![CAPSEE — Center for Analysis of Postsecondary Education and Employment]

Estimating Returns to College Attainment: Comparing Survey and State Administrative Data Based Estimates

A CAPSEE Working Paper

Judith Scott-Clayton
Community College Research Center
Teachers College, Columbia University
and NBER

Qiao Wen
Community College Research Center
Teachers College, Columbia University

January 2017

Abstract

The increasing availability of massive administrative datasets linking postsecondary enrollees with post-college earnings records has stimulated a wealth of new research on the returns to college, and has accelerated state and federal efforts to hold institutions accountable for students' labor market outcomes. Many of these new research and policy efforts rely on state databases limited to postsecondary enrollees who work in the same state post-college, with limited information regarding family background and pre-college ability. In this paper, we use recent waves of data from the National Longitudinal Survey of Youth 1997 (NLSY97) to provide new, nationally representative, non-experimental estimates of the returns to degrees, as well as to assess the possible limitations of single-state, administrative-data-based estimates. To do this we evaluate how the national estimates change, depending upon the baseline comparison group, adjustments for pre-college achievement and family background, and alternative ways of addressing interstate mobility. We conclude with a discussion of the relative advantages and disadvantages of survey versus administrative data for estimating returns to college, as well as implications for research and policy efforts based upon single-state administrative databases.

Table of Contents

1. Introduction

The increasing availability of massive administrative datasets linking postsecondary enrollees with post-college earnings records has stimulated a wealth of new research on the returns to college, and has accelerated state and federal efforts to hold institutions accountable for students' labor market outcomes. Many of these new research and policy efforts rely on state databases limited to postsecondary enrollees who work in the same state post-college, with limited information regarding family background and pre-college ability. Limited controls may tend to upwardly bias the estimated returns to degree completion, while sample limitations may lead to downwardly biased estimates. We are not aware of prior studies that have attempted to assess how these biases net out in practice.

To address this question, we use recent waves of data from the National Longitudinal Survey of Youth 1997 (NLSY97) to provide new, nationally representative, non-experimental estimates of the early-career returns to degrees, when respondents are 25 to 30 years old. In addition to examining differences in average earnings, we also look at employment stability, likelihood of earning at least a "living wage," and propensity to move out of state.[1] We then turn to an assessment of the limitations of single-state, administrative-data-based estimates. To do this we evaluate how our national estimates change, depending upon the baseline comparison group, adjustments for pre-college achievement and family background, and alternative ways of addressing interstate mobility. While our goal is to estimate causal effects, these estimates should be interpreted cautiously because they are not based on any experimental or quasi-experimental variation in college enrollment or attainment. We control for a very rich set of observable pre-college characteristics (including test scores); nonetheless, unobservable biases in who enrolls and completes college may remain.

Our findings add to a large body of evidence documenting substantial returns to associate and bachelor's degrees relative to high school completion alone. People who have a bachelor's or associate degree have annual earnings nearly $21,000 and $8,000 higher than high school graduates respectively. Interestingly, we also find significant returns to enrollment in four-year college without obtaining a degree. Earnings for those with only some two-year college enrollment, or only a certificate but no degree, are not significantly different from the earnings of high school completers. Beyond average earnings, we also estimate positive effects of four- and two-year attainment on the stability of employment and likelihood of earning at least a "living wage."

Our study also explores how college enrollment and attainment relate to interstate mobility. We find that bachelor's degree holders (and those with some four-year college but no degree) are more likely to live outside their home state after college. Bachelor's degree holders (and four-year college non-completers) who live out of state also earn more than their counterparts who remain in their home state, suggesting that part of the payoff from four-year

---

[1] We define this as annual earnings equivalent to $15/hour, for 35 hours/week, 50 weeks/year.

college comes from interstate migration. In contrast, other levels of attainment appear to have no significant correlation with interstate mobility (if anything, associate degree holders who leave their home state appear to earn less than their counterparts who remain).

The final contribution of our study is to explore the sensitivity of the estimates of returns to college, depending upon sample restrictions, inclusion of different sets of covariates, and alternative ways of treating out-of-state earnings. While the broad pattern of findings is generally quite robust to alternative specifications, the magnitude of the point estimates can vary substantially. For example, estimates from our richest model, controlling for high school GPA and test scores, reduces estimated returns to bachelor's degrees by 13 percent, associate degrees by 32 percent, and certificates by 67 percent. On the other hand, estimates of total returns are biased downward when samples are restricted to college enrollees only (versus using high school graduates as the comparison group). On balance, these two countervailing influences roughly balance out.

The effect of out-of-state mobility is harder to fix. Mobility does not appear to be a major obstacle to accurately estimating returns to associate degrees and other sub-baccalaureate credentials. But simply treating all out-of-state earnings as zeros, as some state-level analyses do, substantially underestimates the returns to bachelor's degrees. A common fix is to estimate returns conditional on in-state employment. We find that this provides a reasonable estimate of the returns conditioned on any employment—but because bachelor's degrees also positively influence employment rates, this still understates the returns to a bachelor's degree.

The remainder of the paper proceeds as follows. In Section 2, we place this analysis in context of the existing literature. Section 3 describes our methodology, and Section 4 presents our main results. Section 5 concludes with a discussion of the relative advantages and disadvantages of survey versus administrative data for estimating returns to college, as well as implications for research and policy efforts based upon single-state administrative databases.

2. Prior Literature

The prior literature examining the returns to schooling is extensive and long-standing, and a full review is beyond the scope of this paper (for detailed reviews, see Card, 1999; Oreopoulos & Petronijevic, 2013). Here we provide a brief overview of the recent literature most relevant to the present analysis (see Appendix E for a detailed listing of relevant studies).

Among studies using national longitudinal survey data with controls for student characteristics and ability, the average earnings premium for associate degrees (relative to high school credentials only) averages approximately 13 and 22 percent for men and women respectively, but with wide variation across studies (Belfield & Bailey, 2011). Very few studies have used national datasets to examine returns to certificates (in part due to the relatively low prevalence of such degrees until recently). But two studies that have examined returns to

certificates have found a premium, relative to high school completion, of 7–8 percent for men and 20–24 percent for women (Grubb, 1997; Marcotte, Bailey, Borkoski, & Kienzl, 2005).

Recently, more studies have utilized state administrative databases that include academic records for students who enrolled in the state's community college system, merged with quarterly wage data from state Unemployment Insurance (UI) systems. These studies typically start with a sample of college enrollees, so returns to degrees are estimated relative to those who have at least some college rather than relative to those who completed high school. Although they typically have limited background variables to include as controls, these studies often use an individual fixed-effect approach to estimating returns, using pre-college earnings and sometimes also enrollment and course-taking patterns to account for across-student variations. These more recent studies find consistent results in terms of the substantial earnings increase for obtaining an associate degree, ranging from 24 and 56 percent in Kentucky (Jepsen, Troske, & Coomes, 2014), 16 and 32 percent in North Carolina (Liu, Belfield, & Trimble, 2015) and 4 and 9 percent in Washington (Dadgar & Trimble, 2015) for men and women respectively.

