Which College Types Increase Earnings? Estimates from Geographic Proximity

Jennifer L. Steele

American University [1]

steele@american.edu

January 2023

[1]School of Education, 4400 Massachusetts Ave., NW, Washington, DC 20016-8030 Desk: (202) 885-3762

Abstract

The question of why postsecondary institutions produce different labor market outcomes is difficult to answer due to unobserved student characteristics that drive institutional choice as well as earnings. I leverage students' geographic proximity to three extant classifications of postsecondary institutions—earnings-enhancing, competitive, and Historically Black Colleges and Universities. Using a nationally representative sample, I estimate attainment and earnings effects of first attending each type. Attending an institution classified as earnings-enhancing raises early-career wages, whereas underrepresented students' HBCU enrollment lowers wages. Initial attendance at both types yields higher bachelor's degree and humanities credit completion, and HBCU enrollment yields additional STEM credits.

Highlights

- The attributes of the nearest college predict first choice of college for many high school students.
- Colleges previously linked to students' wage mobility yield higher earnings by students' mid-20s.
- Higher earnings effects coincide with higher humanities credit completion and bachelor's degree completion.
- HBCU attendance yields higher degree completion, as well as humanities and STEM credit completion, but lower early-career wages.

Keywords: human capital, salary wage differentials, institutional effects, instrumental variables, college proximity
JEL codes: I23, I26, R32
Abbreviations: *STEM*: Science, Technology, Engineering, and Mathematics; *HBCU*: Historically Black College or University; *HSI*: High-Success Institution; *Competitive*: Barron's Top 3 Selectivity Tier Institution; *Underrepresented Minority (URM)*: Black, Indigenous, or Hispanic/Latinx

1 Introduction

It is widely understood that graduation rates and labor market earnings vary by the prestige and selectivity of postsecondary institutions, but there is also evidence that such variation is not due entirely to the attributes of the students who attend them (Carnevale and Strohl, 2010). For instance, Dale and Krueger (2002) compared high school graduates who applied to and were accepted by similar colleges but matriculated to more- or less-competitive colleges as defined by average SAT scores or Barron's selectivity category ratings. They found that attending a more selective college conferred few earnings advantages, except among low-income students, whose earnings benefited as much as 8% by attending a college with 200-point-higher average SAT score. Along similar lines, **?** used a multi-attribute composite of institutional quality measures in a non-parametric, generalized method of moments framework, finding that attending an institution with higher composite quality is linked to higher labor market earnings of up to 4% in the following decade. In addition, Long (2008) conducted a comparison of these methods using the National Education Longitudinal Study of 1988, finding that a standard deviation difference in composite institutional quality predicted 8% to 18% higher earnings for males in ordinary least squares and generalized methods of moments specifications, though effects for females and for instrumental variable models (instrumented by nearest institution average quality) were not statistically significant. Internationally, regression discontinuity evidence in France suggests that access to higher-quality institutions for disadvantaged students increased peer quality and STEM coursetaking and conferred a 13% earnings advantage (Canaan and Mouganie, 2018).

Fortunately, students do have greater information than those in past generations about how much graduates of particular institutions can expect to earn after college. The College Scorecard, which was launched by the U.S. Department of Education in 2015, provides public data about costs, graduate earnings, and loan default rates for more than 5,000

U.S. postsecondary institutions (Rothwell, 2015). To better gauge the earnings effects of particular institutions, Chetty et al. (2017) described the economic mobility of the students who attended them. Linking federal tax return data for 30 million U.S. adults to those of their parents, the authors found that college choice strongly mediated the positive association between parents' and children's earnings. Overall, parents' earnings rankings were positively linked to those of their adult offspring, with a coefficient of about 0.29, *but within the offspring's postsecondary institutions*, this relationship largely disappeared. The study also sought to examine the attributes of institutions in which a large share of students from bottom-quintile socioeconomic backgrounds moved to the top quintile by their early 30s. The authors termed the fraction making this transition as the "success rate" of a given institution. But Chetty et al. (2017) did not find institution-level covariates, such as share of STEM graduates, that strongly predicted institutions' success rates. [1]

Of course labor market prospects are not students' sole concern when selecting a postsecondary institution. Institutions that historically existed to serve marginalized groups may hold strong appeal for students whose identities align with those institutional foci. This alignment may, in turn, promote stronger labor market outcomes. For instance, Black students' enrollment at Historically Black Colleges and Universities (HBCUs) in lieu of Predominantly White Institutions (PWIs) has been linked to higher college persistence and graduation (Franke and Deangelo, 2018; Fryer and Greenstone, 2010), to higher socioeconomic status over time (Price et al., 2011), to higher earnings in the 1970s and more recently for female graduates (De Zeeuw et al., 2020), but also to earnings disadvantages in the 1990s and in STEM fields more recently (De Zeeuw et al., 2020; Fryer and Greenstone, 2010). One constraint of these studies is that they have largely employed propensity score matching or weighting methods. Such methods address selection on observable but not on unobservable

[1] Chetty et al. (2017) also found high variation in the "mobility rates" of institutions, which they defined as the product of institutions' success rates and the fraction of bottom-quintile students enrolled in them.

student attributes, and thus may be confounded by any baseline student attributes that are correlated with both institutional choice and the outcomes of interest (Angrist and Pischke, 2008; Steiner et al., 2011). Still, survey-based and qualitative research has shed light on instructional practices at HBCUs that may promote strong outcomes for Black students, including higher rates of undergraduate participation in faculty research (Kim and Conrad, 2006; Perna et al., 2009), as well as institutional practices, such as small class sizes, supportive faculty outreach, accessible faculty offices, and available peer tutoring, that facilitate strong student performance, especially in STEM fields (Palmer and Gasman, 2008; Perna et al., 2009). Qualitative studies have also pointed to supportive and collaborative peers as contributors to student success in HBCUs (Palmer and Gasman, 2008; Perna et al., 2009).

The current paper aims to build on the work of Chetty et al. (2017) by exploring whether the earnings of effects of *"high-success" institutions* appear to be causal. I also examine how such institutions descriptively differ from and overlap with other institution types that compete for college students, including *competitive institutions* (those in the top three of seven *2004 Barron's Guide to Colleges*selectivity ratings), and *minority-serving institutions*, which may be especially appealing to students of underrepresented minority (URM) backgrounds. I define URM students as those of Black, Indigenous, and Hispanic/Latinx heritage, since these groups have historically faced legal and systemic barriers to postsecondary enrollment and remain underrepresented in U.S. higher education settings on average. Because the 2002 dataset employed in the study defines only one type of minority-serving institution—Historically Black Colleges and Universities (HBCUs)—I feature this subset in the institutional category comparison as well.

