6)	(110.9)	(85.4)	(109.0)	(215.4)	(118.3)	(43.0)	(98.3)	(48.0)	(142.5)
Ratio	2.311	1.120	2.187	1.106	1.834	1.395	2.249	1.241	1.819	1.107	1.895	1.450
	(0.226)	(0.187)	(0.207)	(0.166)	(0.152)	(0.259)	(0.318)	(0.221)	(0.087)	(0.183)	(0.122)	(0.349)
Ratio difference		1.191***		1.081***		0.439		1.008***		0.712***		0.445
		(0.293)		(0.265)		(0.300)		(0.387)		(0.203)		(0.370)
Ln college return	83.7	11.3	78.2	10.1	60.6	33.3	81.1	21.6	59.8	10.1	63.9	37.1
	(9.8)	(16.7)	(9.5)	(15.0)	(8.3)	(18.6)	(14.1)	(17.8)	(4.8)	(16.5)	(6.5)	(24.4)
Difference ln return		72.4***		68.1***		27.3		59.4***		49.7***		26.8
		(19.4)		(17.7)		(20.3)		(22.8)		(17.2)		(25.2)
Absolute difference college return		816.5***		889.2***		396.1***		861.5***		479.8***		296.4**
		(172.4)		(190.3)		(138.4)		(245.7)		(107.3)		(150.4)

SOURCE: Authors' calculations from the PSID.
NOTE: See previous tables.

Table A.1 Construction of PSID Analytic Sample

	Total		SRC sample		SEO sample	
Sample restrictions	Unique individuals	Person-year observations	Individuals	Observations	Individuals	Observations
Number of individuals in PSID core samples	61,261		31,829		29,432	
…whose parental income can be identified	9,144		4,775		4,369	
… and were interviewed	9,144	256,614	4,775	139,171	4,369	117,443
… and were at least age 20 in earnings year	9,144	127,495	4,775	68,965	4,369	58,530
… and were at least age 25 in earnings year	8,712	90,108	4,592	50,129	4,120	39,979
… and had nonimputed earnings (but incl. 0s)[a]	8,103	82,540	4,381	47,235	3,722	35,305
… and earnings did not exceed 99th pctile	8,103	81,928	4,381	46,708	3,722	35,220
… and have exactly 12 years of schooling by age 25	3,336	35,321	1,644	18,766	1,692	16,555
… have exactly 16 years of schooling by age 25	1,227	12,400	939	9,857	288	2,543
… have at least 16 years of schooling by age 25	1,559	14,550	1,196	11,627	363	2,923
… have exactly a high school diploma by age 25	2,814	31,628	1,360	16,441	1,454	15,187
… have exactly a bachelor's degree by age 25	1,352	12,822	1,040	10,069	312	2,753
… have at least a bachelor's degree by age 25	1,580	15,161	1,221	12,069	359	3,092

SOURCE: Authors' calculations from the PSID.

[a] Earnings are consistently available in the PSID only for household heads and their spouses; thus, the analytic sample is necessarily limited to adults not living with their parents.

Table A.2 PSID Sample Sizes by Education, Income Background, and Birth Year

Count of individuals

	SRC sample				SEO sample			
	High school grad.		Coll.		High school grad.		Coll.	
Birth year	Poor	Not poor	Poor	Not poor	Poor	Not poor	Poor	Not poor
1950	18	19	5	23	30	14	6	0
1951	14	29	5	26	49	7	4	0
1952	15	36	2	22	52	6	7	1
1953	20	40	3	27	52	7	7	2
1954	13	40	2	27	57	15	8	4
1955	10	41	2	23	55	18	6	8
1956	14	41	1	33	46	22	11	4
1957	13	30	3	15	48	18	8	5
1958	13	36	2	28	35	26	8	13
1959	9	41	1	23	42	25	7	7
1960	15	30	1	28	36	22	7	11
1961	10	33	0	28	39	19	2	6
1962	8	36	1	30	36	24	3	3
1963	8	29	0	16	26	22	1	3
1964	13	17	1	38	37	24	3	8
1965	6	26	1	26	29	18	4	8
1966	9	30	2	25	28	20	2	8
1967	8	12	2	26	27	22	2	3
1968	6	23	1	20	21	10	6	3
1969	7	21	0	16	14	7	2	4
1970	9	22	2	27	11	11	1	5
1971	6	22	2	29	9	6	2	3
1972	6	27	2	24	9	7	2	0
1973	11	15	3	25	8	4	1	2
1974	3	24	1	18	10	6	0	2
1975	8	21	1	28	12	2	1	2
1976	6	21	3	22	10	6	3	2
1977	9	18	3	35	11	9	2	3
1978	11	28	1	24	13	6	3	7
1979	10	24	0	38	17	8	2	3
1980	9	27	4	55	17	9	5	13
1981	9	22	5	43	13	8	3	9
1982	14	22	6	42	10	8	3	7
1983	4	16	0	54	10	6	2	7
1984	12	18	3	43	9	13	3	11
1985	14	13	2	58	11	8	4	12
1986	5	12	4	32	22	7	3	10
1987	7	15	1	46	13	8	5	11
Total	382	977	78	1,143	974	478	149	210

SOURCE: Authors' calculations from the PSID.