These same studies have generally found positive returns for long-term certificates (requiring at least one year of study): compared to community college non-completers, the earnings premium for obtaining these credentials range from 2–21 and 15–45 percent for men and women respectively (Dadgar & Trimble, 2015; Jepsen et al., 2014; Liu et al., 2015). Results for short-term certificates are mixed: modest earnings increases (3–7 percent) relative to community college non-completers are found in North Carolina and Kentucky (Jepsen et al., 2014; Liu et al., 2015; Xu & Trimble, 2015), but other studies conclude that earnings for short-certificate holders are statistically indifferent from that of community college non-completers (Bahr et al., 2015; Dadgar & Trimble, 2015; Zeidenberg, Scott, & Belfield, 2015).

While the availability of state administrative databases has led to a wealth of new estimates of returns to degrees, such databases also have potentially important limitations. First, state databases typically contain wage records only for those people who work in the same state after college graduation. As a result, in such databases people who work out of the state appear indistinguishable from those who remain in the state but are not working.

Second, such databases seldom contain information on students' family background, previous academic performance, or test scores; instead they typically rely on pre-enrollment labor market history to predict students' earnings in the absence of a degree. As a result, omitted variable bias can be a serious threat to validity, and younger students with limited or no pre-college work experience are often excluded from these analyses. Since prior evidence suggests that college attainment positively affects mobility (Bound & Holzer, 2000; Malamud & Wozniak, 2012; Wozniak, 2010), this could lead to an underestimate of returns if some high-earning graduates migrate out of state.

The last major limitation of studies using such databases is that their samples include only students who enrolled for at least one credit in community colleges. Thus, the earnings increases associated with obtaining community college awards are all estimated relative to those

who enrolled in community college but obtained no degree. These comparisons may underestimate the returns to college degrees and attendance relative to those with high school credentials only.

To sum up, the key takeaway of the recent literature on returns to college credentials is that there are substantial returns associated with obtaining bachelor's and associate degrees and in some cases, also with obtaining long-term certificates. But, the range of the estimated earnings premiums is large and highly sensitive to restrictions regarding databases, samples, covariates and measurements of earnings. In this paper, by using nationally representative survey data to simulate the limitations of state administrative databases, we are able to examine how sensitive such estimates are to sample limitations, treatment of out-of-state earnings, and alternative controls.

3. Methodology

Data and Sample

This study uses data from the National Longitudinal Survey of Youth 1997 (NLSY97). The NLSY97 is a longitudinal project that follows a sample of youth in the United States who were born in 1980–1984 and were 13–17 years old at the time when the first round of the survey was conducted in 1997. The initial NLSY97 cohort consists of 8,984 young people, including a nationally representative sample of 6,748 youths and an oversample of 2,236 Black and Hispanic youths.[2] This cohort has been surveyed for 16 times to date, of which the first 15 surveys were conducted annually since 1997 and were replaced with biennial surveys starting in 2013 (round 16).

A major challenge in any survey-based dataset is sample attrition, either from survey or item non-response. Survey retention rates are high for the NLSY: for round 15 (which is the most recent round we use), retention is almost 82 percent, and 58 percent respond in every single round. Even for respondents who do not participate in every follow-up round, however, most respond in at least some: 99 percent responded in at least one follow-up round.

This study uses the first 15 rounds of the survey, which covers the years of 1997 through 2011.[3] NLSY97 contains rich data on education and labor market outcomes. It also contains very detailed information on respondents' family background, test scores, and academic performance at each schooling level. These data are publicly available through the Bureau of Labor Statistics

[2] The NLSY97 provides sampling weights for each survey round, and we use the baseline survey sampling weights in both our descriptive and regression analysis.
[3] We do not use the most recent round (i.e., round 16) because the structure of some key survey questions is different from the previous 15 rounds (for example, instead of asking for educational credential receipt information and employment status in the previous year, the round 16 asks whether the respondent has obtained any new credentials and has had any type of jobs in the past two years), which makes it difficult to compose a clear-cut cross-sectional sample for our analysis.

(BLS). In addition, our study also uses the NLSY97's restricted geocode data to study the relationship between educational attainment and interstate mobility. NLSY97 geocode data document geographic information for each respondent, including the state, county, and metropolitan area of residence for each respondent in each survey year, as well as migration history across two rounds. Except as otherwise noted, our study focuses on respondents who have obtained at least a high school diploma or GED as of 2009 and who were not enrolled in 2010 and afterwards.

Estimation Strategy

Our estimation strategy is based on the human capital earnings function proposed by Mincer (1974). Specifically, we use ordinary least squares (OLS) regressions to estimate the relationship between educational attainment and annual earnings, conditional on other personal characteristics that might affect both educational attainment and earnings:

$$(1) \quad Y_i = \alpha + \boldsymbol{\beta Educ_i} + \lambda Exp_i + \varphi Exp_i^2 + \boldsymbol{\theta X_i} + \boldsymbol{\pi Geo_i} + \boldsymbol{\rho SES_i} + \boldsymbol{\sigma Abil_i} + \varepsilon_i$$

The outcome variable in equation (1), Y_i, indicates an earnings or other post-college outcome for individual i. Our primary specification uses earnings (or log earnings) from all jobs in 2010 as the dependent variable.[4] $Educ_i$ is a set of indicators for having attended two-year college but obtaining no credential, attended four-year college but obtaining no credential, obtained an undergraduate certificate, obtained an associate degree, and obtained a bachelor's degree and above, respectively. Unless stated otherwise, the reference group for equation (1) is respondents who have obtained a high school diploma or GED but have not attended any postsecondary educational institutions.[5] X_i is a vector of demographic characteristics, such as gender, race/ethnicity, and age; Geo_i is a vector of geographic characteristics, such as region of residence at age 17 and a series of urban/rural/suburban and MSA dummies for the residence; SES_i is a vector of family background characteristics, which include parental educational levels, household size, and household net worth at age 17, and $Abil_i$ is a vector of proxies for respondents' ability, including type of high school attended, high school GPA, and ASVAB test score.

Apart from earnings, we also use equation (1) to look at other labor market outcomes, such as the probability of being employed and having none-zero earnings, hours worked during a calendar year, and the probability of having earnings above thresholds calculated using both federal minimum hourly wage and "living wage" rates. Moreover, our study further delves into the relationship between educational attainment and interstate mobility. We examine how educational attainment relates to interstate mobility, measured by the probability of residing in a

---

[4] Job earnings include wages, salary, commissions, or tips, and jobs include either self-employed jobs or employee-type jobs. We report job earnings in 2015 dollars, but also perform regressions using the natural logarithm form.
[5] We also perform an alternative specification in which only college enrollees are included. In this specification, the reference group is those who have attended college but obtained no postsecondary credentials.

different state from the respondents' home state in 2010, as well as the probability of ever having moved across states between age 18 and the 2010 survey round.

The primary threat to validity of causal inference in the OLS approach is the potential endogeneity in schooling choices. If schooling choices are endogenous—for example, if some individuals have both higher potential earnings and higher educational attainment for reasons that cannot be explained by observable characteristics—then the estimated returns to educational attainment will be biased upwards. Common omitted variables include ability and family background, which often bias the estimates of education upwards. Researchers have long been trying to overcome these limitations and use techniques such as instrumental variables, difference-in-differences, and propensity score matching to make causal claims about the relationship between education and earning. However, results from previous studies suggest that a simple cross-sectional regression of earnings on education can well reflect the average return to education in a given population (Card, 1999; Griliches, 1977). The best available evidence from the latest studies of identical twins shows a small upward bias on the order of 10 percent in the simple OLS estimates. Moreover, our model incorporates very detailed control variables including family net worth, parental educational attainment, high school GPA, and aptitude test scores, not all of which were utilized in early studies. This could further reduce the biases due to omitted variables.