I define "high-success" institutions as those that Chetty et al. (2017) previously identified with IRS data as having success rates in the top quartile of four-year institutions in 2000, four years before students in the current study typically enrolled in postsecondary education. The success rate is defined as the fraction of students from families with incomes in the bottom

U.S. quintile who have earnings in the top quintile by their 30s. "High-success" therefore refers to extant institution-level calculations by Chetty et al. (2017) using observational analyses.

Because they are defined as the top 3 Barron's selectivity tiers, "competitive" institutions here represent institutions that are relatively selective, helping us understand the extent to which the effects of high-success institutions and HBCUs align with or diverge from the effects of competitive institutions as a group. Barron's rates four-year institutions based on entrance examination requirements and average scores, high school grade point averages of entering students, and admissions rates. The rating tiers in 2004 ranged from 1 (most competitive) to 6 (non-competitive), with the tier of 7 indicating special institutions to whom the criteria did not apply, such as art or music schools that do not base their admissions on academic criteria *per se* (Schmitt, 2016). For descriptive statistics, I impute a rating of 8 to institutions, including all two-year colleges, that did not receive Barron's ratings in 2004.

HBCUs, which opened postsecondary pathways to Black students when other paths were foreclosed by segregation, and which continue to serve large shares of Black students, are identified as such through their 2004 classification in the Integrated Postsecondary Education Data System (IPEDS) produced by the National Center for Education Statistics.

The aim of this analysis is to estimate the effects of first-postsecondary institution type on various measures of attainment, including STEM and humanities credit completion, bachelor's degree attainment, pursuit of postbaccalaureate education, and earnings in 2012. Of course a key challenge of estimating such effects is that the choice of institution type is endogenous. To address this problem, I take advantage of variation in tenth graders' geographic distances to their nearest postsecondary institution and to that institution's classification as "high-success," competitive, HBCU, or other. If I can capture a source of randomness that affects students' institutional choices but is *conditionally uncorrelated* with their pre-college preparedness or motivation, then I can estimate the ways in which institution types influence student outcomes at the margin. One instrumental variable (IV) that has been used for this purpose in prior studies is students' geographic proximity to postsecondary institutions. For instance, Rouse (1995) used proximity to students' nearest two-year college to estimate access and diversion effects of two-year college enrollment on educational attainment, finding that it increased educational attainment by 1 to 1.5 years, though effects on bachelor's degree completion were less evident. Card (1999) found that geographic proximity to a four-year college and the interaction of this proximity with parental education strongly predicted educational attainment. From this, he estimated an earnings return of 9.7% to to each year of schooling. Herbst and Tekin (2016) used the distance families lived from social service agencies as an instrument for subsidy use, finding negative initial effects of subsidized pre-K on children's cognitive skills and behavior, possibly due to low program quality. And Jaggars and Xu (2013) used distance from nearest community college to instrument for online enrollment, finding negative effects of online coursework on persistence and grades. In this study, I adapt an approach more similar to Long (2008), who used the average institutional quality of postsecondary institutions near students' homes to instrument for the quality of institution attended, finding mostly null effects of postsecondary quality relative to OLS and generalized method of moments estimators.

The IV estimates pertain to students whose probability of choosing a given institution type is increased by geographic proximity. As noted by Card (2001), such students may be of special concern to institutions and policymakers because they may be more credit-constrained and risk-averse than the average first-time college entrant. IV methods do not allow me to directly observe the *identities* of individuals who were influenced by geographic proximity (compliers) versus individuals who would have chosen their institution types anyhow (always-takers) (Angrist and Pischke, 2008). However, I observe that within the analytic sample in the current study, students who enrolled in the institution closest to their tenth grade zip codes—about 12.7% of individuals who had enrolled in postsecondary by age 2012—had a composite socioeconomic status level that was 0.3 SD below that of the analytic sample as a whole, as well as tenth grade math scores that were 0.22 SD below that of the full sample. They were also 2 percentage points more likely to be underrepresented minority students, and 2 points less likely to be Asian American or Pacific Islander, as compared to the sample as a whole. These measures imply that students whose choices are geographically sensitive may have greater needs for college information and support than their college-going counterparts, on average.

In the next section, I describe my my data sources, including geographic proximity indicators, outcome variables in terms of attainment and wages by roughly age 26, and a variety of student-level and school-level control variables. This is followed by a description of my analytic approach. In the subsequent section, I present results from OLS models and first- and second-stage IV models. Finally, I discuss the limitations and implications of these results.

2 Data

I use data from a restricted-use version of the ELS:2002, a nationally representative, longitudinal sample of individuals surveyed for the first time as tenth graders in the spring of 2002, and then in three subsequent waves in the spring of 2004, 2006, and 2012. High school and postsecondary transcript data were collected from the students' institutions in 2005 and 2013, respectively. As with many NCES surveys, sampling for the ELS:2002 proceeded in two phases (Ingels et al., 2004). First, 976 schools serving tenth graders were randomly sampled across 50 states and the District of Columbia, with probability proportional to their size. Sampling was stratified by school sector (public, Catholic, private), U.S. Census division, and urbanicity, and school participation was about 77%. Approximately 26 tenth grade students were then randomly sampled from each participating randomly drawn school. Student sampling was stratified by race and ethnicity, with oversampling of Asian and Hispanic students. Participating students (about 87.3%) completed a survey about their school experiences, educational plans, and career goals, as well as cognitive tests in reading and mathematics. Parents, teachers, school administrators, and school librarians of selected students were also surveyed. The full ELS:2002 dataset includes 16,197 students. [2]

To measure the demographics of the zip codes in which students lived when they were in tenth grade and of the first postsecondary institutions they attended, I use selected economic and demographic data from the 2000 decennial U.S. Census. I employ 5-digit Zip Code Tabulation Areas within the United States and Puerto Rico, retrieved from the American FactFinder (U.S. Census Bureau, 2020).