Table A.2 PSID Sample Sizes by Education, Income Background, and Birth Year, cont'd

Count of observations

	SRC sample				SEO sample			
	High school grad.		Coll.		High school grad.		Coll.	
Birth year	Poor	Not poor	Poor	Not poor	Poor	Not poor	Poor	Not poor
1950	379	384	105	612	536	225	149	0
1951	230	673	104	636	857	122	76	0
1952	345	861	54	548	843	112	142	19
1953	428	760	72	628	881	86	124	44
1954	255	828	44	565	947	223	147	64
1955	153	916	47	537	897	280	88	125
1956	246	838	21	631	657	333	216	60
1957	224	494	65	288	550	228	111	85
1958	229	570	41	529	452	360	98	183
1959	153	727	21	440	551	289	89	79
1960	209	476	2	484	466	240	98	118
1961	147	511	0	445	424	203	11	53
1962	130	570	18	488	451	237	35	26
1963	132	407	0	250	331	178	7	27
1964	164	211	16	530	332	192	35	56
1965	69	316	15	331	240	155	23	66
1966	104	332	28	305	200	152	26	46
1967	91	122	25	302	187	139	8	11
1968	53	181	10	202	170	62	59	22
1969	46	192	0	158	84	36	4	14
1970	74	194	19	243	69	37	10	14
1971	44	166	18	235	59	49	16	23
1972	40	177	14	180	57	40	14	0
1973	64	99	24	196	56	27	8	7
1974	20	132	7	116	52	31	0	10
1975	50	127	7	182	79	14	7	9
1976	30	110	17	114	56	31	14	12
1977	47	96	10	187	50	35	11	18
1978	50	131	5	110	53	20	12	29
1979	45	103	0	181	75	33	10	12
1980	35	94	16	209	60	31	16	47
1981	33	79	20	165	49	28	12	28
1982	37	51	18	112	26	19	7	19
1983	12	37	0	153	25	12	5	17
1984	22	34	6	83	15	25	6	20
1985	27	25	4	110	21	15	6	23
1986	5	12	4	32	22	7	3	10
1987	7	15	1	46	13	8	5	11
Total:	**4,429**	**12,051**	**878**	**11,563**	**10,893**	**4,314**	**1,708**	**1,407**

SOURCE: Authors' calculations from the PSID.

NOTE: SRC sample refers to the "Survey Research Center" or cross-sectional sample; SEO refers to the "Survey of Economic Opportunity" or low-income sample. "HSG" refers to individuals with exactly a high school diploma at age 25; "Coll" refers to those with a bachelor's degree or higher at age 25. "Poor" refers to individuals whose parental income was less than 185% of the federal poverty threshold when the individuals were aged 13–17; "Not poor" refers to those whose parental income was above that threshold.

Table A.3 Differential Attrition in the PSID

	(1)	(2)	(3)	(4)
Family below 185% FPL	0.0757***	0.0538***	0.0477***	0.0458***
	[0.0109]	[0.0104]	[0.0126]	[0.0123]
College graduate	−0.1053***	−0.0720***	−0.0848***	−0.0631***
	[0.0101]	[0.0096]	[0.0095]	[0.0093]
College grad × Family below 185% FPL	0.0479**	0.0219	0.0019	−0.0223
	[0.0229]	[0.0221]	[0.0241]	[0.0235]
Observations	87,286	87,286	87,286	87,286
R-squared	0.1134	0.2023	0.0834	0.1466
Include age dummies	No	Yes	No	Yes
Include weights	No	No	Yes	Yes
Mean of attrition	0.200	0.200	0.200	0.200

SOURCE: Authors' calculations from the PSID.
NOTE: The estimates represent the effect of the indicated variables from a linear probability model where the outcome equals 1 if the individual did not complete an interview in a survey wave and 0 if the individual did. The regression also includes survey year dummies (estimates not shown). The sample includes respondents who would be at least age 20 (not 25) in the calendar year prior to the survey year. Standard errors robust to heteroskedasticity and clustered on individual shown in brackets. *** p<0.01, ** p<0.05, * p<0.1.

Table A.4 Differential Selection into Head/Spouse in the PSID

	(1)	(2)	(3)	(4)
Family below 185% FPL	0.0056	−0.0082	0.0134*	0.0129*
	[0.0077]	[0.0072]	[0.0080]	[0.0073]
College graduate	−0.0612***	−0.0450***	−0.0642***	−0.0480***
	[0.0077]	[0.0068]	[0.0080]	[0.0071]
College grad × Family below 185% FPL	−0.0017	−0.0165	0.0115	−0.0139
	[0.0159]	[0.0137]	[0.0166]	[0.0136]
Observations	69,828	69,828	69,828	69,828
R-squared	0.0364	0.2398	0.0396	0.2708
Include age dummies	No	Yes	No	Yes
Include weights	No	No	Yes	Yes
Mean of head/spouse	0.873	0.873	0.873	0.873

SOURCE: Authors' calculations from the PSID.
NOTE: *** p<0.01, ** p<0.05, * p<0.1. The estimates represent the effect of the indicated variables from a linear probability model where the outcome equals 1 if the individual was listed as a head or spouse in a given survey year and 0 if the individual was not, conditional on the individual being successfully interviewed. The regression also includes survey year dummies (estimates not shown). The sample includes respondents who would be at least age 20 (not 25) in the calendar year prior to the survey year. Standard errors robust to heteroskedasticity and clustered on individual shown in brackets.