A second concern with OLS estimates is measurement error: misreporting of years of schooling (or degree levels or majors) will bias the estimates of returns toward zero.[6] We minimize this issue by using postsecondary transcripts data to cross-validate the self-reported educational information, which can reduce misreporting behaviors (more detailed discussions on this topic can be found in Appendix D). We also use highest degree obtained, instead of years of schooling, which is less prone to measurement error (Black, Sanders, & Taylor, 2003).

One of the fundamental goals of our analysis is to explore the sensitivity of OLS estimates of the returns to degrees, depending upon sample restrictions and availability of data on family background, pre-college achievement, and test scores. To the extent that our findings are robust across a range of specifications and sets of controls, this helps assuage concerns that the observed patterns are purely spurious.

Key Variables

Table 1 provides summary statistics on the key variables (controls, predictors, and outcomes) used in our analysis, both for the full NLSY97 sample and for our analysis sample, which is limited to respondents who obtained at least a high school diploma or GED as of 2009 and who were not enrolled in school in 2010 and beyond. We focus on only key variables in Table 1. Additional information on other variables and extra notes on data cleaning are provided in Appendices A and B.

---

[6] Griliches (1977) argues that measurement error in schooling partially offsets the upward bias from omitted variables.

Table 1: Summary Statistic for the Full and Analysis Samples

Variable	Full Sample Mean	N	Analysis Sample (HS completers, not currently enrolled) Mean	N
Control Variables:				
Female	0.487	8984	0.485	5760
White non-Hispanic	0.666	8984	0.673	5760
Black non-Hispanic	0.154	8984	0.154	5760
Hispanic	0.129	8984	0.123	5760
Other or mixed Race	0.051	8984	0.050	5760
Age (in 2010)	28.0	8984	28.0	5760
Urban area	0.692	8984	0.688	5760
Parent's highest education:				
No credential	0.151	8984	0.142	5760
HS diploma/GED	0.480	8984	0.513	5760
Associate degree	0.092	8984	0.103	5760
Bachelor's degree	0.120	8984	0.134	5760
Advanced degree	0.052	8984	0.055	5760
Missing	0.104	8984	0.052	5760
Family background in 1997:				
Household net worth (in 1997 $)	$108,806	6713	$110,973	4358
Household size	4.5	8984	4.4	5760
Attended public high school	0.902	8984	0.903	5760
Overall HS GPA	2.87	6004	2.89	4167
ASVAB score	50.4	7093	51.8	4697
Predictor Variables:				
Work experience (in years since age 18)	7.0	8984	7.9	5760
Ever Enrolled in College	0.662	8966	0.713	5760
Educational attainment:				
No credential	0.108	8966	-	5760
HS diploma/GED	0.230	8966	0.287	5760
Some 2-yr coll., no degree	0.159	8966	0.184	5760
Some 4-yr coll., no degree	0.128	8966	0.113	5760
Some coll. (type unknown), no degree	0.012	8966	0.007	5760
Undergraduate certificate	0.020	8966	0.027	5760
Associate degree	0.064	8966	0.066	5760
Bachelor's degree and above	0.279	8966	0.316	5760
Outcomes				
Still enrolled in school	0.283	8966	-	5760
Income from all jobs in 2010 (if not in school, 2015 dollars)	$31,589	6429	$33,444	5735
Have non-zero job income in 2010 (if not in school)	0.814	6429	0.836	5735
Number of hours worked in 2010 (if not in school)	1658	6258	1716	5588
Worked FTYR in 2010 (if not in school)*	0.579	6258	0.604	5588
Income above min. level in 2010 (if not in school)*	0.685	6429	0.716	5735
Ever moved across states since 18	0.307	6448	0.320	5753
State of residence in 2010 different from home state	0.200	5990	0.208	5371

Note. Source is NLSY97 restricted data, rounds 1–15 (1997–2011). The full sample includes all NLSY97 respondents, and the analysis sample is limited to respondents with at least a high school diploma or GED as of 2009 and no enrollment in school in 2010. Sampling weight in round 1 is used for all analyses.

*"Worked FTYR in 2010" refers to the probability of working "full-time and year-round" in 2010, which is defined as working 1,750 or more hours (i.e., 35 hours/week * 50 weeks). "Income above min. level in 2010" refers to the probability of having annual job earnings in 2010 above the federal minimum level, which is calculated as $12,687.50 in 2010 dollar (i.e., $7.25/hour * 1,750 hours).

Control variables. Our model specifications include four sets of controls. Basic controls include gender, age, and race;[7] the second set of controls further includes the level and squared forms of working experience; the third set of controls further includes geographic and family background controls. Geographic information captures individuals' residence in the first round when they were 13–17 years old.[8] Family background controls include parental educational attainment, household net worth in 1997, and household size; the last set of controls further include pre-college performance, including type of high school attended, overall high school GPA and test score on the Armed Services Vocational Aptitude Battery (ASVAB).[9]

In 2010, the average age of respondents in our analysis sample was 28. Average work experience, measured as cumulative years worked since age 18 through 2009, was 7.9 years.[10] Only about 1 in 5 respondents in our analysis sample had a parent with a bachelor's degree or higher; about 1 in 3 had a parent with any college degree.

Educational attainment. We determine respondents' educational attainment using both the self-reported survey data and the postsecondary transcript study, which was administered by the NLSY97 research team during 2010–2011. A detailed description of this study and crosschecking results regarding postsecondary enrollment and attainment between the self-reports and transcripts data are presented in Appendices C and D. We classify educational attainment into six categories: high school diploma/GED but no postsecondary enrollment, two-year college enrollment but no degree, four-year college enrollment but no degree, undergraduate certificate, associate degree, and bachelor's degree and above.[11]

Approximately 89 percent of the total NLSY97 sample have at least obtained a high school diploma or GED, 66 percent have at least some college enrollment, 34 percent have at least an associate degree, and 28 percent have at least a bachelor's degree. The estimated high school graduation rates and college entry rates are very similar to rates estimated in the Current Population Survey for 25–29 year olds. Degree completion rates, however, are slightly lower than CPS estimates (e.g., CPS suggests 32 percent of 25–29 year olds had a bachelor's degree in 2010).[12] It is possible that some of this discrepancy reflects over-reporting in the CPS data.

---

[7] We classify respondents' race/ethnicity as "White non-Hispanic," "Black non-Hispanic," "Hispanic," and "Other Races/ethnicities."

[8] It includes regions of residence (i.e., Northeast, South, West, and North Central) and characteristics of the residence (i.e., urban/rural/unknown area and metropolitan statistical area [MSA]/non-MSA area).

[9] High school type includes public, private/parochial, and other types. Overall high school GPA indicates grade point averages across all high school courses on a 5-point grading scale and is weighted by course credits. The ASVAB score has been adjusted to indicate respondents' percentile ranking on the four math and verbal subtests that take into consideration the respondents' age.

[10] The original data are in hours; they are converted to years by dividing by 1,750 (i.e., 50 weeks/year * 35 hours/week).

[11] There are 459 respondents (approximately 5 percent) who have obtained a graduate degree (including master's degrees, Ph.D.s, and professional degrees). Due to this small size, we combine them with respondents who have bachelor's degrees and create the "bachelor's degree and above" group.