Data on postsecondary institutional characteristics are drawn from the IPEDS, which as noted above is a national dataset of characteristics for postsecondary institutions in the U.S. I focus on IPEDS data from the year 2004, which is the anticipated high school graduation year for students in the ELS:2002 sample. Data on Barron's college selectivity come from a supplement to the restricted-use version of the ELS:2002 provided by the National Center for Education Statistics, and also pertain to the year 2004. In classifying undergraduate courses as STEM, I follow the U.S. Department of Education SMART grant definitions (U.S. Department of Education, 2010) to include agriculture and natural resources, computer and information sciences; engineering and engineering technologies, mathematics and statistics, biological and biomedical sciences, physical sciences, science technology and technicians, and health professions and clinical sciences. Because STEM majors at the institution level are reported by IPEDS only in the four core STEM areas of engineering, biological and

[2] ELS:2002 provides cross-sectional base-year weights for each school and student to reflect both the inverse probability of selection, which is known from the sampling design, and the probability of nonresponse, which is estimated from student and school attributes at baseline. The dataset also includes panel weights for use in longitudinal analyses across the other survey waves. I do not employ the ELS weights in this analysis because my identification strategy, instrumental variables analysis, in effect assigns greater weight to respondents who are sensitive to the set of geographic instrumental variables. Applying sampling and non-response weights may therefore distort the internal validity of the IV analysis (Solon et al., 2015).

biomedical sciences, mathematics/statistics, and physical sciences, I use this definition to calculate the percentage of undergraduate STEM majors and STEM-specific field majors by institution.

2.1 Measures

Descriptive statistics for the analytic samples are displayed in Table 1. The full analytic sample includes 9,053 students, of whom 7,840 (87%) had attended any postsecondary institution after high school and 1,213 had not. The tables presents means and standard deviations for each variable for the full sample and for the subsample of underrepresented minority students, defined as students from Black, Indigenous (American Indian or Alaska Native), or Hispanic (including Latinx) backgrounds. I combine Black, Indigenous, and Hispanic students into a single group to preserve statistical power for the subgroup analysis, and because these groups have been similarly underrepresented in U.S. postsecondary settings. Constraining the subgroup analyses to Black students yields similar results, but with less statistical precision. The table also includes the range of each variable for the full sample.

Among the individual attributes, the socioeconomic status variable (SES) and base math score variable are each standardized among all ELS respondents to have a mean of 0 and SD of 1. The ELS:2002 SES composite includes parents' income, occupational prestige scores, and levels of education. The high school variables denote the percent of students in respondents' tenth grade school who qualified for free or reduced-price lunches (FRL) and who took at least one Advanced Placement (AP) or International Baccalaureate (IB) course during high school as of 2004. Base-year residential zip code variables capture the percent of jobs in the student's residential zip code in 2002 that were classified by the Bureau of Labor Statistics as being in the professional, scientific, or technical services sector (sector 54), as well as residential zip code average years of education among adults, and the zip code median income in thousands of dollars. The first-college attributes indicate the share of the

Table 1: Descriptive statistics for full (n=9,053) and underrepresented (UR) (n=2,776) analytic samples

	Mean All	**SD All**	**Mean URM**	**SD URM**	**Min All**	**Max All**
Female	.53	.5	.54	.5	0	1
Underrep Minority	.31	.46	1	0	0	1
Asian	.093	.29	0	0	0	1
White	.6	.49	0	0	0	1
Native English	.84	.36	.75	.43	0	1
SES	.065	1	-.31	.97	-2.9	2.6
Base Math	.093	1	-.4	.94	-3.2	3.6
HS GPA	2.7	.88	2.3	.85	0	4.6
HS Pct FRL	21	23	33	29	0	100
HS Pct AP	12	14	9.8	12	0	81
Zip Pct Profess	.089	.052	.079	.051	.0069	1
Zip Med Edu Yrs	13	1.1	13	1.1	8.9	17
Zip Med Inc	53	19	46	17	18	175
High-Success	.19	.39	.12	.32	0	1
Competitive	.21	.41	.11	.31	0	1
HBCU	.018	.13	.048	.21	0	1
Nearest is HSI	.098	.3	.13	.33	0	1
Nearest is Compet	.089	.29	.095	.29	0	1
Nearest is HBCU	.031	.17	.057	.23	0	1
Miles to Nearest	7.5	11	6.8	10	0	158
State Share 4=Yr	.57	.091	.55	.098	.32	1
Nearest Tuit Ratio	.89	.81	.9	.91	0	6.3
Attended Postsec	.87	.34	.82	.39	0	1
Credits Humanit	29	33	23	31	0	329
Credits STEM	23	33	18	29	0	409
Bach Degree	.43	.5	.27	.45	0	1
Post-Bac Pursuit	.13	.33	.067	.25	0	1
Hourly Wage	16	10	14	9.4	0	125

analytic sample whose first college was classified in 2004 as high-success (Chetty et al., 2017), competitive (Barron's Top 3 selectivity tiers), or an HBCU. In this study, I do not code other types of minority-serving institutions, such as those serving predominantly Hispanic students [Hispanic Serving Institutions—not to be confused with "high-success" institutions as defined by (Chetty et al., 2017)] or Asian American, Native American, and Pacific Islander students [Aanapisi] because these designations were not documented in the IPEDS data from 2004, though they would be useful to examine with later datasets. The geographic, distance-based instruments include dichotomous indicators of whether the institution nearest the student is high-success, comparative, or an HBCU, and distance in miles from the student's tenth grade residence to the nearest postsecondary institution. For instance, 10% of the full sample and 13% of the URM sample had an HSI as their nearest institution; the comparable figures for HBCUs were about 3% and 6%, respectively. My analysis includes controls for two other geographic attributes from IPEDS that may affect college access, including the share of institutions in the students' tenth-grade residential state that were four-year institutions in 2002, and the ratio of the tuition at the nearest institution to that of the state average. Finally, the dependent variables comprising the lower five rows of Table 1 show that only about 43% the full sample and 27% of the URM sample had obtained a bachelor's degree or higher by 2012. Average hourly wages in 2012 were $16 in the full sample and and $14 in the URM sample.

Figure 1 presents the institutional sample size and Barron's rating distribution of each of the key institutional types of interest in the student-level dataset, alongside those of all four-year institutions and all postsecondary institutions, two-year and four-year, in the dataset. The shaded boxes show the interquartile range of the Barron's ratings, where a vertical bar in the midst of a shaded box indicates the median, and the bar and whiskers outside the boxes, where visible, extend to the 5th and 95th percentiles, respectively. By definition, competitive institutions have ratings of 1 to 3; their median and lower quartile are 3. High-