[12] For CPS based tabulations, see Digest of Education Statistics (2010): https://nces.ed.gov/programs/digest/d10/tables/dt10_008.asp

Earnings and employment. Our analysis uses total earnings from all jobs in 2010, which includes both self-employed jobs and employee-type jobs. Job earnings include wages, salary, commissions, and tips.[13] Individuals who refused to report the amount of job income were asked to choose from several income ranges. Individuals who reported categorical earnings are assigned values equal to the mid-point of the range reported. As mentioned before, individuals who were still enrolled in school in 2010 and after are excluded from our analysis. The mean earnings in 2010 for people no longer enrolled in school for our overall sample and analysis sample are $31,589 and $33,444, both inflated to 2015 dollars. For our analysis, we use both unlogged and logged earnings to estimate the returns to degrees and majors; we also both include and exclude zero earnings when using unlogged earnings in order to capture the potential biases caused by variation in the probability of employment by degree level and major.

Apart from earnings, we also look at other labor market outcomes in order to capture variations in employment and labor supply characteristics by degree level and major. These alternative employment outcomes include the probability of having positive job earnings and the probability of working full-time and year-round. Among those no longer enrolled in school, 81 percent of the overall sample had positive job incomes in 2010, and 58 percent worked full-time, year-round. To examine impacts on earnings in different parts of the earnings distribution, we also examine the probability of earning above the full-time, full-year, minimum-wage level (equivalent to $12,688 in 2010, which is about 17 percent above the poverty line for a single individual) as well as the probability of earning a "living wage" (equivalent to full-time, full-year employment at $15/hour, or about $26,250 annually in 2015 dollars).

Interstate mobility. We derive migration information from the restricted-use geocode data, and construct two variables to capture respondents' interstate migration history. We create two dummies to indicate whether the respondents have ever moved across states between two survey rounds since he or she turned 18 years old, and whether each respondent's state of residence in 2010 is different from his or her state of residence at 17 years old.[14] Thirty-one percent of the overall sample have moved across states at least once since they were 18 years old, and 20 percent lived in a different state from their home state in 2010.

---

[13] For respondents who were not interviewed in 2011 and thus do not have income information for year 2010, we use job earnings in 2009 (inflated to 2010 dollars) if the individuals were not enrolled in school since 2009. This affects 4.44 percent (approximately 300 individuals) of the total analysis sample. We create a variable to indicate which income year is used for each individual.

[14] The "ever moved across states since 18 years old" dummy equals to 1 if individuals have changed their state of residence at least once starting from when they were 18 years old to the year 2011; this dummy equals to 0 if "state of residence" is the same for every year when the respondent was 18 years old and older, or "state of residence" is the same for all available years since 18 years old and "state of residence" information for the remaining years is missing. This dummy equals to *missing* if information for "state of residence" is missing for every single year or it is available for only one year since the respondent was 18 years old.

4. Results

Returns to College Enrollment and Completion

Table 2 provides our basic results regarding the returns to college enrollment and completion, relative to a high school credential only. To avoid confounding the extensive and intensive employment margins, we focus on returns conditional on positive earnings (and this excludes roughly 16 percent of our sample); we examine effects on the extensive employment margin later. The top panel runs regressions with level earnings as the outcome, while the bottom panel shows log earnings.[15] Model 1 includes only the educational attainment indicators, along with controls for age, race/ethnicity, and gender. Model 2 adds experience and experience squared measured as total years worked since 18 years old through 2009, Model 3 adds geographic indicators and controls for family socioeconomic characteristics (measured at baseline). Model 4 adds high school GPA and ASVAB scores (along with relevant missing data flags).

Consistent with prior literature, Table 2 indicates large positive returns to associate and bachelor's degrees, regardless of specification. Interestingly, those with some college but no degree also receive a substantial payoff, but only if their enrollment was at a four-year college. Notice that there is significant difference in credits completed between those enrolled in two-year and four-year colleges. Those with some four-year enrollment have completed, on average, more than two years of college, while those with only some two-year enrollment have, on average, slightly less than one year of college.[16] But since the magnitude of the returns to some four-year college is more than double that for some two-year college, this still suggests that enrollment in four-year college has a higher per-credit payoff than enrollment in two-year college. Coefficients for other levels of attainment are generally positive, but smaller and statistically insignificant.

By comparing the coefficients in Model 3 and Model 4, we can examine the role of ability bias in a way that is not always feasible in administrative datasets. Adding controls for high school GPA and ASVAB scores reduces the estimated return to a bachelor's degree only modestly, by about 15 percent. Ability controls appear to matter more for lower levels of attainment: estimated returns to associate degrees shrink by 25 percent and returns to "some four-year college" shrink by over 30 percent. (Returns to certificates and "some two-year college" are also substantially reduced, although the original estimates were not statistically significant in any case.)

---

[15] Note that coefficients for *ln(earnings)* specifications approximate percentage changes only when log coefficients are less than about 0.20 in absolute value. For larger magnitudes, the percentage change will be larger than the coefficient; the percentage change can be obtained by the formula, $e^B - 1$.

[16] These estimates are based on variables that indicate the fraction of the corresponding degree program that had been completed. Those with some four-year college on average had completed 54 percent of their (four-year) degree, while those with some two-year college had on average completed 46 percent of their (two-year) degree.

Table 2: Returns to College Versus High School Only (Job Earnings Conditional on Employment)

Variable	Model 1 Coef.	(S.E.)		Model 2 Coef.	(S.E.)		Model 3 Coef.	(S.E.)		Model 4 Coef.	(S.E.)	
Earnings in 2010 (inflated to 2015 $)				(HS/GED only mean = $31,419)								
Some 2-yr coll. (only)	$1,060	(1278)		$839	(1227)		-$254	(1228)		-$1,122	(1228)	
Some 4-yr coll. (only)	$7,354	(1468)	***	$8,785	(1410)	***	$7,271	(1423)	***	$4,981	(1452)	***
Certificate	$1,911	(2688)		$2,345	(2580)		$1,444	(2564)		$624	(2553)	
Associate degree	$9,944	(1758)	***	$10,360	(1688)	***	$8,953	(1690)	***	$6,794	(1708)	***
Bachelor's degree	$20,620	(1084)	***	$24,355	(1053)	***	$21,215	(1130)	***	$17,942	(1217)	***
Ln(earnings)												
Some 2-yr coll. (only)	0.11	(0.04)	***	0.09	(0.04)	**	0.07	(0.04)	*	0.05	(0.04)	
Some 4-yr coll. (only)	0.25	(0.05)	***	0.30	(0.04)	***	0.27	(0.05)	***	0.21	(0.05)	***
Certificate	0.03	(0.09)		0.04	(0.08)		0.02	(0.08)		0.00	(0.08)	
Associate's degree	0.35	(0.06)	***	0.34	(0.05)	***	0.31	(0.05)	***	0.25	(0.05)	***
Bachelor's degree	0.63	(0.03)	***	0.75	(0.03)	***	0.67	(0.04)	***	0.59	(0.04)	***
Model controls:												
Incl. age, race, gender	X			X			X			X		
Incl. exp. and exp. squared				X			X			X		
Incl. geographic and family background							X			X		
Incl. HS GPA and ASVAB										X		
N	4699			4699			4699			4699		

Note. Source is authors' analysis using NLSY97 restricted data, rounds 1–15. Analysis sample limited to respondents with at least a high school diploma or GED as of 2009 and no enrollment in school in 2010. Job sampling weight in round 1 is used for all analyses. Regressions limited to 83.6 percent of respondents with positive job earnings in 2010. Job earnings refer to self-reported income from all jobs held in 2010, including self-employment, wages, salary, commissions, and tips. Omitted attainment category is high school diploma/GED only. *R*-squared range from 0.12–0.14 in Model 1 and in Model 4. We do not report estimates for the small number (*n* < 50) of people who enrolled in some unknown type of college without obtaining a degree.

*** *p* < .01, ** *p* < .05, * *p* < .1

Other Labor Market Outcomes

Our main models, which examine average earnings conditional upon at least some employment, may miss important margins of impact. Table 3 presents several alternative labor market outcomes of interest. The first column shows effects on unconditional job earnings; that is, including those with zero job earnings in the regression. Similar to results shown in column 4 of Table 2, associate and bachelor's degree holders, as well as those who enrolled in some four-year college without obtaining a degree, earn substantially more, relative to those with high school credentials only. However, using conditional earnings, instead of unconditional earnings, slightly underestimates the returns to four-year college enrollment and substantially underestimates returns to both associate and bachelor's degrees by about 14 percent.

Table 3: Other Employment-Related Outcomes (College Versus High School Only)

Variable	Unconditional job earnings in 2015 dollar Coef.	(S.E.)		Probability of having positive job earnings Coef.	(S.E.)		Probability of working FTYR Coef.	(S.E.)		Probability of job earnings>min. level Coef.	(S.E.)		Probability of job earnings>living wage Coef.	(S.E.)	
Some 2-yr coll. (only)	-$499	(1060)		0.023	(0.013)	*	-0.001	(0.017)		0.041	(0.016)	***	0.027	(0.017)	
Some 4-yr coll. (only)	$5,037	(1278)	***	0.061	(0.016)	***	0.030	(0.021)		0.081	(0.019)	***	0.074	(0.021)	***
Certificate	$575	(2225)		0.036	(0.028)		-0.075	(0.036)	**	0.017	(0.033)		0.019	(0.036)	
Associate degree	$7,917	(1552)	***	0.089	(0.020)	***	0.057	(0.025)	**	0.146	(0.023)	***	0.152	(0.025)	***
Bachelor's degree	$20,939	(1095)	***	0.171	(0.014)	***	0.271	(0.018)	***	0.268	(0.016)	***	0.334	(0.018)	***
HS/GED (only) mean	$23,199			0.738			0.507			0.578			0.349		
N	5735			5735			5588			5735			5735		
r2	0.302			0.191			0.261			0.253			0.266		

Note. Source is authors' analysis using NLSY97 restricted data, rounds 1–15. Analysis sample limited to respondents with at least a high school diploma or GED as of 2009 and no enrollment in school in 2010. Sampling weight in round 1 is used for all analyses. Job earnings refer to self-reported income from all jobs held in 2010, including self-employment, wages, salary, commissions, and tips. The third column presents the probability of working "full-time and year-round" in 2010, which is defined as working 1,750 or more hours in 2010 (i.e., 35 hours/week * 50 weeks). The fourth column presents the probability of having annual job earnings in 2010 above the federal minimum level, which is calculated as $12,687.50 in 2010 dollars (i.e., $7.25/hour * 1,750 hours). The last column presents the probability of having annual job earnings in 2010 (inflated to 2015 dollars) above the "living wage" level, which equals $26,250 in 2015 (i.e., $15/hour * 1,750 hours). Controls in the above include: gender, race/ethnicity, work experience, work experience squared, geographic and family background variables, high school GPA and ASVAB test score. Omitted attainment category is high school diploma/GED only. *R*-squared range from 0.12–0.14 in Model 1 and in Model 4. We do not report estimates for the small number ($n < 50$) of people who enrolled in some unknown type of college without obtaining a degree.

*** $p < .01$, ** $p < .05$, * $p < .1$

The next two columns show effects on the probability on having any job income during the year, and the probability of working full-time, year round (based on reported total hours worked). Results show that associate and bachelor's degree holders are both more likely to have any job and work full-time, year-round during a year. Moreover, enrollments in both two-year and four-year colleges even without obtaining a degree are associated with higher probabilities of having any job income during a year. However, people with some college exposure but no degree are not more likely to work full-time and year-round, relative to high school graduates only. Interestingly for certificate holders, they are *less* likely to work full-time and year-round by 7.5 percent, relative to high school graduates only—possibly indicating that certificates are disproportionately earned by individuals already at a disadvantage in the labor market (as it is difficult to take the negative effect as causal at face value).

The last two columns show the probabilities of earning at least as much as a full-time, full-year worker would earn at the federal minimum wage ($7.25/hour, or $12,678 in 2010) and at the "living wage" rate of $15 per hour (approximately $26,250 annually). Results show that bachelor's and associate degree holders, as well as those with some college enrollment, are all more likely to earn above the minimum wage threshold. However, the probabilities of earning above the "living wage" are significantly higher only for bachelor's degree and associate degree holders, as well as people who enroll in four-year colleges. The minimum wage outcome captures shifts at the low end of the earnings distribution, which may be of particular value to policymakers (note that this earnings threshold is only slightly above the federal poverty level for a single individual and only slightly below the maximum level to qualify for the EITC as a single individual with no dependents).

The results across these four additional outcomes generally reinforce the findings from Table 2. Associate and bachelor's degree holders, as well as those with some four-year college experience, not only earn more conditional on working, but are substantially more likely to be employed. The magnitude of these coefficients (6 percentage points for those with some four-year college experience, 9 percentage points for associate degree holders and 17 percentage points for bachelor's degree holders) is particularly notable given that employment rates for high school graduates are already quite high, at 73.8 percent. In contrast to Table 2, though, in Table 3 we also see significant positive benefits to some two-year college experience (even without a degree), in terms of the probability of having any income and the probability of earning above the full-time, full-year minimum wage level.

Interstate Mobility

Table 4 explores how college enrollment and attainment relate to interstate mobility. The first column shows the probability of ever moving across states between rounds (excluding moves that occur at age 17 or younger), with full sets of controls. The second column shows the probability of living in a different state in 2010 than at age 17.

Table 4: Relation Between Education and Interstate Mobility (College Versus High School Only)

Variable	Probability of ever moved across states at or after 18 years old			Probability of residing in a different state in 2010 from state of residence at 17 years old		
	Coef.	(S.E.)		Coef.	(S.E.)	
Some coll. 2-yr	0.045	(0.018)	**	0.013	(0.017)	
Some coll. 4-yr	0.099	(0.022)	***	0.091	(0.020)	***
Certificate (main effect)	0.113	(0.039)	***	-0.007	(0.035)	
AA (main effect)	0.073	(0.027)	***	0.022	(0.024)	
BA (main effect)	0.140	(0.019)	***	0.130	(0.017)	***
HS/GED (only) mean	0.212			0.128		
N	5753			5371		

Note. Source is authors' analysis using NLSY97 restricted data, rounds 1–15. Analysis sample limited to respondents with at least a high school diploma or GED as of 2009 and no enrollment in school in 2010. Sampling weight in round 1 is used for all analyses. The first model examines the probability of ever changed "state of residence" since 18 years old to year 2011, and the second model examines the probability of residing in a different state in 2010 from the state of residence at age 17. Controls in the above models include: gender, race/ethnicity, work experience, work experience squared, geographic and family background variables, high school GPA, and ASVAB test score. Omitted attainment category is high school diploma/GED only. We do not report estimates for the small number ($n < 50$) of people who enrolled in some unknown type of college without obtaining a degree.