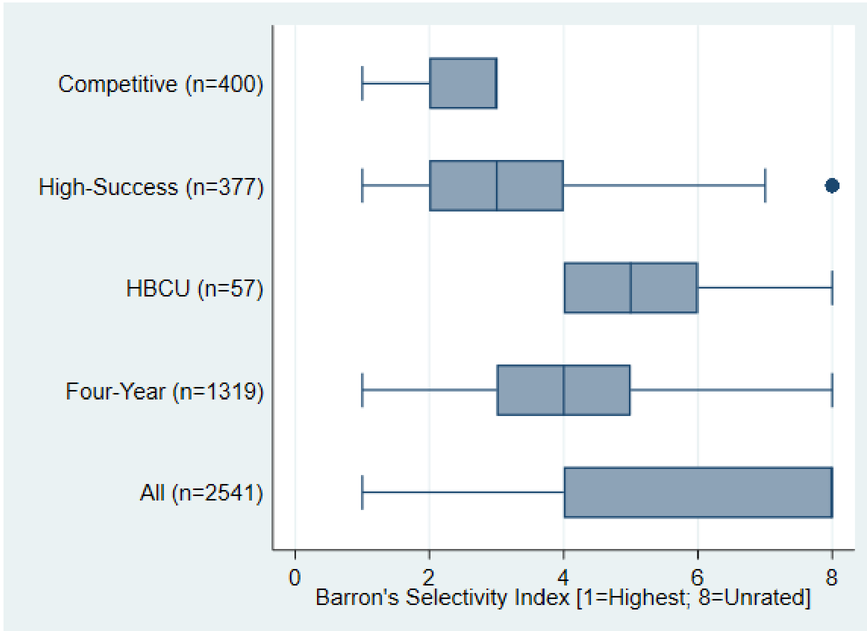

Figure 1: Barron's selectivity ratings of postsecondary institutions attended by ELS:2002 respondents

success institutions are relatively competitive but have a broader distribution of ratings than the competitive institutions, and the HBCUs have a distribution that reflects their mission of expanding postsecondary access to Black students whose access to educational opportunity in the U.S. has been systematically constrained.

Table 2 breaks the partially overlapping institutional categories in Figure 1 into mutually exclusive subcategories to illuminate how their mean attributes differ. The first subcategory (High-Success and Competitive) includes the intersection of these two subcategories. The next two subcategories (High-Success Only and Competitive Only) represent the institutions in each group that do not overlap with the other group. The *Other* category includes all other institutions (including two-year) that are not classified into one of the other categories in the table. The intersection of HBCU with HSI is shown in the fifth data row, comprising four institutions, and the remaining 53 HBCUs are represented as HBCU Other. [3] It

[3]The four HBCUs also classified as high-success institutions are Howard University, Morehouse College, Spelman College, and Xavier University of Louisiana.

Table 2: Mean Institutional Attributes by Subcategory

Category	N	Priv NP	Four Yr	Per Pupil	Pct STEM	Pct Busine	Barron	Fam Inc
HSI+ Compet	245	0.74	1.00	10604	16.57	15.90	2.25	125,511
HSI Only	128	0.52	0.98	5638	11.17	17.61	4.65	92,784
Compet Only	155	0.69	1.00	6496	11.32	13.89	2.74	102,438
Other	1956	0.19	0.41	3654	2.28	22.27	7.03	68,983
HBCU+ HSI	4	1.00	1.00	9395	27.42	18.24	4.00	76,575
Other HBCU	53	0.26	0.87	4326	9.22	17.46	5.34	45,642
Total	2541	0.30	0.55	4845	4.84	19.61	6.15	79,656

is noteworthy that the intersection of HSIs with competitive institutions has the highest median family incomes and per-pupil expenditures. Also, the fraction of STEM majors is higher in HBCU-HSIs than in any other institutional category, at 27.4%.

3 Analytic Strategy

To estimate the causal effects of the attributes of students' first postsecondary institution, I leverage the attributes of the postsecondary institution that is geographically nearest to their residential zip code in tenth grade. The rationale is that students do not typically choose where they live in high school, and that families of tenth graders are unlikely to have chosen their residences based on their proximities to particular postsecondary institutions that they aspired for their children to attend. Meanwhile, several studies have shown that geographic proximity does influence college going by partially determining the cost of college, in terms of housing and commuting costs as well as informational access Card (2001); Rouse (1995); **?**. I use the `geonear` command in Stata to calculate geodetic distances between 2004 IPEDS institutional zip codes and the tenth grade residential zip codes of each student in the

ELS:2002 sample of students who attended any postsecondary institution between 2002 and 2012. I instrument for the attributes of the student's first postsecondary institution using attributes of the geographically closest institution, including whether it is competitive, high-success, or an HBCU. [4] I also instrument for miles to nearest postsecondary institution, a decision that modestly improves first-stage precision, but to which second-stage results are not highly sensitive. I use an F-test of instrument strength so that noisy and trivial associations between instruments and endogenous predictors of interest are not exaggerated into biased estimates (Stock et al., 2002; Small, 2008).

The two-stage instrumental variable model is specified as in the following pair of equations:

$$m^r_{ik4} = \alpha_2 + \beta_2 N_{ik0} + \delta_2 X_{ik} + \epsilon_{2ik4} \tag{1}$$

$$y_{ik4} = \alpha_3 + \nu_3 \hat{m}^r_{ik4} + \delta 3 X_{ik} + \epsilon_{3ik4} \tag{2}$$

In the first stage equation, the dependent variable m_{ik4} is a dichotomous indicator of whether student i from tenth grade school k first enrolled in a postsecondary institution of type r, conditional on having ever enrolled in a two-year or four-year institution by 2012, when most survey participants were about age 26. The types of r are competitive, high-success, and HBCU, and the model is estimated separately for each type. The set of aforementioned, geographically based instrumental variables is included in vector N_{ik0}. Control vector X_{ik0} includes controls for individual attributes of race/ethnicity; gender; a standardized family socioeconomic status composite based on parents' occupation, earnings, and education levels; an indicator of whether the student's first language is English; and a standardized mathematics skills score on a test administered in tenth grade. The control vector also includes school-based controls for share of students in the school who qualified for subsidized meals, and the share enrolled in Advanced Placement or International Baccalaureate courses in high

[4] I attempted to instrument for entry into a high-STEM institution as well, but the set of geographic instruments did not predict entry into this type of institution.

school, a measure of the share of employers in the student's residential zip code that were in the professional sector, as well as residential zip code median income and average years of education among adults. The latter serve as controls for the socioeconomic milieu of the residential zip code, to help disentangle effects of geographic proximity to college attributes from the effects of the geographic socioeconomic environment. The IV models also control for two other geographic variables—the ratio of the nearest institution's tuition ratio to the state average tuition, and the fraction of four-year institutions in the state, since this may influence average tuition in the state.