*** $p < .01$, ** $p < .05$, * $p < .1$

Interestingly, the first set of results show that all levels of college attainment are positively associated with having ever moved, but only four-year enrollment/attainment are associated with remaining out of state. It is possible that some of the moves picked up by the first measure are related to the college enrollment itself (i.e., students moving out of state to attend college, but then later moving home).

The results in the second column generally confirm prior evidence that bachelor's degree holders are more mobile: those with a bachelor's degree are 13 percentage points more likely to live outside their home state than those with only a high school credential. Interestingly, even those with only some four-year college exposure without obtaining a degree are significantly more likely to live outside of their home state in 2010. The coefficients on other levels of enrollment/attainment are insignificant and close to zero, but people with certificates, associate

degrees, or some college exposure are more likely to have moved across states at least once since age 18, relative to high school graduates only.

In Table 5, we further explore how patterns of interstate mobility interact with estimated returns to credentials. In the first column, we repeat estimated returns based on Model 4 from Table 2; in the second column, we add interactions of attainment levels with an indicator for living outside the home state in 2010, when earnings are measured. We find that employed bachelor's degree holders and four-year college non-completers who live outside their home state have modestly higher earnings than their counterparts who remain in-state ($7,724 more for bachelor's degrees and $5,899 more for four-year college non-completers).[17] As for employed associate degree holders, those who live outside their home state earn $5,249 *less* than associate degree holders who remain in-state (though this difference is not statistically significant).[18] Mobility is not significantly related to earnings returns for other attainment levels. When we use unconditional earnings, results are similar, except that the "mobility dividend" for four-year college enrollees and completers is notably larger when using the unconditional measure.

---

[17] Kline and Moretti (2013) show that when using real wages rather than nominal wages, part of the payoff of moving to states that offer higher average salaries is offset by higher housing costs in such states. In our analysis sample, respondents who worked in different states than their home states most often lived in California (7.7 percent), Texas (6.8 percent), Florida (6.5 percent), or New York (5.9 percent).
[18] There could be self-selection issues. For example, respondents with associate degrees might all prefer to remain in state to work, and those who could not do that may have ended up moving out of their own states to work.

Table 5: How Estimated Returns Interact With Interstate Mobility (College Versus High School Only)

Variable	Conditional on positive earnings						Unconditional earnings					
	Baseline			Add interactions			Baseline			Add interactions		
	Coef.	(S.E.)		Coef.	(S.E.)		Coef.	(S.E.)		Coef.	(S.E.)	
Earnings in 2010 (in 2015 dollars)												
Some 2-yr coll. (only)	-$1,122	(1228)		-$1,069	(1333)		-$499	(1060)		-$529	(1152)	
Some 4-yr coll. (only)	$4,981	(1452)	***	$3,540	(1647)	**	$5,037	(1278)	***	$3,491	(1451)	**
Certificate	$624	(2553)		$544	(2858)		$575	(2225)		$23	(2462)	
Associate degree	$6,794	(1708)	***	$8,017	(1884)	***	$7,917	(1552)	***	$9,191	(1720)	***
Bachelor's degree	$17,942	(1217)	***	$15,832	(1324)	***	$20,939	(1095)	***	$18,605	(1199)	***
Some 2-yr coll., working OOS				-$525	(2653)					-$303	(2348)	
Some 4-yr coll., working OOS				$5,899	(2698)	**				$7,225	(2461)	***
Certificate, working OOS				-$7,317	(7264)					-$4,877	(6420)	
Associate degree, working OOS*				-$5,249	(3968)					-$4,928	(3650)	
Bachelor's degree, working OOS				$7,724	(1433)	***				$8,528	(1378)	***
HS/GED (only) mean		$31,420						$23,199				
N		4,699			4,417			5,735			5,371		

Note. Source is authors' analysis using NLSY97 restricted data, rounds 1–15. Analysis sample limited to respondents with at least a high school diploma or GED as of 2009 and no enrollment in school in 2010. Sampling weight in round 1 is used for all analyses. Columns 2 and 4 add interactions between educational levels and a "working out-of-state (OOS)" dummy. Controls in the above models include: gender, race/ethnicity, work experience, work experience squared, geographic and family background variables, high school GPA, and ASVAB test score. Omitted attainment category is high school diploma/GED only. We do not report estimates for the small number ($n < 50$) of people who enrolled in some unknown type of college without obtaining a degree, as well as for those who attended some unknown type of college and also worked out-of-state in 2010 ($n < 10$). If earnings are expressed in logarithmic form instead of in dollar values, associate degree holders who worked out-of-state earn 20 percent less than associate degree holders who worked in their home state, and the coefficient is statistically significant.

*** $p < .01$, ** $p < .05$, * $p < .1$

5. How Do Typical Limitations of Administrative Data Influence Estimated Returns?

Sample Restrictions and Control Variables

In Table 6, we directly evaluate two common limitations of state administrative databases that are increasingly used to estimate returns to degrees: first, they often lack of rich control variables, particularly for incoming student ability; second, they typically include data only on college entrants, meaning any returns to college enrollment itself will not be captured in resulting estimates. To address these, the top panel of Table 6 restricts the sample to only college enrollees and repeats the four models originally shown in Table 2 (estimates from Table 2 are reproduced in the bottom panel, to facilitate comparisons).

Table 6: Comparing Estimates With Different Sample Limitations and Control Variables

Variable	Model 1 Coef.	(S.E.)		Model 2 Coef.	(S.E.)		Model 3 Coef.	(S.E.)		Model 4 Coef.	(S.E.)	
College enrollees only				(College non-completers mean = $34,496)								
Certificate	**-$1,614**	**(2852)**		-$1,645	(2760)		-$1,327	(2737)		-$630	(2725)	
Associate degree	**$6,502**	**(1849)**	***	$6,369	(1791)	***	$6,375	(1781)	***	$5,568	(1774)	***
Bachelor's degree	**$17,221**	**(1116)**	***	$20,519	(1098)	***	$18,546	(1131)	***	$16,569	(1167)	***
All high school completers				(HS/GED only mean = $31,419)								
Some 2-yr coll. (only)	$1,060	(1278)		$839	(1227)		-$254	(1228)		-$1,122	(1228)	
Some 4-yr coll. (only)	$7,354	(1468)	***	$8,785	(1410)	***	$7,271	(1423)	***	$4,981	(1452)	***
Some coll. (only, type unk.)	$1,988	(4112)		$3,671	(3945)		$2,991	(3918)		$2,130	(3900)	
Certificate	$1,911	(2688)		$2,345	(2580)		$1,444	(2564)		**$624**	**(2553)**	
Associate degree	$9,944	(1758)	***	$10,360	(1688)	***	$8,953	(1690)	***	**$6,794**	**(1708)**	***
Bachelor's degree	$20,620	(1084)	***	$24,355	(1053)	***	$21,215	(1130)	***	**$17,942**	**(1217)**	***
Model controls:												
Incl. age, race, gender	X			X			X			X		
Incl. exp. and exp. squared				X			X			X		
Incl. geographic and family background							X			X		
Incl. HS GPA and ASVAB										X		

Note. Source is authors' analysis using NLSY97 restricted data, rounds 1–15. Analysis sample limited to respondents with at least a high school diploma or GED as of 2009 and no enrollment in school in 2010. Sampling weight in round 1 is used for all analyses. Regressions limited to 83.6 percent of respondents with positive job earnings in 2010. Job earnings refer to self-reported income from all jobs held in 2010, including self-employment, wages, salary, commissions, and tips. Omitted attainment category for top panel is college non-completers (i.e., respondents who attended some college without obtaining a credential); for bottom panel it is high school graduates.

*** $p < .01$, ** $p < .05$, * $p < .1$

A few interesting patterns emerge. First, and unsurprisingly, estimated returns are consistently smaller in the top panel restricted to college enrollees regardless of the model. The underestimation is less dramatic for bachelor's degree holders (estimates in the top panel are 8–15 percent smaller than in the bottom panel) than for associate degree holders (top panel underestimates by 18–39 percent) and certificate completers (estimates in the top panel flip sign to become negative). These patterns reflect the fact that for different degree levels, college enrollment itself represents a different proportion of the overall return. For example, college non-completers may earn *more* credits than typically accrue to certificate completers, but still have significantly fewer credits than the typical bachelor's degree holder.

Second, while controlling for student characteristics still affects the magnitude of estimated coefficients in the sample limited to college enrollees, the impact of these controls is notably smaller than in the unrestricted sample. For example, moving from Model 1 to Model 4 reduces the estimated return to an associate degree by only 14 percent in the top panel (compared to 32 percent in the bottom panel) and reduces the estimated return to a bachelor's degree by only 4 percent in the top panel (compared to 13 percent in the bottom panel). This suggests that

much of the impact of student characteristics is on the college enrollment decision itself, rather than what happens afterward.

Putting these two countervailing biases together, it appears that administrative-data based estimates, which are restricted to college enrollees and typically have only basic demographic controls, are actually quite similar to fully controlled estimates from an unrestricted sample, at least for associate and bachelor's degrees. Our associate degree and bachelor's degree estimates restricted to college enrollees with only age, race/ethnicity, and gender controls (Model 1, top panel) are virtually indistinguishable from our fully controlled estimates in the unrestricted sample (Model 4, bottom panel). Returns to certificate holders appear to be substantially understated on net, because they are simply not that different from college non-completers. But the estimates for certificate holders are very noisy, and the changes across models and samples are not statistically significant despite the shifts in magnitude.

Treatment of Out-of-State Earnings

Another potentially serious limitation of state administrative datasets relative to survey data like the NLSY is that administrative datasets typically do not enable analysts to distinguish who remains in state and who has left. If a respondent is working in-state, the analyst knows they are still in-state; however, for those not working in-state, it is not possible to distinguish who may be working out-of-state from those who may still be in the state but simply not working.

Thus, in Table 7 we compare results using full information in the NLSY to results that approximate the data limitations in state datasets, and different ways of dealing with them. Columns 1 and 2 show unconditional earnings, while columns 3 and 4 show earnings conditional on any earnings. Columns 1 and 3 are based on a fully controlled model with earnings data from all states (i.e., our best estimates of the returns to degrees, above and beyond college enrollment alone). Columns 2 and 4 simulate estimates that could be realistically obtained with state administrative data: both use only basic controls for age, race/ethnicity, and gender. Column 2 further assumes that analysts take a naive approach and simply treat all missing earnings observations as zeros (to proxy this, we recode all earnings to zero if respondents live outside their home state). Column 4 provides estimates of effects for those who are working and living in their home state.

Table 7: Alternative Approaches to Treating Out-of-State Earnings

Variable	Full Sample (unconditional earnings)				Conditional on positive earnings from…			
	(1) baseline		(2) naïve treatment (OOS=>0)		(3) …any state (baseline)		(4) …home state only	
	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)	Coef.	(S.E.)
College enrollees only								
Certificate	-$1,000	(2414)	$996	(2561)	-$630	(2725)	-$682	(2849)
Associate degree	$6,355	(1647) ***	$8,592	(1738) ***	$5,568	(1774) ***	$7,746	(1889) ***
Bachelor's degree	$19,151	(1087) ***	$10,649	(1049) ***	$16,569	(1167) ***	$15,129	(1188) ***
All high school completers								
Some 2-yr coll. (only)	-$499	(1060)			-$1,122	(1228)		
Some 4-yr coll. (only)	$5,037	(1278) ***			$4,981	(1452) ***		
Some coll. (only, type unk.)	$2,012	(3523)			$2,130	(4250)		
Certificate	$575	(2225)			$624	(2553)		
Associate degree	$7,917	(1552) ***			$6,794	(1708) ***		
Bachelor's degree	$20,939	(1095) ***			$17,942	(1217) ***		
Model controls:								
Incl. age, race, gender	X		X		X		X	
Incl. exp. and exp. squared	X				X			
Background	X				X			
Incl. HS GPA and ASVAB	X				X			

Note. Source is authors' analysis using NLSY97 restricted data, rounds 1–15. Analysis sample limited to respondents with at least a high school diploma or GED as of 2009 and no enrollment in school in 2010. Sampling weight in round 1 is used for all analyses. The first column presents unconditional earnings for the full sample, and the second column uses the same sample but replaces earnings to zero for individuals working out-of-state. The third and fourth columns condition on earning positive job income, and the fourth column also conditions on working in his/her home state. The last two models condition on working full-time and year-round, as proxied by earning at least the equivalent of minimum wage for 35 hours/week, 50 weeks/year ($12,688 annually); the last model also conditions on working in his/her home state. Controls in the above models include: gender, race/ethnicity, work experience, work experience squared, geographic and family background variables, high school GPA, and ASVAB test score. Omitted attainment category is high school diploma/GED only.

*** $p < .01$, ** $p < .05$, * $p < .1$

The results indicate that the estimates are affected quite differently for bachelor's degree versus associate degree attainment. Comparing columns 1 and 2, naively treating missing earnings as zeros underestimates the returns to a bachelor's degree by about half (whether or not we restrict the sample to college enrollees). For associate degree completion, doing this substantially *overestimates* returns—by about 35 percent in the college enrollee sample—

because more college non-completers migrate out of state.[19] Still, if we compare the feasible estimates of associate degree returns in column 2 (top panel) to our best estimates of the full returns in our high school completer sample (column 2, bottom panel), the returns are only overestimated by 8 percent.

Comparing column 4 to column 3 shows that at least for bachelor's degree holders, conditioning on positive earnings helps reduce the bias from out-of-state mobility substantially, but it does not eliminate it completely. Keep in mind that as discussed previously, conditioning on employment itself will tend to underestimate the total returns to a bachelor's degree, but at least column 4 does a much better job of estimating the conditional impact it purports to estimate. For associate degree holders, conditioning on employment does not matter much, and if anything, it makes the bias from out-of-state mobility slightly worse.