In the second-stage equation, the predicted probability of initially enrolling in institution type r from the first stage, $\hat{m}^r_{ik4}$, becomes the substantive predictor of interest. Its coefficient, ν_3, captures the Local Average Treatment Effect (LATE) of first enrolling in institution type r on each dependent variable of interest, adjusting for the other variables in the model. The LATE in this context refers to the causal effect for students whose choice of first institution is influenced by the geographic proximity and type of the institution nearest their tenth grade residence. Because local proximity influences cost and convenience of college, it is likely to be most influential for students who are less-affluent and who are on the margin of college attendance in general, as noted above. Estimates of the LATE may be especially relevant for students who are already at higher risk of not earning a postsecondary credential, and thus to policy decisions about institutional support and financial aid for particular types of institutions and institutional practices. The dependent variables in the second-stage include the number college credits obtained by 2012 in (1) humanities and (2) STEM courses (respectively); dichotomous indicators of whether the participant had (3) obtained at least a bachelor's degree by 2012, or (4) had ever pursued post-baccalaureate study by 2012, and an indicator of the participant's (5) hourly wages from work (scaled in 2012 \$) in the year leading up to the 2012 survey. In stage two, the error term is given by ϵ_{3ik4}, and the intercept by α_3, with corresponding parameters shown for the first stage as well.

Standard errors are clustered at the base high school level as an adjustment for possible interdependence of students' postsecondary decision-making within schools. The IV model assumes that the instruments influence the outcomes of interest only through their effects on first-institutional enrollment (the exclusion restriction) and that proximity works in only a single (positive) direction, with no effects on the choice of other focal institution types (unordered monotonicity) (?).

4 Findings

4.1 Institution-Subcategory Relationships to Courses, Attainment, and Earnings

As a first step, I describe the attainment and earnings outcomes associated with each of the mutually exclusive institutional subcategories shown in Table 2. Table 3 presents Ordinary Least Squares (OLS) estimates of these relationships. I include statistical controls for all aforementioned covariates including individual, high school, and residential zip code characteristic controls. The omitted reference category in each model is having never attended a postsecondary institution.

OLS estimates express linear associations between initial institution types and the outcomes of interest, holding constant numerous attributes of individuals, schools, and base-year residential zip codes. If the control variables captured all relevant differences between students choosing different institution types, they would provide causal estimates. Yet, because students and colleges choose one another based on many considerations, only some of which are captured in large datasets, these OLS estimates cannot be construed as causal.

Table 3 presents estimates for each institutional subcategory of first attendance, relative to respondents who never attended a postsecondary institution. The top panel refers to the

Table 3: OLS Estimates Relating Institutional Attributes to Outcomes

	(1) Humanities	(2) STEM	(3) Bachelor	(4) Postbac	(5) Wage
Full Sample					
High-Success + Compet	19.382***	11.945***	0.396***	0.166***	6.246***
	(1.401)	(1.377)	(0.016)	(0.015)	(0.505)
High-Success Only & Less Compet.	22.443***	16.964***	0.361***	0.049*	5.857***
	(1.459)	(1.742)	(0.020)	(0.020)	(0.647)
Competitive Only	27.946***	14.769***	0.378***	0.120***	4.583***
	(1.852)	(1.614)	(0.018)	(0.017)	(0.528)
Less-Competitive	18.103***	13.490***	0.139***	-0.005	2.361***
	(0.568)	(0.535)	(0.009)	(0.004)	(0.279)
Observations	9,384	9,384	9,367	9,975	9,488
R-squared	0.238	0.190	0.439	0.167	0.130
F-all	357.1	205.8	761.1	102.3	84.46
Schools	713	713	713	717	715
Underrepresented Minority Students					
High-Success + Compet	21.758***	13.109***	0.476***	0.195***	6.752***
	(3.103)	(2.789)	(0.032)	(0.030)	(0.926)
High-Success Only	24.840***	18.383***	0.331***	0.028	4.842***
	(2.749)	(3.880)	(0.048)	(0.031)	(1.141)
Competitive Only	33.251***	15.608***	0.397***	0.075+	6.085***
	(4.487)	(3.846)	(0.045)	(0.040)	(1.124)
Less-Competitive	14.554***	10.145***	0.078***	-0.006	2.234***
	(0.808)	(0.751)	(0.012)	(0.005)	(0.416)
High-Success HBCU	36.975***	37.698**	0.418***	0.036	2.404
	(11.057)	(12.190)	(0.122)	(0.091)	(2.334)
Other HBCU	17.448***	15.589***	0.202***	0.033	1.842*
	(2.655)	(2.534)	(0.041)	(0.022)	(0.726)
Observations	2,957	2,957	2,949	3,221	3,041
R-squared	0.272	0.214	0.398	0.155	0.126
F-all	102.3	66.33	177.6	19.14	23.09
Schools	617	617	617	629	619

Cluster robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Omitted category: never attended postsecondary institution

Controls include race/ethnicity where possible, gender, family SES composite, first language not English, honors-weighted GPA in high school, 10th grade math test score, subsidized meal share in baseline high school, share of students in AP or IB classes in baseline high school, share of employers in professional sector in baseline residential zip code, median zip code income and average zip code years of education.

full analytic sample. The bottom panel refers to underrepresented minority students only. Because each institutional subcategory is evaluated relative to those who did not attend college, most estimates are large and statistically significant. But we can gain some insight by comparing estimates among categories. For instance, in the full sample, attending a high-success institution (even a less-competitive one) predicts hourly wages more strongly than attending a competitive college that is not classified as high-success, by a difference of about \$1.5 per hour, or about 15% of a standard deviation, as shown in column 5. High-success, less-competitive colleges are also more strongly predictive of STEM credit completion, as shown in the top and bottom panels of column 2, whereas competitive colleges not classified as high-success re more strongly predictive of humanities credit completion, as shown in the top and bottom panels of column 1. But for underrepresented minority students, it is also important to note that the strongest wage estimates come from competitive colleges with and without the high-success designation, and that by far the largest estimates in terms of STEM and humanities credit completion come from high-success HBCU institutions.

4.2 First-Stage Predictions of Initial Institution Type

The OLS estimates of initial institution subcategory relationships to outcomes are useful in that they describe the outcomes that one might expect to see from each subcategory, on average. They also show how these outcomes by subcategory may differ for underrepresented minority students as compared to the full sample. But because first-institution choice is a bi-directional process driven by individual qualifications, preferences, and financial resources (individual, governmental, institutional, etc.), the OLS estimates have limited use in helping prospective students understand how first-institution choice might affect their subsequent outcomes, holding constant their own skills, resources, and priorities. For that, we need an element of plausible randomness—a factor outside a young person's control and independent of their broader opportunity set and preferences, that may drive their choice of first institution type. For this, I use the two aforementioned instrumental variables—distance in miles from the nearest postsecondary institution, and whether the nearest institution is one of the institution types of interest here. That includes, namely, a high-success institution (and for comparison's sake, a Barron's top-3 tier competitive institution), and an HBCU. I instrument for enrollment in these broad categories instead of for the subcategories in the OLS estimation for improved precision in the instrumental variable models, and because the substantive question of interest concerns to what extent high-success institutions and HBCUs predict different outcomes from other institution types (including competitive institutions) and why. I estimate these effects separately for the full and underrepresented minority samples.

First-stage estimates from two-stage least squares regression models are shown in Table 4. These estimates tell us how well these two instruments (distance to nearest institution and its match with the type of substantive interest) predict first enrollment in each institution type, relative to all other possible choices, including no postsecondary enrollment.

What might we logically anticipate in terms of the strength of these instruments? Competitive colleges are widely known as such and also are access-restricted by definition, meaning that geographic proximity may play a limited role in enrollment in these institutions. In contrast, high-success institutions are a newer and still-obscure designation that did not exist when the ELS:2002 cohort finished high school, nor are all of these institutions highly competitive. Thus, we might expect geographic proximity to be a stronger predictor of enrollment in this category. Finally, we might expect geographic proximity to an HBCU to exert a draw toward HBCUs for Black students in particular, to a greater extent than for the full analytic sample.

The first-stage estimates are consistent with logial predictions. Closeness predicts first-institution type in all columns, and having the closest institution match the focal type of interest strongly raises the probability of enrolling in that institutional type among the full sample. In the sample of underrepresented minority students, having the nearest institution be an HBCU raises the probability of enrolling in an HBCU by about 13 percentage points, but having the nearest institution be one of the other two institution types is not a predictive factor.

If we are to have confidence that our instrumental variable estimates precisely capture randomly generated variation in first-institution type rather than noise (leading to a biased second stage), we need a strong first stage. Based on Stock et al. (2002), to keep relative bias less than 10% of what we would anticipate with OLS, we prefer for the joint first-stage F-statistic to be at least 9 with up to three instruments, rising toward a threshold of 12 with fifteen or more instruments (see Stock et al. (2002), p. 522, Table 1). In this case, because I jointly employ two instruments, I use 9 as the target threshold of instrument strength. The models that meet this threshold are those predicting first enrollment in a high-success institution in the full sample, with an F-statistic of 15.37 on the excluded instruments, and predicting first enrollment in an HBCU in the sample of underrepresented minority students, with an F-statistic of 9.69. As such, these are the models in which the second-stage estimates appear robust to weak instrument bias.

4.3 Instrumented Institution-Type Effects on Courses, Attainment, and Earnings

Table 5 presents IV estimates for the full sample in the top panel and for the subsample of underrepresented minority students in the lower panel. The IV estimates capture only those effects that are driven by variation in students' geographic proximity to institutions and institution types in grade 10. They are plausibly causal estimates for "compliers," meaning for students whose initial enrollment in a given institution type is influenced by living closest to an institution of that type, along with proximity to the nearest institution. In light of

Table 4: First-Stage Estimates of Instrument Effects on Category of Initial Institution

	Full Sample			**Underrepresented Minority**		
	(1)	(2)	(3)	(4)	(5)	(6)
	High-Success	Compet.	HBCU	High-Success	Compet.	HBCU
Miles to Nearest	-0.001***	-0.001**	-0.000**	-0.001*	-0.001+	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Nearest is Column Category	0.068***	0.042*	0.121**	0.020	0.026	0.133***
	(0.018)	(0.020)	(0.039)	(0.021)	(0.025)	(0.036)
F-instrum (2)	15.37	7.532	5.547	3.22	2.08	9.69
F-all	97.71	117	5.863	25.93	27.06	5.453
R-squared	0.265	0.309	0.061	0.238	0.244	0.044
Observations	9,053	9,053	8,870	2,822	2,822	2,744
Schools	703	703	697	599	599	595

Cluster robust standard errors in parentheses
*** p<0.001, ** p<0.01, * p<0.05, + p<0.1
Controls include race/ethnicity where possible, gender, family SES composite, first language not English, honors-weighted GPA in high school, 10th grade math test score, subsidized meal share in baseline high school, share of students in AP or IB classes in baseline high school, share of employers in professional sector in baseline residential zip code, median zip code income and average zip code years of education. They also include ratio of nearest-institution tuition to that of the state, and share of four-year institutions in the state.

the finding in Table 4 that the first stage is reasonably strong for high-success institutions in the full sample and for HBCUs in the sample of underrepresented minority students, we should focus particular attention on these estimates.

Focusing on the full analytic sample, I find that, among compliers, first attending a high-success institution is linked to a substantial earnings advantage at roughly age 26 of about \$8.66 per hour, or 0.87 SD. The comparison condition here is attending any other institution type *or* never attending college. Omitting those who never attended yields a similar but slightly smaller estimate of about 0.65 SD. This earnings advantage does not seem to be mediated by taking more STEM credits, as there appears to be no instrumented effect on STEM credits. In contrast, there is a positive instrumented effect on humanities course completion of 35 credits, roughly equivalent to a major. This is smaller than the marginally significant estimate of 42 humanities credits in competitive colleges, though again, competitive college effects are more vulnerable to exaggeration by weak instruments. Beginning college at a high-success institution is also linked to a marginally significant 33 percentage-point increase in the probabilty of bachelor's degree completion, but not to an increase in postbaccalaureate pursuits. It is possible that high-success institutions lead to higher early-career earnings in part due to graduates entering the workforce more quickly than those of other types of competitive institutions.

Turning to the subsample of underrepresented minority students, we find compelling and well-instrumented effects for first-enrollment at an HBCU. These include a marked increases in the number of humanities and STEM credits completed (39 and 28, respectively), as well as a marginally significant increase in the probabilty of bachelor's completion fo 48 percentage points ($p<0.1$). The positive effect on STEM credit completion is the only significant and positive estimate in the IV analysis and is consistent with prior findings suggesting that HBCUs provide enhanced support for Black students in STEM majors Perna et al. (2009); Kim and Conrad (2006).

Of concern, however, is the finding that first-attending an HBCU is linked to lower earnings, by just over \$6 per hour or 0.6 SD, in wages at age 26. Though the finding is only marginally significant ($p<0.1$), it is quite robust to other IV model specifications, including the use of more and fewer geographic instruments and the omission of individuals who never attended college. Thus, for underrepresented students geographically induced to attend HBCUs, I observe a paradox, which includes positive effects on STEM as well as humanities credit completion and bachelor's degree completion, but negative effects on earnings. This finding suggests that employer bias against credentials from HBCUs may be a factor, as I discuss further below.

As noted, the IV estimates are robust in direction and magnitude to omitting distance in miles as an instrument. Estimates (not shown) are similar in relative size but slightly more muted if individuals who never attended college are omitted from the comparison group.