6. Discussion

In this paper, we use recent waves of data from the National Longitudinal Survey of Youth 1997 (NLSY97) to provide new, nationally representative, non-experimental estimates of the returns to degrees, as well as to explore the relationship between college attainment and interstate mobility. The NLSY allows us to control for rich information on students' family background, high school academic performance, and test scores, which are often unavailable in state administrative databases. Moreover, we use information on individuals' state of residence in each year to construct individuals' interstate migration pattern and thus explore the relationship between college attainment and interstate mobility. Lastly, we examine the sensitivity of our estimates of the returns to college enrollment and degree attainment, depending upon sample restrictions, adjustments for pre-college achievement and family background, and alternative ways of addressing interstate mobility.

Our findings indicate that relative to high school credentials only, there are substantial returns to bachelor's and associate degrees, as well as to enrollment in four-year colleges without obtaining a degree. People who have a bachelor's or associate degree have annual earnings nearly $21,000 and $8,000 higher than high school graduates respectively, when zero earnings are included. Conditional on employment, bachelor's and associate degree holders still earn $18,000 and $6,800 more respectively. People who enrolled in four-year colleges without obtaining a degree earn about $5,000 more than high school graduates. We also find positive effects on other labor market outcomes including likelihood of employment, full-time, full-year employment, and earning an annual "living wage."

---

[19] Note: thus far, we have not made any sample restrictions based on state of college enrollment. The impact of out-of-state mobility is likely overstated here because four-year non-completers are more likely than associate degree holders to move out of state. However, many of these individuals also may have migrated out of state for college itself. If we were able to condition on enrolling in a specific state, it is not clear whether four-year non-completers would still have higher rates of out-of-state migration.

Our study also explores how college enrollment and attainment relate to interstate mobility and finds that four-year college enrollees and graduates are significantly more likely to work outside their home state after college. These out-of-state migrants also earn more than their counterparts who work in their home state. In contrast, two-year enrollment and attainment have no effect on the likelihood of living outside one's home state, but students with this level of education who do migrate tend to earn *less* than their counterparts who remain in their home state, possibly due to state-specific licensing requirements for some of the fields (such as health) that are most common among two-year enrollees and graduates.

Lastly, we explore the sensitivity of estimated returns to college, by simulating different sample restrictions, inclusion of different sets of covariates, and alternative ways of treating out-of-state earnings to approximate the real-world limitations of state administrative databases. We find that controlling for ability and other background characteristics matters more for lower degree levels, and matters more when we estimate returns among all high school graduates rather than limiting our sample to college enrollees. Failure to control for measures of student ability leads to upward bias, while limiting the sample to college enrollees only leads to an understatement of degree returns. On net, these two biases roughly balance out, suggesting that administrative data-based estimates may reasonably approximate true returns.

The challenge of interstate mobility, combined with the typical limitation of state databases to in-state employment, is more complicated. A simplistic approach that treats all out-of-state earnings as zero earnings (as many state databases, by default, do) will severely understate the returns to bachelor's degrees, but overstate the returns to associate degrees. Conditioning on in-state employment greatly ameliorates the problem for bachelor's degrees, but does not address the problem for associate degree estimates. Still, it is worth noting one reason why our simulation may overstate the problem of interstate mobility: we consider enrollment and degree completion in any state, but state databases typically start with a cohort of in-state college enrollees. Those who choose to enroll in-state may be less likely to later migrate out-of-state.

Overall, the estimates and analysis provided in this paper are reassuring about the validity of administrative-data based estimates. Even when magnitudes of estimated returns vary somewhat across samples and specifications, the broad patterns are quite stable, and our NLSY-based estimates are quite consistent with prior estimates based on state administrative data. This is good news, because state administrative data has many practical advantages: their large size and frequency make it feasible for policymakers to track outcomes over time for many smaller subgroups than is possible in the NLSY. State administrative databases have a particularly strong comparative advantage when it comes to estimating returns to specific degree-fields, especially at the sub-baccalaureate level (for which numbers of observations are quite small in the NLSY). The main comparative advantage of NLSY and other survey-based data may not be in estimating earnings returns to college attainment, though we could not have concluded that before undertaking the analyses presented here. Rather, the comparative advantage of survey data may be in examining some of the nonpecuniary returns to attainment—including better fringe

benefits, greater job satisfaction, and more schedule flexibility—that are not included in administrative data (Schudde, 2016).

References

Bahr, P. R., Dynarski, S., Jacob, B., Kreisman, D., Sosa, A., & Wiederspan, M. (2015). *Labor market returns to community college awards: Evidence from Michigan* (CAPSEE Working Paper). New York, NY: Center for Analysis of Postsecondary Education and Employment.

Belfield, C. R., & Bailey, T. (2011). The benefits of attending community college: A review of the evidence. *Community College Review*, *39*(1), 46–68.

Black, D., Sanders, S., & Taylor, L. (2011). Measurement of higher education in the census and current population survey. *Journal of the American Statistical Association*, *98*(463), 545–554.

Bound, J., & Holzer, H. J. (2000). Demand shifts, population adjustments, and labor market outcomes during the 1980s. *Journal of Labor Economics*, *18*(1), 20–54.

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (3rd ed., pp. 1801–1863). Amsterdam, NL: North Holland.

Dadgar, M., & Trimble, M. J. (2015). Labor market returns to sub-baccalaureate credentials: How much does a community college degree or certificate pay? *Educational Evaluation and Policy Analysis*, *37*(4), 399–418.

Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems. *Econometrica*, *45*(1), 1–22.

Grubb, W. N. (1997). The returns to education in the sub-baccalaureate labor market, 1984–1990. *Economics of Education Review*, *16*(3), 231–245.

Jepsen, C., Troske, K., & Coomes, P. (2014). The labor-market returns to community college degrees, diplomas, and certificates. *Journal of Labor Economics*, *32*(1), 95–121.

Kline, P., & Moretti, E. (2013). Place based policies with unemployment. *The American Economic Review*, *103*(3), 238–243.

Liu, V. Y. T., Belfield, C. R., & Trimble, M. J. (2015). The medium-term labor market returns to community college awards: Evidence from North Carolina. *Economics of Education Review*, *44*, 42–55.

Malamud, O., & Wozniak, A. (2012). The impact of college on migration: Evidence from the Vietnam Generation. *Journal of Human Resources*, *47*(4), 913–950.

Marcotte, D. E., Bailey, T., Borkoski, C., & Kienzl, G. S. (2005). The returns of a community college education: Evidence from the National Education Longitudinal Survey. *Educational Evaluation and Policy Analysis*, *27*(2), 157–175.

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York, NY: Columbia University Press.

Oreopoulos, P., & Petronijevic, U. (2013). Making college worth it: A review of the returns to higher education. *The Future of Children*, *23*(1), 41–65.

Schudde, L. (2016). *Nonpecuniary returns to postsecondary education: Examining early non-wage labor market outcomes among college-goers in the United States* (CAPSEE Working Paper). New York, NY: Center for Analysis of Postsecondary Education and Employment.

Wozniak, A. (2010). Are college graduates more responsive to distant labor market opportunities? *Journal of Human Resources*, *45*(4), 944–970.

Xu, D., & Trimble, M. J. (2015). What about certificates? Evidence on the labor market returns to nondegree community college awards in two states. *Educational Evaluation and Policy Analysis*, *38*(2), 272–292.

Zeidenberg, M., Scott, M., & Belfield, C. R. (2015). What about the non-completers? The labor market returns to progress in community college. *Economics of Education Review*, *49*, 142–156.