Notably, some coefficients in the IV analysis exceed those from the OLS analysis, such as the earnings estimates for high-success institutions. If we presume that OLS estimates are biased away from 0 by unobserved ability and motivation, then this finding is counterintuitive. On the other hand, Card (2001) observe that numerous studies of returns to postsecondary education have found IV estimates that exceed OLS estimates. They note that a possible explanation is that benefits may be larger for the subset of students who are particularly sensitive to quasi-random facilitators of education access, such as (in this case) geographic proximity. This suggests that students on the margin of choosing a particular institution type may be especially sensitive to institutional effects on coursetaking, attainment, and earnings, whereas students' whose decisions are less geographically dependent may also be less susceptible to the effects of other institutional factors on their educational choices.

Table 5: Instrumented Effects of Institution Attributes on Outcomes

	(1) Humanities	(2) STEM	(3) Bachelor	(4) Postbac	(5) Wage
Full Sample					
High-Success Inst.	34.954*	-12.513	0.333+	0.159	8.666*
	(14.521)	(12.835)	(0.174)	(0.119)	(4.147)
Competitive Inst.	41.940+	-9.669	0.448	0.460*	7.143
	(25.066)	(21.912)	(0.300)	(0.203)	(5.791)
HBCU	23.187	18.076	0.275	0.119	-0.527
	(21.418)	(14.685)	(0.273)	(0.139)	(5.564)
Observations	8,957	8,957	8,942	9,524	9,053
Schools	702	702	702	707	703
Underrepresented Minority Students					
High-Success Inst.	63.419	-32.864	0.632	0.156	9.658
	(49.012)	(41.580)	(0.605)	(0.312)	(11.246)
Competitive Inst.	33.464	-5.919	0.794	-0.049	11.120
	(57.110)	(51.323)	(0.839)	(0.384)	(13.776)
HBCU	39.004+	27.543*	0.483+	0.144	-6.085+
	(22.731)	(13.526)	(0.274)	(0.128)	(3.449)
Observations	2,739	2,739	2,732	2,987	2,822
Schools	598	598	598	611	599

Cluster robust standard errors in parentheses

*** p<0.001, ** p<0.01, * p<0.05, + p<0.1

Controls include race/ethnicity where possible, gender, family SES composite, first language not English, honors-weighted GPA in high school, 10th grade math test score, subsidized meal share in baseline high school, share of students in AP or IB classes in baseline high school, share of employers in professional sector in baseline residential zip code, median zip code income and average zip code years of education. They also include ratio of nearest-institution tuition to that of the state, and share of four-year institutions in the state.

5 Discussion

As the U.S. economy has shifted from industrial to knowledge-and-service-based, returns to postsecondary education have risen (Goldin and Katz, 2008), and U.S. education policy has increasingly emphasized the aim of postsecondary education for all (Carnevale and Strohl, 2010; Oreopoulos and Petronijevic, 2019; Perna, 2015). But postsecondary education is not truly dichotomous. Student outcomes vary substantially as a function of education level, institution, and program of study (Dale and Krueger, 2014; Webber, 2014). For young people wishing to make wise postsecondary investments, it would be useful to know the attributes of institutions that lead to higher earnings, all else being equal. It would also be useful for educators at the secondary and postsecondary levels to understand the mechanisms by which some colleges yield higher earnings than others. To what extent do these institutions promote coursetaking in higher versus lower paying fields, or simply higher completion of bachelor's degrees and beyond?

Using longitudinal, student-level data, I estimate institutional category effects on the coursetaking, attainment, and earnings of individuals who have reached their mid-20s. These first-institution category effects may be considered causal to the extent that the type of institution nearest to the student's residence in grade 10 affects their first institution choice and, conditional on family background covariates, is not systematically chosen by families with particular institutional or career preferences. As highlighted by Card (2001), an important question in any IV analysis based on geographic proximity is whether such proximity is actually an exogenous predictor of first postsecondary institution. Hillman (2016) finds that geographic access to colleges is not equally distributed in terms of student-level variables, like ethnicity and socioeconomic status, though Rouse (1995) and Card (1999) find that it is plausibly exogenous, conditional on families' socioeconomic status.

Even if IV estimates do not purge all unobserved geographic preferences by families, they can still be used in juxtaposition with OLS to remove any bias that is not associated with nearest-institution attributes, conditional on all other controls. The rationale is that even if geographic proximity instruments are not perfect, they do leverage the ways in which accidents of geography shape people's—especially children's—choice sets and sense of possibility, independent of the many other accidents of birth that the models can adjust for, such as those linked to race, socioeconomic status, community affluence and education levels, and individual academic proclivity.

This paper complements the college value-added analyses of Rothwell (2015) and the college mobility scorecard work of Chetty et al. (2017) to further unpack the role that institutional attributes play in raising students' attainment levels and earnings capacity, including the role that coursework in STEM and humanities fields may play in mediating institutional effects on earnings. The current study differs from that of Rothwell (2015) in that he examines college scorecard data to identify average earnings differentials and their correlation with institution-level attributes, such as share of STEM majors. This study builds on the the work of Chetty et al. (2017) by examining the effects of the institutions their team finds to have top-quartile success rates on the actual coursetaking, attainment, and earnings of individuals geographically induced to attend them. This upacking effort is important if we take seriously the question of what colleges can do to raise students' human capital.

Though my findings align with the Chetty et al. (2017) success rates, corroborating the role of high-success institutions in promoting higher earnings, they leave open the question of what these institutions are doing differently to facilitate higher earnings among early-career graduates. For instance, these institutions appear to promote higher course-taking in humanities rather than STEM fields, a finding at odds with their effects on higher pay. On the other hand, they do appear to promote bachelor's degree completion and not postbaccalaureate education, where the former would indeed be expected to raise wages, and where the latter could delay the ability to reap earnings benefits from one's degrees. In considering the effects of high-success institutions along the null earnings estimates from competitive colleges, the fact that the dataset ends at roughly age 26 could play a role. For instance, Chetty et al. (2017) show that earnings for students from most four-year and two year colleges stabilize by age 25, but that average earnings rise until about age 30 for students at Ivy League and Barron's tier 1 colleges. This is because students from these institutions are especially likely to pursue postbaccalaureate education.

Importantly, the findings suggest a marked undervaluing of HBCU credentials by the labor market, despite evidence that these institutions substantially raise STEM coursetaking and postsecondary attainment levels for students of underrepresented backgrounds. The reason that first attendance at an HBCU might negatively affect wages despite having positively affected STEM course-taking and bachelor's attainment is not clear. One possibility is that employers are systematically biased against credits and credentials from HBCUs, meaning that the higher rates of STEM coursework and attainment that students acquire as a result of attending an HBCU are undervalued by the labor market. The systematic bias explanation is all the more plausible given that respondents' academic skills in high school (mathematics scale scores and grade point averages) are held constant in all of the statistical models.

In summary, this paper provides new evidence from instrumental variable estimation to support the notion that what Chetty et al. (2017) call high-success institutions do indeed play a causal role in enhancing the early-career earnings of those who enroll in them, relative to the choice of a different institution or no institution. I do not find similar evidence that traditionally competitive colleges—those in the top three tiers of the Barron's selectivity index—enhance early-career earnings. This suggests that there may be something distinctive about the high-success colleges, and it does not appear to be related to the completion of more STEM coursework, though greater inducement to complete a bachelor's degree may play a role. What this means for families is that attending a college traditionally classified as competitive may not be the career-optimizing choice for many young people, especially for young people who prefer to stay close to home for college, and that other as-yet-unidentified factors related to college completion or skill development may play a key role. Our findings also corroborate smaller studies suggesting a causal effect of HBCU attendance on STEM coursework completion and bachelor's degree attainment among underrepresented minority students, even if these advantages are not manifested as earnings advantages early in the career. The lesson here is that the attributes that make an effective college environment may depend on the needs of individual students. At the same time, researchers should continue aiming to illuminate these attributes on average so that institutions may adjust their own best practices and so that young people benefit from clearer guidance in facing the array of education choices before them.

Acknowledgements

[redacted for peer review]

References

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, Princeton, NJ.

Canaan, S. and Mouganie, P. (2018). Returns to education quality for low-skilled students: Evidence from a discontinuity. *Journal of Labor Economics*, 36(2):395–436.

Card, D. (1999). The causal effect of education on earnings. In Ashenfelter, O. C. and Card, D., editors, *Handbook of Labor Economics*, volume 3, chapter 30, pages 1801–1863. Elsevier.

Card, D. (2001). Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems. *Econometrica*, 69(5):1127–1160.

Carnevale, A. P. and Strohl, J. (2010). How increasing college access is increasing inequality, and what to do about it. In Kahlenberg, R. D., editor, *Rewarding strivers: Helping low-income students succeed in college*, chapter 3, pages 71–183.

Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2017). Mobility report cards: The role of colleges in intergenerational mobility. Technical report, Cambridge, MA.

Dale, S. B. and Krueger, A. B. (2002). Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *Quarterly Journal of Economics*, 117(4):1491–1527.

Dale, S. B. and Krueger, A. B. (2014). Estimating the effects of college characteristics over the career using administrative earnings data. *Journal of Human Resources*, 49(2):323–358.

De Zeeuw, M., Fazili, S., and Hotchkiss, J. L. (2020). Decomposing Outcome Differences between HBCU and Non-HBCU Institutions. Technical report, Federal Reserve Bank of Atlanta Working Paper Series, Atlanta, GA.

Franke, R. and Deangelo, L. (2018). The Power of HBCUs : A Propensity Score Matching Approach Examining Persistence and Degree Completion for Black Students at PWIs and HBCUs. (April).

Fryer, R. G. J. and Greenstone, M. (2010). The changing consequences of attending Historically Black Colleges and Universities. *American Economic Journal: Applied Economics*, 2(1):116–148.

Goldin, C. and Katz, L. F. (2008). *The race between education and technology.* Belknap Press, Cambridge, MA.

Herbst, C. M. and Tekin, E. (2016). The Impact of Child-Care Subsidies on Child Development: Evidence from Geographic Variation in the Distance to Social Service Agencies. *Journal of Policy Analysis and Management*, 35(1):94–116.

Hillman, N. W. (2016). Geography of college opportunity: The case of education deserts. *American Educational Research Journal*, 53(4):987–1021.

Ingels, S. J., Pratt, D. J., Rogers, J. E., Siegel, P. H., Stutts, E. S., and Owings, J. A. (2004). Education Longitudinal Study of 2002: Base year data file user's manual. Technical report, Washington, DC.

Jaggars, S. and Xu, D. (2013). Examining the effectiveness of online learning within a community college system: An instrumental variable approach (CCRC Working Paper 56). Technical report, New York.

Kim, M. M. and Conrad, C. F. (2006). The impact of historically Black colleges and universities on the academic success of African-American students. *Research in Higher Education*, 47(4):399–427.

Long, M. C. (2008). College quality and early adult outcomes. *Economics of Education Review*, 27(5):588–602.

Oreopoulos, P. and Petronijevic, U. (2019). The remarkable unresponsiveness of college students to nudging and what we can learn from it. Technical report, Cambridge, MA.

Palmer, R. and Gasman, M. (2008). "It takes a village to raise a child": The role of social capital in promoting academic success for African American men at a Black college. *Journal of College Student Development*, 49(1):52–70.

Perna, L., Lundy-Wagner, V., Drezner, N. D., Gasman, M., Yoon, S., Bose, E., and Gary, S. (2009). The contribution of HBCUS to the preparation of African American women for STEM careers: A case study. *Research in Higher Education*, 50(1):1–23.

Perna, L. W. (2015). Improving college access and completion for low-income and first-generation students: The role of college access and success programs.

Price, G. N., Spriggs, W., and Swinton, O. H. (2011). The Relative Returns to Graduating from a Historically Black College/University: Propensity Score Matching Estimates from the National Survey of Black Americans:. *https://doi.org/10.1007/s12114-011-9088-0*, 38(2):103–130.

Rothwell, J. (2015). Using earnings data to rank colleges: A value-added approach updated with College Scorecard data. Technical report, Brookings Institution, Washington, DC.

Rouse, C. E. (1995). Democratization or diversion? The effect of community colleges on educational attainment. *Journal of Business & Economic Statistics*, 13(2):217–224.

Schmitt, C. M. (2016). Documentation for the restricted-use NCES Barron's admissions competitiveness index data files: 1972, 1982, 1992, 2004, 2008, and 2014. Technical report, National Center for Education Statistics, Washington, DC.

Small, D. (2008). The strength of instrumental variables and their sensitivity to unobserved biases.

Solon, G., Haider, S. J., and Wooldridge, J. (2015). What are we weighting for? *Journal of Human Resources*, 50(2):301–316.

Steiner, P. M., Cook, T. D., and Shadish, W. R. (2011). On the importance of reliable covariate measurement in selection bias adjustments using propensity scores. *Journal of Educational and Behavioral Statistics*.

Stock, J. H., Wright, J. H., and Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments.

U.S. Department of Education (2010). Academic Competitiveness and National SMART Grants.

Webber, D. A. (2014). The lifetime earnings premia of different majors: Correcting for selection based on cognitive, noncognitive, and unobserved factors. *Labour Economics*, 28:14–23.